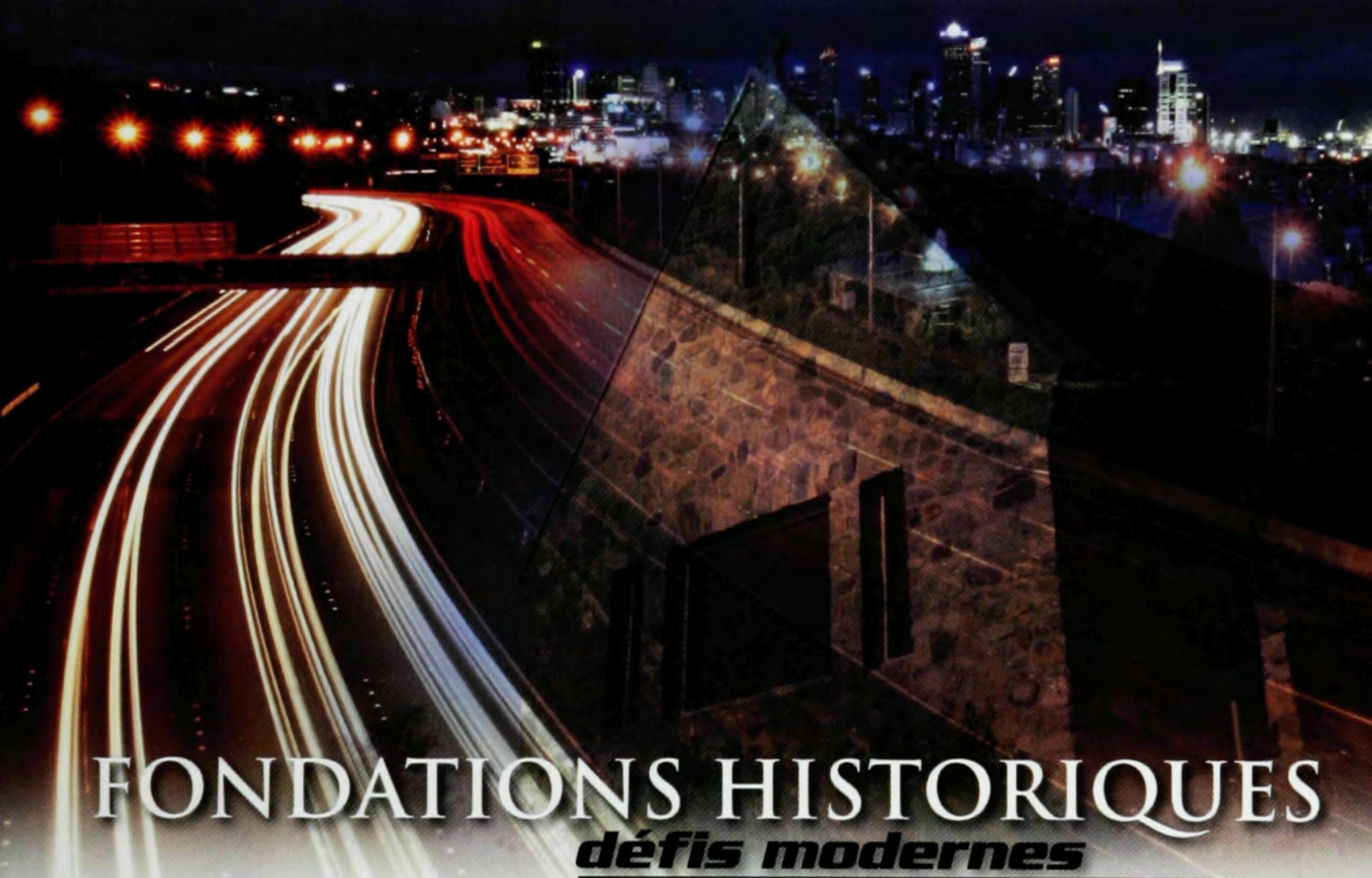
CANADIAN TRANSPORTATION RESEARCH FORUM | LE GROUPE DE RECHERCHES SUR LES TRANSPORTS AU CANADA

OLD FOUNDATIONS

modern challenges



FONDATIONS HISTORIQUES

défis modernes

Hamilton, Ontario • May 8 -11 mai • 2005

Cover Design By: University of Saskatchewan Printing Services

The Canadian Transportation Research Forum assumes no responsibility for the content of the papers published in these Proceedings or for any errors or omissions in the printing of this volume. These papers are published by the CTRF as a means of disseminating transportation research and to encourage discussion on a broad range of issues during the CTRF Annual Conference.

ISBN: 0-9737440-0-6
ISSN: 1183-2770

No part of this book may be reproduced in any manner whatsoever for commercial purposes. For information write to CTRF, 209-15 Innovation Blvd., Saskatoon, Saskatchewan, S7N 2XB. www.ctrf.ca

Proceedings Printed by
University of Saskatchewan Printing Services

TABLE OF CONTENTS

The 40th Annual Meeting of the CTRF
Actes de la 40è Conférence annuelle du GRTC
Hamilton, Ontario, Canada
May 8 -11mai 2005

VOLUME I

PRESIDENT'S MESSAGE / MESSAGE DU PRÉSIDENT

xi

I. Transportation Planning/Modelling - 1

1. *Modeling the location choice behaviour of retail trade business establishments in the City of Hamilton, Canada*
Hanna Maoh and Pavlos Kanaroglou 1

2. *Examining impacts of transportation on residential property values using geographical information system: a hedonic price model approach*
Muhammad Ahsanul Habib, Farhana Ferdous and K. M. Maniruzzaman 14

3. *Land use and capital costs at border crossings: Pacific Highway*
Philip Davies 29

II. Trucking

1. *"Too many trucks on the Road"?*
Gordon Baldwin 44

2. *Un portrait du marché du travail des conducteurs dans le*

secteur du transport routier des marchandises au Québec
Yvon Bigras and Guy Normandeau 55

3. *Encore sur la route: profil des camionneurs au Canada*
Vincent Dubé and Denis Pilon 70

III. Transportation Safety

1. *Comparing highway speed consistency measures using analytical method*
Young-Jin Park 85

2. *Estimation of vehicle-to-vehicle accident risk at signalized intersections*
Dewan Masud Karim 100

3. *On a necessity of sensor redundancy for safety integrity improvement in GNSS/INS based train location*
Lubor Bazant and Hyneck Mocek 115

IV. Trucking and Highway Issues

1. *The relationship of collision rates and dynamic performance of heavy trucks*
Alfonso Corredor, Karl Groskopf and Ron Madill 130

2. *A framework for winter road maintenance optimisation*
Mathieu Trudel 144

V. Transportation and the Economy

1. *The role of transportation in trade competitiveness: understanding the policy context through quantitative measures*
Nilam Bedi and Atif Kubursi 158

2. *Sourcing of transportation inputs: comparative tire pricing on the Internet and other channels*
Louis A. LeBlanc 172

Short Sea Shipping Panel

1. *Short sea shipping and the supply chain: a review of cross lake ferry economics and benefits*
Christopher Wright 186

VI. Transportation and the Elderly

1. *Transportation implications of Canada's aging population*
Darren M. Scott, K. Bruce Newbold, Antonio Páez and Pavlos S. Kanaroglou 201

2. *Transport policy in an ageing society: an international comparison*
Ruben Mercado, Antonio Páez, K. Bruce Newbold, Darren M. Scott and Pavlos S. Kanaroglou 209

3. *Canada's elderly population travel behaviour: experience from the Hamilton Census Metropolitan Area*
Dimitris Potoglou, Antonio Páez, Darren M. Scott, Pavlos S. Kanaroglou and K. Bruce Newbold 224

4. *A GIS-based decision support tool to study the impact of elderly population on the transportation system in Hamilton Census Metropolitan Area (CMA)*
Hanna Maoh, Pavlos Kanaroglou, Darren M. Scott, K. Bruce Newbold and Antonio Páez 239

VII. Air - 1

1. *Air transportation - are new destinations on the horizon?*
Joseph Monteiro and David Krause 253

2. *Computer reservation systems in Canada - changing regulations and their economic rationale*
Joseph Monteiro and David Krause 270

3. *Capstone Avionics in Alaska: an economic appraisal*
Darren Prokop 286

VIII. Transportation and the Environment

1. *Purolator Courier's "greening the fleet" initiative*
Douglas J. Kube 311

2. *A Kyoto Protocol update: opportunities and challenges for Canadian freight rail*
Malcolm Cairns 323

3. *Climate change, productivity shocks and demand for freight ground transportation in Atlantic Canada*
Yuri V. Yevdokimov and Oleg Byelyayev 341

IX. Transportation Costs

1. *Challenges in measuring the external costs of transport*
A. E. Boardman, David Gillen, W.G. Waters II and Anming Zhang 356

2. *Measuring congestion costs*
W. G. Waters II and Markus von Wartburg 371

3. *Estimating traffic delay costs caused by freeway crashes*

Jaehyun H. Noh, Frank Saccomanno and Bruce Hellinga 385

X. Marine

1. *The Canadian shipbuilding industry, world market distorting practices and mechanisms to deal with it*
Joseph Monteiro, Gerald Robertson and Keith Dawson 400

2. *Midstream opportunities and challenges for Canada's west coast ports*
Darryl Anderson 417

3. *Canadian-US port competition and cargo diversion*
Joseph Monteiro 432

XI. Transportation Financing

1. *The effects of railway investments in a polycentric city*
Ghebreegziabiher Debrezion, Eric Pels and Piet Rietveld 449

2. *Auslink: a new approach to intergovernmental financial arrangements for land transport funding*
Raluca Raicu, Derek Scrafton and Michael Taylor 465

XII. Air - 2

1. *Cargo airships: civilian applications and national defence*
Barry Prentice 480

2. *Some second thoughts on the governance of Canadian airports*
Michel Boucher 495

3. *Airport congestion and market structure: the case of*

variable passenger time costs
Anming Zhang and Andrew Yuen 516

Geographic Inforamtion Systems Panel

1. *National Road Network (NRN)*
(Canada's digital road centreline framework data)
Marcel Sabourin 540

2. *Reviving the activity space: can "old" behavioural geography provide "new" insights to the urban transport policy debate*
Ron N. Buliung 555

3. *Developing a GIS-based detailed traffic sumulation model for the Montreal region: opportunities and challenges*
Timothy Spurr and Murtaza Haider 571

XIII. Trade and Border Issues

1. *The importance of security in the service offerings of Canada's customs brokers*
Michael A. Haughton 585

2. *The Eurasian land bridge - an economic perspective of a new transportation corridor between Europe and China*
Barry E. Prentice and L. Fan 597

XIV. Accessible Transportation and Developing Issues

1. *Determining relationships between regional characteristics and paratransit demand*
Peter Allaby 611

2. *Public transportation: a solution to social exclusion*
Sean Nix 625

3. *Developing media partnerships for crisis communications during transportation disasters*
Jill A. Lynn 634

XV. Transportation Safety and the Environment

1. *Potential demand for household alternative fuelled vehicles in Hamilton, Canada: a stated choices experiment and survey*
Dimitris Potoglou and Pavlos S. Kanaroglou 647

2. *Hydrogen hybrid internal combustion engine bus demonstration*
Allister Hickson and Christy Sokol 662

3. *An analytical model for determination of road safety index*
Mahmoud Saffrzadeh, Maghsood Pooryari and Hamidreza Bahramian 673

XVI. Transportation Planning/Modelling - 2

1. *Travel demand model in railway transportation network, using the vector autocorrelation routine (VAR)*
Mahmoud Saffarzadech and Abdolreza Rezaee-Arjroody 687

2. *Ship traffic simulation in the port of a developing country: Evidence from Nigeria*
Mobolaji S. Stephens and Innocent C. Ogwude 703

3. *Impacts of road improvements on generated traffic*

congestion and emissions: case of the Hamilton CMA, Canada
Hejun Kang, Darren M. Scott, Pavlos S. Kanaroglou and Hanna Maoh 718

XVII. Urban Issues

1. *Keeping Peel's economy on the move - advancing the agenda for goods movement in Peel*
Murray Mcleod 733

2. *Urban transport issue and improvement options in Bangladesh*
M. Mazharul Hoque, Bidoura Khondaker and M. J. B Alam 748

CTRF 2003-2004 Student Paper Competition
Award Winning Papers

Undergraduate Category
First Prize – Past Presidents Award ($1000)
Aircraft Noise: A survey of the regulatory framework and Enforcement mechanisms in Canada
Micheal T. Kariya 763

Undergraduate Category
Third Prize ($250)
An urban governance model to achieve sustainable urban transportation and development: a case study of Vancouver
Jesse W.R. Coleman 778

AUTHOR INDEX **793**

APPENDIX I

STATEMENT OF PURPOSE	I-2
HISTORY OF CTRF	I-3
CTRF / GRTC - EXECUTIVE/EXECUTIF	I-4
CTRF PAST PRESIDENTS / PRÉSIDENTS ANTÉRIEURS DU GRTC	I-5
2005 ANNUAL CONFERENCE SPONSORS / COMMANDITAIRES DE LA CONFÉRENCE 2005	I-6
2004 SEMI-ANNUAL CONFERENCE SPONSORS / COMMANDITAIRES DE LA CONFÉRENCE 2004	I-6
SCHOLARSHIP PROGRAM AND AWARDS / PROGRAMME DE BOURSE	I-7
2003-2004 STUDENT PAPER COMPETITION DONORS / PARRAINS DU CONCOURS DE ÉTUDIANTS EN 2003-2004	I-8
MEMBERSHIP FORM / FACTURE DU MEMBRES	I-9
ANNOUNCEMENT / INFORMATION	I-11

PRESIDENT'S MESSAGE MESSAGE DU PRÉSIDENT



Gordon Baldwin, CTRF President

At this conference and with these proceedings, the Canadian Transportation Research Forum is celebrating its 40th anniversary. The Canadian transportation scene was very different 40 years ago. We had a country where:

- on January 1st 1965 Trans-Canada Air Lines became Air Canada;
- the Trans-Canada Highway was only 3 years old;
- the Canadian Transport Commission had not yet been established;
- the Minister of Transport was the Hon. John Whitney Pickersgill;
- this was 12 years before the creation of VIA Rail;
- in this year the US-Canada Auto pact was signed resulting in a huge trade in cars and car parts between the two

Au cours de cette réunion annuelle et dans ces actes de la conférence, le Groupe de recherches sur les transports au Canada fête ses 40 ans d'existence. La scène des transports au Canada était fort différente il y a 40 ans. En voici quelques exemples :

- Le 1er janvier 1965, Lignes aériennes Trans-Canada devenait Air Canada;
- La route transcanadienne existait depuis seulement trois ans;
- La Commission canadienne des transports n'avait pas encore vu le jour;
- Le ministre des Transports était l'honorable John Whitney Pickersgill;
- Via Rail allait être lancé douze ans plus tard;
- Le Pacte de l'automobile canado-américain était ratifié, donnant lieu à un d'importants mouvements d'automobiles et de pièces d'automobile entre les deux pays;

- the cost structure was very different, where here in Hamilton, the site of this conference, the Ontario minimum wage for women in southern Ontario was $0.95/hour and for men it was $1.00/hour (yes the minimum wage varied by sex);
- all of the above was being done under the new Canadian flag of red and white with the maple leaf and side bars.

On to this scene, Canadian transportation professionals, encouraged by the success of the American Transportation Research Forum (TRF) and with the financial assistance of the federal Department of Transport, planned the CTRF. On behalf of us all, I would like to thank that visionary group. The planning committee consisted of:

- Ray Cope - Department of Transport
- Bill Hughes - University of British Columbia
- Harvey Romoff - Canadian Pacific Railway
- Eric Gracey - Canadian Industrial Traffic League
- Peter Williams - National Harbours Board
- Claude Taylor - Air Canada
- Fred Woodward - Air Transport Board
- Benoit Savard - Canadian Trucking Association

- La structure des coûts était très différente, en effet, ici même à Hamilton, le salaire minimum était, dans le sud de la province, de 0,95 $ pour les femmes et de 1,00 $ pour les hommes (le salaire minimum variait effectivement en fonction du sexe);
- Tout cela se déroulait sous le nouveau drapeau canadien rouge et blanc qui arborait l'unifolié et deux barres latérales.

Dans notre contexte, les professionnels du transport au Canada, incités par le succès du *Transportation Research Forum (TRF)* aux États-Unis, et avec l'appui financier du ministère fédéral des Transports, ont mis sur pied le Groupe de recherches sur les transports au Canada. Au nom de tous ici présents, j'aimerais remercier ce groupe visionnaire. Le comité de planification était constitué de:

- Ray Cope - ministère des Transports
- Bill Hughes - Université de la Colombie-Britannique
- Harvey Romoff - Canadien Pacifique
- Eric Gracey - Ligue canadienne de transport industriel
- Peter Williams - Société canadienne des ports
- Claude Taylor - Air Canada
- Fred Woodward - Commission des transports aériens du Canada
- Benoit Savard - Association canadienne du camionage

_ Don MacKinnon - Canadian National Railways
_ Ron Wilson - Hawker Siddeley
_ Peter Sandiford - McGill University

Bob Bandeen, a veteran of the early development of the TRF and a past TRF President, provided experience, enthusiasm and direction to the new CTRF.

The first meeting of the CTRF was held at Manoir St. Castain in Lac Beauport, Quebec on September 9, 1965. That first year, the CTRF attracted 120 members to join the organization.

The organization has grown and evolved. The CTRF scholarship Fund was launched in 1988 to encourage university students to pursue studies leading to careers in transportation. The annual conference was joined by a one-day semi-annual conference.

The organization has prospered due to its strong mixture of industry, academic and government members, providing a unique forum to present, discuss, confront and learn about the leading transportation research.

_ Don MacKinnon - Chemins de fer nationaux du Canada
_ Ron Wilson - Hawker Siddeley
_ Peter Sandiford - Université McGill

Bob Bandeen, un pionnier et ancien président du *Transportation Research Forum*, a fourni expérience, enthousiasme et direction en vue de l'établissement du nouveau Groupe de recherches sur les transports au Canada.

La première rencontre du Groupe a eu lieu au Manoir St-Castain au Lac Beauport, Québec, le 9 septembre 1965. À sa première année d'existence, le Groupe a attiré 120 membres.

Depuis, l'organisme a prospéré et a évolué. En 1988 on lançait le fonds de bourse d'études du Groupe de recherches sur les transports au Canada afin d'inciter les étudiants universitaires à poursuivre des études qui mèneraient à une carrière dans le domaine des transports. Outre la conférence annuelle, on compte maintenant une conférence semestrielle d'un jour.

As the CTRF returns to its roots, with the annual conference planned for Quebec City in 2006, I encourage you all to invite a younger colleague, business associate or student to come help launch our next 40 years.

L'organisation, qui a pu s'épanouir grâce à la grande diversité de ses membres issus de milieux industriels, universitaires et gouvernementaux, offre une tribune unique où l'on peut présenter, discuter et débattre des recherches de pointe sur le transport tout en apprenant sur le sujet.

Alors que le Groupe de recherches sur les transports au Canada fait un retour aux sources dans le cadre de sa réunion annuelle de 2006, qui aura lieu à Québec, je vous incite à inviter un jeune collègue, associé ou étudiant à participer au lancement des 40 prochaines années.

Modeling the Location Choice Behavior of Retail Trade Business Establishments in the City of Hamilton, Canada

Hanna MAOH and Pavlos KANAROGLOU
Center for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University, 1280 Main Street West, Hamilton, ON, Canada, L8S 4K1

1. INTRODUCTION

In recent years, there has been a widespread recognition among transportation analysts that the configuration of urban land uses can determine a large portion of urban daily travel patterns. Integrated Urban Modeling Systems (IUMS) were developed for many cities around the world to model this process (Wegener, 2004). IUMS are planning tools used to evaluate the long-term repercussions of proposed land use and transport network changes. These include the differential estimation of land development, location decision of households and firms and the trips generated between the different parts of the city. These trips are then used to estimate congestion and associated tailpipe emissions in all the links of the urban transport network. The results achieved from simulating the urban system could enable decision makers to carry on, rectify or abandon certain aspects of their infrastructure projects.

Research to develop and improve the performance of IUMS received more attention in recent years in response to the environmental concerns raised by the different governments on the green house gas emissions problem. Research initiatives are underway to develop an agent-based microsimulation IUMS for the city of Hamilton, Canada. Part of these initiatives is to study the location choice behavior of business establishments at the micro-level.

This paper reports on the efforts we undertook to model the location choice behavior of retail trade business establishments, one of the fast growing economic sectors in Hamilton. Following Lee (1982) and Martinez (1992), each business establishment is assumed to face the decision of selecting a location that will maximize its profit. Accordingly, discrete choice theory is applied and a multinomial logit model is specified and estimated for the period 1996 - 1997. We offer a novel framework for utilizing different sources of datasets to study the location behavior of business establishments at the micro level. We also provide new empirical evidence of the factors influencing the location behavior of individual retail trade business establishments in the city of Hamilton in Ontario, Canada.

The remainder of this paper is organized as follows. Section two provides an overview on location choice modeling and highlights some of the significant factors used to explain the location behavior of firms. Section three discusses the methods of analysis used to formulate and specify the location choice models. It also highlights the different data sources used in the empirical research. Section four discusses the estimation results to inform the recent trends of retail trade location choice behavior in the City of Hamilton. The last section provides a conclusion to our study.

2. BACKGROUND

Employment location modeling has made remarkable advancements over the past four decades. The theoretical advances achieved by McFadden (1978) in discrete choice models and the success in applying them to residential location problems has paved the road for empirical research on firm location. Later, the work of Martinez (1992) provided a theoretical basis for using discrete choice theory in the firm location choice problem. He showed that bid-rent and profit maximization theories are equivalent and complement each other. Within the discrete choice-modeling framework, urban space is typically partitioned into a set of discrete zones that form a location choice set. Firm location behavior is modeled as the choice of zone that maximizes utility by way of maximizing profit. Utility, in firm

location studies that have made use of discrete choice models, is typically specified as a composite measure that includes both site and firm attributes (see for example Lee, 1982; Shukla and Waddell, 1991; Wu, 1999; Frenkel, 2001; Maoh et al., 2005).

Among site characteristics that are often used to explain the location of firms, distance from the Central Business District (CBD) appears to explain firm centralization or decentralization. Proximity to major roads and highways also appears in the urban economic literature as an important determinant of firm location choice. While highways and major roads remain an important accessibility factor for the flow of people and goods in cities, recent dependence on airline shipping has brought land in close proximity to airports into focus. Land use development practices and the establishment of regional malls in many North American cities has affected the way certain firms react to location in the city. Studies have shown that certain firms are attracted to regional malls or areas in close proximity to them. Agglomeration economies are central to urban processes. As such, their representation in firm location modeling is instrumental. Empirical studies have shown that agglomeration economies play a key role in firm location and that firms accrue numerous benefits from clustering and localizing in space.

Land use zoning is also a factor that draws firms to certain locations. Therefore, the type of a dominant land use could attract certain firms and encourage them to gravitate to a particular location. For example, Wu (1999) shows that foreign direct investment FDI firms are more likely to locate in the Economic and Technology development zone (ETDZ). The prestige level of a given zone is also a determinant of firm location. Wu (1999) and Frenkel (2001) show that regional prestige is a particularly important location factor for FDI and high-tech firms. Other location factors used in the literature include population size and income. Shukla and Waddell (1991) found that all types of firms show a tendency toward locating in populated areas. Their results also show that population income and spatial variation in population income affect where firms locate in cities.

Significant firm-specific factors used in formulating the utility function of location models include size, age and

organizational structure (e.g: single and multi plants). Waddell and Shukla (1993) found that large and medium size firms locate in one subcenter as compared to the CBD. Frenkel (2001) found that the central part of the metropolitan area holds strong attraction for small plants with high technological capabilities. In addition, he found that, 58% of all multiplants were located outside the metropolitan area; of these, 83% were young plants. The findings also show that the majority of multiplants (69%), located within the metropolitan area, were older plants.

3. METHODS OF ANALYSIS

3.1 Study area and data sources

The City of Hamilton in Ontario, Canada forms the geographic domain for our empirical analysis. The city is located south west of the Greater Toronto. The topography is shaped by the Niagara escarpment, which separates the city's geography into a lower coastal area between the escarpment and the harbor to the North and a highland beyond the escarpment to the south. While early urban development took place on the lower city, the continuous population and economic growth over time led to urban development above the escarpment. Recently, the city expanded its boundary and annexed the municipalities of Ancaster, Dundas, Flamborough, Glanbrook and Stoney Creek as shown in Figure 1.

The empirical analysis is based on three sources of data that were geo-coded to the generated grid cells. These are: (1) Statistics Canada Business Register (BR), (2) Statistics Canada Census data and (3) Desktop Mapping Technologies Incorporated (DMTI) GIS data. The BR was used to extract annual information on individual business establishments for the years 1996 – 1997. The extracted data included small and medium size (SME) retail trade establishments with less than 200 employees. Larger establishments with more than 199 employees were excluded from the analysis since they have a small number and their location behavior is predictable given the high monetary value associated with their infrastructure. SME establishments, on the other hand, account for the majority (99%) of business establishment population in Hamilton. They are also more

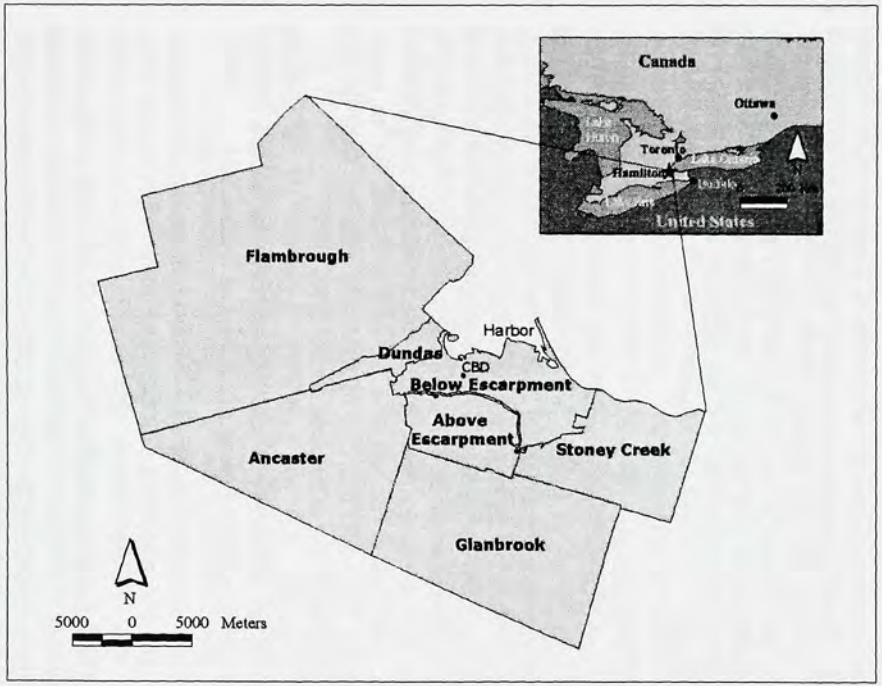


Figure 1: The City of Hamilton in the regional context

footloose and tend to relocate more frequently due to the lower cost required for their infrastructure.

The extracted SME data was constrained to the self-owned non-incorporated (NIP) establishments. This restriction was imposed since annual information is required to develop the agent-based model and is only available for this group of business establishments in the BR. The NIP population is deemed appropriate for the analysis since it accounts for the majority (86%) of SME establishments. However, the absence of the incorporated (IP) establishments from the extracted sample will undermine our understanding of what role organizational structure has on the location choice behavior of establishments.

3.2 Model formulation

The location choice models we seek to formulate will be used to estimate the probability of choosing a particular developed site in the city as a place for the firm to be established. Relocating, newly established and in-migrating retail trade firms are grouped into one pool. Each firm is assumed to be searching for an urban location that will maximize its profit. We divide the urban space into grid cells of 200 by 200 meters and then identify the developed land available to firms. This spatial resolution is deemed sufficient for an agent-based model. The result is a total of 2635 grid cells that form the sets of location alternatives in the time period 1996-1997.

Assuming that the grid cells form the set of location alternatives, the probability of selecting grid cell i as the new location of a firm can be estimated using a multinomial logit model:

$$P_n(i) = \frac{\exp(V_{ni})}{\sum_{j=1}^{J_n} \exp(V_{nj})}$$

Here V_{ni} is a linear-in-parameter function that depends on covariates characterizing firm n and the attributes of site i. Following McFadden (1978), the estimation of the logit model is based on a random sample of alternatives that is drawn from the choice set of grid cells. We use a uniform distribution to randomly sample a set of nine grid cells in addition to the chosen cell. The location choice problem is therefore

modeled among 10 alternative locations. The method overcomes the problem of dealing with a large choice set and produces consistent coefficient estimates that can be used to understand the location behavior of retail trade firms.

3.3 Model specification

The specification of the models involved using information from the BR, the population census and GIS layers to create variables at the grid cell level. GIS tools were employed to create a number of spatial measures to reflect the variables listed in Table 1. These were devised based on the information we gathered from the urban economic and urban geography literature and the exploration of economic clustering and urban form in the city of Hamilton (Maoh and Kanaroglou, 2004). Distance to CBD was initially included to examine how centrality influences the location of firms in the city of Hamilton. The variable was calculated as the Euclidean distance between the core and the centroid of each grid cell. However, we opt to replace the CBD variable with the SUBURB variable listed in table 1 to capture the effect of suburbanization, which appeared to be more influential than the decentralization effects.

Proximity to highways and main roads is used to test what influence land accessibility has on the location behavior of firms. The variable was based on intersecting grid cells with highways and main roads in the city. Developed land within 1500 meters from regional malls is specified since it is an attraction to certain types of firms.

Agglomeration economies are accounted for using a spatial measure that counts the number of firms of a particular sector within a 500 meters radius from the center of a given grid cell. The value of 500 meters radius was selected after experimentation with values at 1500, 1000 and 500 meters. The chosen radius provided more plausible and statistically significant results. Land use in grid cells is used as a control for the type of land development. It is used to examine if land specialization influences the location behavior in the city of Hamilton. Percentage of land use in a given grid cell was

Table 1: List of covariates used in the specification of the location choice models

Location Specific Factors	
CBD	Euclidian distance between the centroid of grid cell i and the centroid of the CBD
HWY	1 if grid cell i intersects with a main road or a highway, 0 otherwise
MALL	1 if grid cell i is within 1500 metes from a regional mall, 0 otherwise
AGGLOM	Number of retail trade business establishments within 500 meters from the center of grid cell i
SUBURB	1 if grid cell i is located in the suburban municipalities of the city of Hamilton, 0 otherwise
NEWDEV	Density of residential dwellings built in grid cell i between 1991 and 1996
OLDDEV	Density of residential dwellings built in grid cell i between 1981 and 1990
HHINCD	Density of average household income in grid cell i
LANDU1	Percentage of residential land use in grid cell i
LANDU2	Percentage of commercial land use in grid cell i
Firm Specific factors (IND_x variables are based on the 1980 Standard Industrial Classification Coding)	
IND1	1 if the firm belongs to SIC 60, 61 or 65, 0 otherwise
IND2	1 if the firm belongs to SIC 61, 62 or 63, 0 otherwise
IND3	1 if the firm belongs to SIC 64 or 65, 0 otherwise
IND4	1 if the firm belongs to SIC 601, 633, 635 or 641, 0 otherwise

calculated from intersecting the grid cell layer with the land use GIS cover. The procedure provided the land use classification of each grid cell and the associated land use area in the grid cell. Our tests led us to retain the commercial and residential land uses in the specification of the final model since other land use types turned out to be insignificant.

Due to the aggregate nature of the census population, a disaggregation was required to obtain information at the grid cell level. We use cross-area interpolation and kernel estimation to transform spatial population distribution from census centroid data into a continuous surface of population density. This surface is then superimposed on a grid to make it compatible with other GIS data. Using this method we generated a number of variables at the grid cell level that reflect the average household income density and the intensity of new developed dwellings in the time periods 1991 – 1996, as well as the intensity of old developed dwelling in the period 1981 – 1990. Finally we derived a number of categorical variables to represent the industry where a firm belongs. These were interacted with other variables to account for firm heterogeneity and to examine if location factors vary among the different firms.

4. ESTIMATION RESULTS

Estimation results for the retail trade location choice model in the period 1996–1997 are presented in Table 2. Since the coefficients are based on random sampling of alternatives, alternative specific constants were not specified in the models. Land accessibility in the model is measured by the *HWY* covariate. The estimated parameters of this variable suggest that accessibility plays a role in the location decision of the firms. The magnitude of the estimated parameters suggests that retail trade firms favor location on land in close proximity to main roads and highways. Land in close proximity to regional malls seems to attract retail trade firms that specialize in selling food, beverage and drugs (SIC 60), shoe, apparel, fabric and yarn (SIC 61), and general retailing stores (other retail store SIC 65). The specified agglomeration variable is highly significant and meets our a priori expectation with regard to its sign. The result suggests

that other things being equal, firms of the same type appreciate the benefits of agglomeration economies, leading to economic clustering in the city of Hamilton. Such results coincide with previous results obtained by the authors on the recent economic clustering trends in the City of Hamilton.

Table 2: Parameter estimates for the location choice model

Retail trade			
Covariate		*Beta*	*t-stats*
HWY		0.376	3.168
AGGLOM		0.022	6.829
SUBURB		0.426	2.294
NEWDEV*		3.345	3.189
OLDDEV*		-0.629	-2.734
HHINCD*		0.002	1.698
LANDU1		-0.560	-2.719
LANDU2		1.152	2.946
IND1 x MALL		0.650	2.890
IND2 x HHINCD*		-0.002	-1.535
IND3 x LANDU1		0.644	1.888
IND4 x SUBURBS		-0.425	-1.752
N	*L(0)*	*L(B)*	*Rho²*
369	-911.8	-812.9	0.1085

** Parameters are in thousands*

The *SUBURB* covariate was specified to capture urban sprawl in Hamilton. The estimated parameter of this covariate suggests a clear suburbanization trend among the retail trade firms. However, the interaction term IND4 x SUBURB indicates that, other things being equal, food stores (SIC 601), gasoline service stations (SIC 633), motor vehicle repair shops (SIC 635) and general merchandize stores (SIC 641) are still central and tend to avoid locating on land in the suburbs.

The variable *NEWDEV* reflects the intensity of recent residential development that took place in 1991–1996. On the other hand, *OLDDEV* reflects the intensity of previous residential development that took place in 1981–1990. The specification of these two variables suggests that retail trade firms favor land in close proximity to new residential development. These results complement our findings on the suburbanization trends, since most of the recent residential development is happening in the suburbs. We also believe that new residential development is a catalyst for commercial development. Retail trade firms gravitate towards newly formed residential nodes to benefit from proximity to these populated areas. The results also indicate that retail trade firms avoid areas with old residential development, given that these are more saturated compared to newly developed areas.

Results related to the density of household income suggest that retail trade firms favor richer areas in the city. However, the interaction term *IND2* x *HHINCD* suggests that, other things being equal, shoe, apparel, fabric and yarn retailers (SIC 61), household furniture, appliances and furnishing retailers (SIC 62) and automotive vehicles, parts and accessories, sales and service (SIC 63) do not prefer high income areas when they make their location decisions.

The land use variables were specified to control for the effect of land use type within a given grid cell. Not surprisingly, retail trade firms show a strong affiliation with commercial land use, as suggested by the *LANDU2* parameter. Also the interaction term *IND3* x *LANDU1* suggests that population oriented retail firms, such as general merchandize stores (SIC 64) and other retail stores (SIC 65), show affiliation with grid cells characterized by residential land use. However, other retailers show no affiliation with areas that are mainly residential, as suggested by the negative sign of the *LANDU1* parameter.

5. CONCLUSION

This paper presented a micro-analytical location choice model for retail trade firms in the city of Hamilton in Canada. The model will be used within an agent-based firm demographic model to

simulate the life trajectory of individual establishments over time and space. The research presented here offers a new way to approach the firm location problem at the micro level. Using spatial analysis and GIS, it was possible to develop critical variables that enabled us to extend previous firm location modeling techniques and adapt them for the use in microsimulation models. This research also highlights the potential of employing the BR data in the development of urban firm demographic agent-based models.

The results highlight the factors influencing the location choice behavior of retail trade firms in the city. They also emphasize the decentralization and suburbanization processes that have manifested themselves in the city of Hamilton in recent years. The proposed approach sets a good example for modeling the firm location choice problem at the micro-level. Similar effort, therefore, could be undertaken to investigate the factors affecting the location choice behavior of firms from other sectors of the economy such as manufacturing, wholesale trade and services firms.

REFERENCES

Frenkel, A. (2001). Why high technology firms choose to locate in or near metropolitan areas. *Urban Studies* 38, 1083--1102.

Lee, K. (1982). A model of intraurban employment location: an application to Bogota, Colombia. *Journal of Urban Economics* 12, 263-279.

Maoh, H. and P. Kanaroglou (2004). Economic clustering and urban form: the case of Hamilton, Ontario. Working Paper No. Working Paper No. 3, Center for Spatial Analysis, School of Geography and Geology, McMaster University,

Maoh, H. and P. Kanaroglou (2005). Modeling the Location of Firms within an Integrated Transport and Land-use Model for Hamilton, Ontario. Working Paper No. Working Paper No. 6, Center for Spatial Analysis, School of Geography and Geology, McMaster University,

Martinez, F. (1992). The bid-choice land use model: an integrated economic framework. *Environment and Planning A* 24, 871-885.

McFadden, D. (1978). Modeling the choice of residential location. In *Spatial Interaction Theory and Planning Models,* ed.A. Karlqvist, L. Lundqvist, F. Snickars and J. Wiebull, pp. 75-96. Amsterdam: North Holland Amsterdam.

Shukla, V. and P. Waddell (1991). Firm location and land use in discrete urban space: A study of the spatial structure of Dallas-Fort Worth. *Regional Science and Urban Economics* 21, 225-253.

Waddell, P. and V. Shukla (1993). Manufacturing location in a polycentric urban area: a study in the composition and attractivness of employment subcenters. *Urban Geography* 14, 277-296.

Wegener, M. (2004). Overview of land use transport models. In *Handbook of transport geography and spatial systems,* ed.D. Hensher, K. Button, K. Haynes and P. Stopher, pp. 127-146. Amsterdam: Elsevier.

Wu, F. (1999). Intrametropolitan FDI firm location in Guangzhou, China: a poisson and negative binomial analysis. *The Annals of Regional Science* 33, 535-555.

Examining Impacts of Transportation on Residential Property Values using Geographical Information System: A Hedonic Price Model Approach

Muhammad Ahsanul Habib
Graduate Student, Department of Civil Engineering
University of Toronto, Toronto, ON, M5S 1A4
Lecturer (On leave), Department of Urban and Regional Planning
Bangladesh University of Engineering and Technology, Dhaka-1000
Farhana Ferdous
Technical Officer, CARE Bangladesh
Pragati RPR Center, 20-21 Kawran Bazar, Dhaka- 1215, Bangladesh
K. M. Maniruzzaman
Associate Professor, Department of Urban and Regional Planning
Bangladesh University of Engineering and Technology, Dhaka-1000

ABSTRACT

Most hedonic price model studies focused on the impact of transportation system on land uses, particularly on land values. This study was an attempt following hedonic price modeling approach to provide empirical evidence regarding the nature and magnitude of accessibility impacts of transportation system on asking rental price of residential properties. It also investigated and compared impacts of structural attributes and neighborhood attributes on residential property values with that of transportation system.

Both linear and semi-log model specifications were estimated based on data collected by field visual inspection of advertisements for rent during May 2004 in the city of Rajshahi, Bangladesh. The study used Geographical Information System (GIS) as a tool for the

estimation of hedonic models. It was facilitated by a large GIS dataset collected by the Rajshahi Master Plan Project.

In the study, stepwise regression technique was used to determine best-reduced model. Moran's *I* and Geary's *c* were estimated to identify the effects of spatial autocorrelation in the residuals of predicted property values. The final model explained 58.8 percent of the variability in the residential asking rental prices. Number of bedrooms proved to be the highest contributor as a predictor accounting for a standardized coefficient of 0.427.

One of the most interesting findings was that increase in network distance to CBD from the residential properties increases property values in the study area. However, the estimated hedonic model suggested that moving 0.5-mile closer to the major arterials adds 585.858 Taka (CAD $13) premiums, all else held equal, on residential asking rental price at Rajshahi City.

1. Introduction

This study attempted to estimate hedonic price models to explore determinant factors on residential property values. With the hedonic price model technique, observable behavior of the real estate market is used to derive effects of housing attributes that are implicit in this behavior. Residential units are treated as multi-attribute goods. Although these attributes are not explicitly exchanged in observable market transactions, their value is inferred by the willingness to pay for each configuration of property attributes/characteristics. These attributes can include not only structural and site attributes, but also measures of local services and neighborhood quality, and most importantly, the accessibility (Rodriguez and Targa, 2004). In this study, three distinct categories of residential property attributes were considered: structural attributes, neighborhood attributes and transportation attributes (accessibility).

2. Brief literature review

Based on land rent theory (Alonso, 1964; Muth, 1969) and urban economics it is hypothesized that transportation improvements enhance accessibility of properties and by doing so increase land

values. Literature review showed that some researchers defined accessibility as the ease of reaching desired activities such as employment, retail shopping, and healthcare (Dimitriou, 1992; Hanson, 1995; Armstrong, 1997). Accessibility can also be defined as the economic benefits derived from one's ability to reach these activities. However, this study treated accessibility as the relative advantage of one location over another in the ease of reaching other locations at which activities take place (as in Armstrong, 1997; Targa, 2003). Moreover, this study investigated solely the basic access to the transportation system and also through the transportation network to the major activity centers.

There is a wide range of researches exploring impacts of transportation infrastructure and other locational elements on residential and commercial property values. Diamond (1980) investigated 414 parcels sold in the Chicago area from 1969 to 1971 and found that property value increase by 2.1% for each additional mile away from the nearest commuter rail station. Lewis-Workman, et al. (1997) considered census of one-mile radius from a single station (Pleasant Hill) in San Francisco and estimated a premium of $1,578 for every 100 feet closer to BART station. On the other hand, Gatzlaff and smith (1993) studied 912 single-family home sales in Miami from 1971 to 1990 and found no effect of accessibility measured in terms of distance to the nearest Metrorail station.

Most of the similar kind researches reviewed by this study were found to be devoted to examine effect of rail transit systems only. Yet Haider and Miller (2000) considered accessibility effects of highways, shopping malls as well as subway line. They found that propinquity to a subway line adds $4000 to property values. But they concluded that in the presence of other explanatory variables, locational and transportation factors were not strong determinants of housing values (Haider and Miller, 2000). In a similar type study, Abelson (1997) found same kind of result. It also revealed that the house and land prices fall dramatically with the distance from CBD (Abelson, 1997). On the other hand, Simons et al. (1998) drew completely opposite result of increasing housing value with CBD distance.

3. Methodology

As mentioned earlier this study followed hedonic price model approach and used ordinary least square regression method. The basic empirical relationship to be evaluated in this study is the equilibrium implicit price of the property as a function of its attributes, which typically is of the general form:

$$P_i = c + \sum_j \beta_j * X_{ij} + \varepsilon_i \qquad (1.1)$$

where P_i is the sales or rental price of the ith residential property, X_{ij} is the jth attribute for the ith property, β_j is the parameter to be estimated for the jth attribute (implicit empirical marginal price for the attribute), c is the intercept constant term. The intercept term could be interpreted as the fixed start up costs of housing production if sales price is used in the hedonic model. ε_i is the random error term for the ith property (Targa, 2003).

However, the estimation of hedonic price functions might have multicollinearity problems. In addition, there are issues concerning choice of functional form for the price function and spatial autocorrelation. The study investigated correlation coefficient for variables before initial modeling and also performed diagnosis tests to understand the nature of the collinearity in the final model. Regarding functional form specification, both simple linear and semi-log specifications were investigated. The step-wise regression technique was used to determine the best-reduced model.

This study explicitly assumed that error terms are randomly and identically distributed in the OLS hedonic models. However, literature review showed that although hedonic price approach is useful in explaining housing prices, some portion of price variability might remain unexplained. Sometimes, this unexplained variance of housing prices violates standard assumptions of the classical linear regression model. A portion of that unexplained variance might be caused by interdependence among observations in geographic space attributable to their relative locations (Rodriguez and Targa, 2004). Therefore, spatial autocorrelation in the observations may represent a potential problem for the hedonic price function estimation and

convinced this study to run diagnosis tests for the potential autocorrelation.

Table 1.1: Description of the variables

Variable	Definition	Spatial level of data
MEASURE OF VALUE		
RENT	Rent offered price (Taka)	Property
STRUCTURAL ATTRIBUTES		
UL_AREA	Usable living area (sq. ft)	Property
BEDS	Total number of bedrooms	Property
BATHS	Total number of bathrooms	Property
BLD_AGE	Age of the residential property structure	Property
NEIGHBORHOOD ATTRIBUTES		
POP_DENS	Population Density (persons per acre)	Ward
RES_LUSE	Percentage of area dedicated to residential use	Ward
COM_LUSE	Percentage of area dedicated to commercial use	Ward
COMMU_LU	Percentage of area dedicated to community facilities	Ward
DRAINAGE	Euclidian distance to nearest drainage network	Property
TRANSPORTATION ATTRIBUTES		
M_RD_ACC	Network access distance to Major roads (arterials)	Property
CBD_ACC	Network access distance from property to CBD	Property
W_MAR_ACC	Network access distance to wholesale markets	Property
EDU_ACC	Network access distance to primary school	Property
SHOP_ACC	Network access distance to shopping centers	Property

In this study, actual distances between all residential properties were used to define contiguity for the diagnosis of spatial autocorrelation. The distance band was selected as the average of the nearest neighborhood distances of all pairs of properties. Instead of using housing value (as in Haider and Miller, 2000), this study used the residuals of the estimated hedonic price models (as in Targa, 2003). For quantifying the spatial autocorrelations among residuals, this study evaluated both Moran's *I* and Geary's *c*. As vector GIS data model was used in this study, it used irregular lattice-based ROOKCASE (an Excel 97/2000 Visual Basic Add-In) procedure (Sawada, 1999).

4. Data and derived variables

This study was mostly facilitated by large GIS dataset collected by the Rajshahi Master Plan Project. In addition, a field questionnaire survey

was conducted to collect information regarding residential advertised rent prices and structural attributes of the property. In total 550 random sample data were collected by field visual inspection of advertisements for rent during May 2004 at Rajshahi. However, 55 properties were discarded during geo-coding operation and 16 were omitted after a quality control procedure resulting in a 479 samples available for the final analysis.

The study used GIS as a tool for data processing. Thanks to the power, speed and precision of the available GIS software, which enabled this study to derive explanatory variables such as network access distances, euclidian distances and areas of different land uses around properties.

Detailed investigations were carried out to select independent variables in the context of the chosen study area. The independent variables that were finally considered under the three categories are tabulated in Table 1.1.

5. Results and discussions

5.1 Summary statistics

In this study, residential advertised (asking) rental price was selected as the dependent variable. Among the explanatory variables, most of the structural attributes were found through field survey. All of the transportation and neighborhood attributes were measured using ArcInfo, ArcView, and Network Analyst (AVNA) extension module with the aid of ArcAvenue scripting language. Table 1.2 represents the summary statistics for all variables measured in the study.

5.2 Initial modeling

5.2.1 Coefficients of correlations

The first concern in model estimation was to identify relative correlations between the dependent and explanatory variables and also among the explanatory variables themselves. Analysis of coefficient of correlations using commercial statistical software SPSS showed that the explanatory variable BEDS, i.e. the number of bedrooms, has the highest correlation (0.620) with the dependent variable RENT, followed by another structural attribute, the usable

living area (UL_AREA). All the transportation attributes have negative correlation with the dependent variable except EDU_ACC, i.e. network access distance to nearest primary school from the property. However, neighborhood attribute DRAINAGE, the euclidian distance from the property to the nearest drainage network has its expected negative sign in its coefficient of correlation (–0.298), which means as the distance of drainage network increases from the residential property, the rent tends to decrease.

Table 1.2: Summary statistics

Variable	Mean	Std. Deviation	Min	Max	Measurement Unit
Measurement of value					
RENT	1961.86	901.79	300	7000	Taka
Structural attributes					
UL_AREA	1531.96	857.41	200	8000	Sq. feet
BEDS	2.64	0.86	1	4	Number
BATHS	1.48	0.61	0	3	Number
BLD_AGE	18.61	14.06	1	129	No. of years
Neighborhood					
POP_DENS	64.62829	37.81422077	7.433556	161.6547	Persons/acres
RES_LUSE	41.1748	5.6967	27.04	45.35	Percentage
COM_LUSE	5.9542	2.8403	1.36	8.63	Percentage
COMMU_LU	3.1125	1.459	0.19	5.31	Percentage
DRAINAGE	62.68638	119.47954	1.267	760.067	Feet
Transportation					
M_RD_ACC	919.9813	521.52048	48.796	2871.462	Feet
CBD_ACC	2302.315	1355.92145	207.501	5603.634	Feet
W_MAR_AC	1926.987	1320.73124	82.102	5603.447	Feet
EDU_ACC	919.3279	944.1961	3.09	17775.6	Feet
SHOP_ACC	1771.067	1333.1784	88.328	5691.064	Feet

Among the explanatory variables, RES_LUSE has strong negative correlation (–0.880) with SHOP_ACC and W_MAR_AC (–0.833). This implies that residential land uses tend to grow apart from shopping centers and wholesale markets. Moreover, SHOP_ACC is

very strongly correlated with W_MAR_AC (0.913) showing location agglomeration of both wholesale market and retail shopping. The same variable also has significant positive correlation with CBD_ACC (0.886). However, CBD_ACC has most significant correlation (0.972) with the access to wholesale markets (W_MAR_AC).

5.2.2 Initial model

The initial model was generated using the linear functional form specification. The significance of the model was tested by a *F*-test. The null hypothesis of the F-test was "all the regression parameters are not significant", whereas the alternative hypothesis was "at least one parameter is significant." The critical *F*, using $\alpha = 0.05$, is 1.69, which is lower than 42.267. Therefore, the null hypothesis is not rejected and the overall model is significant with a 95% level of confidence.

Table 1.3: Coefficients of linear hedonic model

	Unstandardized Coefficients		Standardized Coefficients	*t* Distribution	*Sig.*
	B	Std. Error	Beta		
(Constant)	-1876.338	614.050		-3.056	.002
UL_AREA	.290	.038	.276	7.657	.000
BEDS	393.278	43.645	.376	9.011	.000
BATHS	185.340	60.592	.125	3.059	.002
BLD_AGE	-.826	2.162	-.013	-.382	.703
POP_DENS	1.819	1.156	.076	1.574	.116
RES_LUSE	34.281	14.501	.217	2.364	.018
COM_LUSE	26.269	28.387	.083	.925	.355
COMMU_LU	69.638	55.223	.113	1.261	.208
DRAINAGE	-.319	.389	-.042	-.819	.413
M_RD_ACC	-.374	.078	-.216	-4.777	.000
CBD_ACC	8.664E-02	.128	.130	.677	.499
W_MAR_AC	.162	.131	.238	1.238	.216
EDU_ACC	-1.118E-02	.031	-.012	-.358	.721
SHOP_ACC	3.082E-02	.085	.046	.365	.716

The coefficient of determination R^2 is 0.560. It implies that 56 percent of the variability is explained by the predictors in the initial model. All of the coefficient estimates were evaluated using a two-

tailed test. The coefficients of the initial model are presented in Table 1.3.

Assessing the usefulness of each predictor in the model is not as simple as comparing the unstandardized coefficients found in the above model. Although all the network and euclidian distance variables were in same units it constituted many other variables that were measured in different units. Even if the explanatory variables are all measured in the same units, a comparison of their size may not be revealing. When the variables are correlated, it is hard to quantify the unique contribution of each variable. However, *beta* coefficients can be an attempt to make the regression coefficient more comparable. The reported standardized coefficient, *beta* can also be found by transforming the data to *z-scores* prior to the regression operation.

In assessing the relative contribution of the explanatory variables based on *beta* coefficients in the initial model, the number of bedrooms of the residential properties (BEDS) is the most influential predictor with a coefficient of 0.376. Usable living area, another structural attribute has second ranked contribution (0.276) followed by a neighborhood attribute, RES_LUSE (0.217). The most important transportation attribute is the network access distance to the major roads, M_RD_ACC (–0.216). The negative sign is expected as the distance to transportation network increases the property value decreases. The number of bathrooms (BATHS) is another significant predictor with a coefficient of 0.125. Other variables are not significant predictor in this model as their t statistics is below the selected $t_{critical}$ ($t_{critical}$ = 1.96 at 95% confidence level).

5.2.3 Functional form specification

Linear ordinary least squares (OLS) regression method is linear in its parameters but not necessary in its variables. Although some of the studies reviewed used only linear model specifications, theoretical issues regarding hedonic price methods suggest a non-linear model specification (Targa, 2003).

Most recent empirical evidence regarding the impact of transportation system on land values indicates that the best specification for the functional form is semi-log (Garrod and Willis,

1992). Indeed, several authors estimated only linear and semi-log specifications, and of the two, better results are produced by the semi-log (Damm et al., 1980; So et al., 1997; Henneberry, 1998).

Therefore, in this study an attempt was made to estimate a semi-log model specification and evaluate its performance with the initial model. In semi-log specification, the dependent variable was transformed (logarithmic transformation) but the independent variables remained untransformed. The results are summarized in Table 1.4.

Table 1.4: Results of OLS semi-log hedonic model

Parameters	Unstandardized coefficient	Std. Error	Beta
(Constant)	2.199	.159	
UL_AREA	6.579E-05***	.000	.234
BEDS	0.121***	.011	.432
BATHS	3.602E-02*	.016	.091
BLD_AGE	-2.969E-06	.001	.000
POP_DENS	2.882E-04	.000	.045
RES_LUSE	9.656E-03**	.004	.228
COM_LUSE	1.185E-02	.007	.139
COMMU_LU	8.420E-03	.014	.051
DRAINAGE	-9.901E-05	.000	-.049
M_RD_ACC	-9.907E-05***	.000	-.214
CBD_ACC	6.985E-05*	.000	.392
W_MAR_AC	-4.586E-06	.000	-.025
EDU_ACC	4.771E-07	.000	.002
SHOP_ACC	-2.438E-06	.000	-.013

*** = 0.1% level of Significance (< 0.001)
** = 1% level of Significance (< 0.01)
* = 5% level of Significance (< 0.05)

ANOVA: $F = 47.714$
$R^2 = 0.590$

The model is significant as its *F*-statistic is 47.714, which is much higher than the critical value of 1.69. The results also indicate that the OLS hedonic price model with semi-log functional form specification explains 59 percents of the variability, which is a three percent increase of the value explained in the initial model using linear functional form specification. This increase is relatively small but the semi-log specification is expected to reduce possible heteroscedasticity in the hedonic price model by compressing the

scale in which the variables are measured. Therefore, semi-log specification was used in the subsequent final model building process.

Table 1.5: Model summary of stepwise regression

Models	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.652	.425	.424	.1833
2	.698	.487	.484	.1734
3	.734	.539	.536	.1645
4	.748	.559	.556	.1610
5	.753	.567	.562	.1597
6	.758	.575	.570	.1584
7	.763	.582	.576	.1572
8	.763	.582	.577	.1571
9	.767	.588	.582	.1562

Table 1.6 Coefficients of final model using stepwise regression

Model		B	Std. Error	Beta
9	(Constant)	2.113	.116	
	BEDS	.119***	.011	.427
	UL_AREA	6.505E-05***	.000	.231
	M_RD_ACC	-1.027E-04***	.000	-.222
	CBD_ACC	6.060E-05***	.000	.340
	RES_LUSE	1.268E-02***	.002	.299
	COM_LUSE	1.408E-02***	.004	.166
	BATHS	3.822E-02*	.015	.096

*** = 0.1% level of Significance (< 0.001)
** = 1% level of Significance (< 0.01)
* = 5% level of Significance (< 0.05)

5.2.4 Best-reduced model

After determining suitable functional form specification, the study aimed to explore the best-reduced model for residential property rent asking price. In total, nine models were constructed in the stepwise regression procedure. Two criteria were used for entering and removing independent variables in the stepwise regression method based on *F* statistic. The first criterion selected for the minimum probability of *F-to-enter* was 0.05 and the second criterion was 0.10 for the maximum probability of *F-to-remove*. The model results after running stepwise regression is presented in the Table 1.5.

Among the nine models, the best-reduced model is Model 9 which is comprised of three structural attributes (BEDS, UL_AREA and BATHS), two neighborhood attributes (RES_LUSE and COM_LUSE) and two transportation attributes (M_RD_ACC and CBD_ACC) with a coefficient of determination, R^2=0.588.

In the best-reduced model, BEDS has the highest contribution as a predictor with a standardized coefficient of 0.427 (Table 1.6). That means, all else being equal, the price of a unit will increase with the increase in number of bedrooms. The second most significant predictor is the CBD_ACC with a positive coefficient of 0.34. This means an increase in distance to CBD increases residential property values at Rajshahi.

Furthermore, the rent of house decreases with the increase of distance to the major roads as the variable M_RD_ACC has a coefficient of –0.222. Another important explanatory variable is the percentage of residential land use (RES_LUSE) having a coefficient of 0.299. Usable living area also has a positive coefficient of 0.231 indicating that larger living spaces demand higher residential rental price. In the best-reduced model, evaluating *t*-statistics we can conclude that all coefficient estimates are significant at least at 99% confidence level except the variable BATHS (significant at 95%).

5.2.5 Diagnosis tests for spatial autocorrelation

The lag used for calculating the Moran's *I* and Geary's *c* is 57.65 meter (Approx. 189 feet), which is the average of the nearest neighborhood distances of the properties concerned. By definition, Moran's *I* has a value between –1 and 1. The expected value of Moran's *I* is $(-1)/(1-n)$ where *n* is the number of observations. In this case, the expected value of Moran's *I* is 0.00209205 whereas the estimated value is 0.058246 (Table 1.7). Hence, it identifies existence of some positive spatial autocorrelation.

On the other hand, Geary's *c* has a value from 0 to 2. The expected value is 1. As the estimated Geary's *c* (0.713191) is close but slightly less than the expected value, it also indicates that there is a little positive spatial autocorrelation in the residuals. However, this

study does not attempt to generate any separate spatial error dependence model.

Table 1.7: Summary result of diagnosis tests for spatial autocorrelation

Summary: Nearest neighborhood statistics						
Min. Dist.	1st Quart.	Median	Mean	3rd Quart.	Max. Dist.	# Neigh. pairs
8.44437	29.10676	44.90971	**57.64587**	75.84801	319.9391	479

Summary: Moran's *I* and Geary's *c*			
Moran's *I* = **0.058246**		Geary's *c* = **0.713191**	
z-Normal *I*	1.076394	z-Normal *c*	3.060892
Var. Normal *I*	0.003142	Var. Normal *c*	0.00878
z-Random *I*	1.077248	z-Random *c*	2.743708
Var. Random *I*	0.003137	Var. Random *c*	0.010927

Total no. of observations, $n = 479$

6. Conclusion

This study revealed that structural attributes are the significant determinants of the residential property values in terms of asking rental price. Number of bedrooms is the most important attribute in the housing attribute bundle to explain housing values at Rajshahi City. However, transportation attributes are also found to be strong determining factors. Accessibility to the CBD and major arterials is affecting the residential price at Rajshahi most. The final model takes the following form:

$P = 2.113 + 0.427(\text{BEDS}) + 0.231\ (\text{UL_AREA}) - 0.222\ (\text{M_RD_ACC}) + 0.340\ (\text{CBD_ACC}) + 0.299\ (\text{RES_LUSE}) + 0.166\ (\text{COM_LUSE}) + 0.960\ (\text{BATHS})$

where P is the asking rental price of the residential property.

One of the most surprising and interesting findings is that the residential property value increases with the increase of network access distances to the Central Business District (CBD). In the context of Bangladesh, the reverse case is expected in most of the cities. The explanation of the result may be that the CBD of Rajshahi (Sayeb Bazar) is not a predominant center of employment. Due to the polycentric nature of the city this finding is not unreasonable.

The study discovered a valuable information that 0.5-mile closer to the major arterials adds 585.858 Taka premiums on residential asking rental price, all else held equal, at Rajshahi City. The capitalization of such accessibility effects has a wide range of

practical applications. Recently, Rajshahi Development Authority (RDA) is going to take an extensive effort for transportation infrastructure investment to increment transportation network through the Rajshahi Master Plan Project 2004. RDA can use the results of this study to collect additional taxes/revenues from the beneficiaries for the implementation of the proposed gridiron based transportation network for the City.

REFERENCES

Abelson, P. (1997) "House and Land Prices in Sydney from 1931 to 1989", *Urban Studies*, Vol. 34: pp. 1381-1400.

Alonso, W. (1964) *Location and Land Use*, Harvard University Press, Cambridge, MA.

Armstrong, R.J. (1997) "Evaluation of the Accessibility Effects and Proximity Related Externalities of Commuter Rail Service", MST Thesis, Massachusetts Institute of Technology, Cambridge, MA.

Damm, D., Lerman, S.R., Learner-Lam, E. and Young, J. (1980) "Response of Urban Real Estate Values in Anticipation of the Washington Metro", *Journal of Transport Economics and Policy*, Vol. 14, No. 3: pp. 315-36.

Diamond, D.B. (1980) "Income and Residential Location: Muth Revisited", *Urban Studies*, Vol. 17: pp. 1-12.

Dimitriou, H. (1992) *Urban Transport Planning: A Developmental Approach*, London: Routledge.

Garrod, G. and Willis, K. (1992) "Valuing Goods' Characteristics: An Application of the Hedonic Price Method to Environmental Attributes", *Journal of Environmental Management*, Vol. 34, Issue 1: pp.59-76.

Gatzlaff, D. and Smith, M.T. (1993) "The Impact of the Miami Metrorail on the Value of Residences Near Station Locations", *Land Economics*, Vol. 69, Issue 1: pp. 54-66.

Haider, M. and Miller, E.J. (2000), "Effects of Transportation Infrastructure and Locational Elements on Residential Real Estate Values: Application of Spatial Autoregressive Techniques", *Transportation Research Record* 1722, pp. 1-8.

Hanson, S. (1995) "Getting There: Urban Transportation in Context" in *The Geography of Urban Transportation*, Second edition. The Guilford Press, New York, NY, pp. 3-25.

Henneberry, J. (1998) "Transport Investment and House Prices", *Journal of Property Valuation & Investment*, Vol. 16, No. 2: pp. 144-158.

Lewis-Workman, S. and Brod, D. (1997) "Measuring the Neighborhood Benefits of Rail Transit Accessibility", *Transportation Research Record* 1576, pp. 147-153.

Muth, R. F. (1969) *Cities and Housing*, University of Chicago Press, Chicago.

Rodriguez, D. A. and Targa, F. (2004), "Value of Accessibility to Bogota's Bus Rapid Transit System", *Transport Reviews*, Vol. 24, No. 5: pp. 587-610.

Sawada, M. (1999) "ROOKCASE: An Excel 97/2000 Visual Basic (VB) Add-In for exploring Global and Local Spatial Autocorrelation" *Bulletin of the Ecological Society of America*, Vol. 80, No. 4: pp. 147-161.

Simons, R.A., Roberto G.Q., and Ivan M. (1998) "The Value Impact of New Residential Construction and Neighbourhood Disinvestment on Residential Sale Price", *Journal of Real Estate Research*, Vol. 15: pp.147-161.

So, H.M., Tse, R.Y.C. and Ganesan, S. (1997) "Estimating the Influence of Transport on House Price: Evidence from Hong Kong", *Journal of Property Valuation & Investment*, forthcoming (cited in Targa, 2003).

Targa, F. (2003), "Examining Accessibility and Proximity-Related Effects of Bogotá's Bus Rapid System Using Spatial Hedonic Price Models", Masters Project, Department of City and Regional Planning, University of North Carolina at Chapel Hill, NC.

Land Use and Capital Costs at Border Crossings: Pacific Highway

This paper will analyze the impact of local land use decisions and community concerns on the capital costs of projects at the Pacific Highway border crossing between Washington State and BC's Lower Mainland. Potential concerns for the future functioning of the crossing will be highlighted. Options for coordinating land use decisions with the operational requirements of major transportation corridors will be discussed, particularly in relation to projects where federal contributions are involved.

The Pacific Highway crossing is one of four in this area. It is located on BC Highway 15 (the City of Surrey's 176 Street). The Douglas/Peace Arch crossing is situated approximately one half mile east on BC's Highway 99 and US Interstate I-5, the major access roads approaching the border, and is restricted to automobile traffic both northbound and southbound. Pacific Highway processes automobiles and commercial traffic (trucks and buses). It is the 4th busiest commercial crossing on the Canada-U.S. border. Rapid increases in truck traffic in the 1990's strained the capacity of the infrastructure at Pacific Highway, and substantial investments have been made to relieve congestion.

Pacific Highway – Northbound System

The northbound approach to Pacific Highway is State Route 543 (SR 543), a two-lane road approximately 1.4 km in length which links the border crossing to the I-5 Interstate highway. An aerial photo of SR 543 from I-5 to the border is shown below. The three intersections (from south to north) are at Boblett Street, H Street and D Street.

Northbound System: SR 543



An Environmental Assessment for proposals to reduce congestion due to border operations was published in August 2000.[1] Problems identified at the border included long delays at the border, and traffic queuing through the "D" Street intersection causing operational problems, and safety concerns regarding vehicles and pedestrians crossing at "D" Street and Boblett Street.[2]
The original proposal for improvements to SR 543 including widening the existing highway, improving the intersection at D

---

[1] SR 543 Interstate 5 to International Boundary Environmental Assessment August 2000; U.S. Department of Transportation Federal Highway Administration and Washington State Department of Transportation.
2 Environmental Assessment, page 5.

Street, and signalization of the intersection at Boblett Street. A Traffic Impact Assessment for this proposal was published in October 2000.[3] Based on traffic forecasts, signalization was recommended at Boblett Street by 2008. D Street was identified as a High Accident Location, and immediate signalization was recommended.[4] It was concluded that this alternative "would allow the D Street and Boblett Street intersections to operate at or above minimum required levels of service".[5]

In addition to this "signalization and channelization" proposal, three additional options were identified following public consultations. They all required lowering of SR 543 and construction of a structure (interchange or grade separation) to maintain the connection for east-west traffic on D Street across SR 543. The option which was selected by Washington State Department of Transportation includes lowering of SR 543 approximately 25 feet to allow traffic to pass under a new Texas "T" interchange at "D" Street; separation of car and truck traffic north of "D" Street; channelization and installation of a signal at Boblett Street; and construction of additional lanes.[6] The cost of the widening and truck bypass option was estimated at $7.9 million; the chosen option was estimated to cost $20.9 million. This has increased to US$33.9[7] million due to unstable geological conditions. Construction of the interchange and retaining walls will require sinking of approximately 1000 drilled shafts 3 feet in diameter, 60 to 70 feet deep.

Closure of the intersection of "D" Street was not considered as an option because it was deemed to be necessary to maintain a direct connection to the neighbourhood to the east, and to maintain access

---

[3] Technical Memorandum SR 543 I-5 to International Boundary Final Traffic Impact Analysis prepared for Washington State Department of Transportation by David Evans and Associates; Bellevue, WA October 2000, p. 30.

[4] Impact Analysis, page 30.

[5] Impact Analysis, page 37.

[6] Environmental Assessment, page 8.

[7] SR 543 – I-5 to Canadian Border Quarterly Project Report Update; WSDOT website

for emergency vehicles.[8] The neighbourhood to the east of SR 543 and north of H Street is zoned Residential Low Density north of D Street and Residential Medium Density south of D street. According to the U.S. census of 2000, this area contains a total of 72 dwellings and a population of 212.[9] Peak hour volumes for east/west movements across SR 543 were forecast at only 15 vehicle movements by 2020.[10] Consideration was given to constructing a missing portion of 14th Street between D and H Streets to provide an alternate route to access SR 543, but this was dropped due to community opposition.[11]

Pacific Highway – Southbound System

The Douglas neighbourhood is bounded by Highway 99 to the West, Highway 15 to the East, 8th Avenue to the North, and 0 Avenue which runs along the U.S. border, to the South. Trucks access the border crossing either via Highway 99 and 8th Avenue (60%), or by Highway 15 (40%). Automobiles using the Douglas/Peace Arch crossing use Highway 99.

Land use planning for the Douglas neighbourhood by the City of Surrey has been shaped by two major objectives: maintaining the agricultural land base, and accommodating and generating economic activity from border-related traffic. The agricultural land base is subject to overriding provincial jurisdiction through the Agricultural Land Reserve (ALR), which was created in the 1970's to preserve the province's agricultural land base.

A special study examining land use options for this area was completed by the City of Surrey's Planning and Development Services branch in 1986.[12] The study examined two options for

---

[8] Personal communication with William Garing, David Evans and Associates.

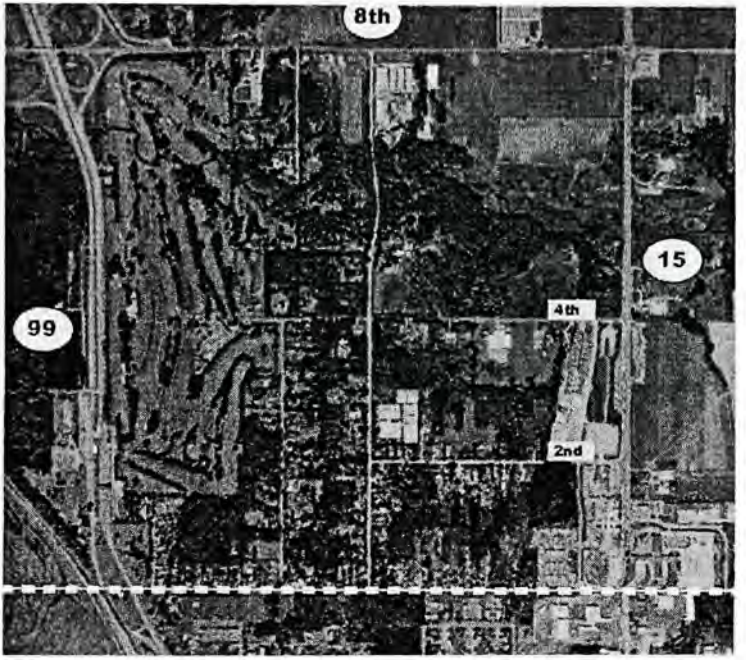
[9] U.S. Census Bureau American Factfinder: Census tract 104.1, Blocks 2000-2005, 2008, 2009.

[10] Impact Analysis page 22.

[11] Personal communication with William Garing, David Evans and Associates.

[12] A Special Study of the Douglas Neighbourhood Report to the Municipal Manager, Surrey Planning and Development Services, March 10, 1986.

Aerial Photo of Douglas Neighbourhood



accommodating commercial activity relating to border operations, subject to the objective of protecting land within the ALR and limiting adverse effects on the Douglas residential area. The option that was adopted by Council was the development of commercial-industrial sites related to border activity to the west of 176 Street (Highway 15) between 4th Avenue and the border. Impacts on the residential area were to be limited by making 0 Avenue and 2nd Avenue into culs-de-sac at the edge of the residential area.[13]
A surge in automobile traffic from 1989 through 1992 caused a crisis in border traffic management. Lineups of southbound automobiles blocked access to the border for trucks and buses, interfered with the

[13] Special Study, p. 13.

operation of the duty free store, and resulted in traffic conflicts as drivers "shortcutted" through the Douglas neighbourhood in an attempt to cut into the queue at 4th or 2nd Avenues. The following recommendations from Surrey's Deputy Municipal Engineer were endorsed by Surrey City Council on March 26, 1990:

"1. That Council support a specifically designated access to the Pacific Highway/U.S. border crossing for commercial traffic ...

2. That Council give approval in principle for the use of a proposed 175A Street for development of a traffic access to the U.S. border

3. That the Province be requested to secure funding and construct 175A Street and it be subsequently designated a secondary highway."[14]

The City endorsed construction of a new truck route from 4th Avenue to the border, paralleling 176 Street to the West along an alignment designated 175 A Street. The section between 2nd Avenue and the border was required to be constructed by the owner of the duty-free store under the terms of his development permit. Avenues 0, 2 and 4 would be closed west of 176 Street to eliminate "short-cutting" through the Douglas neighbourhood. This solution required the Province to fund and construct the new truck access route.

Pressure for changes to the border access roads diminished after 1992 as automobile volumes dwindled. However, growing volumes of commercial vehicle traffic led to renewed calls for improvements by 1996.

In December 1996 the City received an application for rezoning to permit construction of an RV park between 2nd Avenue and 0 Avenue[15]. This application was approved by Council on July 8, 1997.[16] A building permit was issued October 10, 2000 and temporary occupancy was granted on September 16, 2002.[17]

---

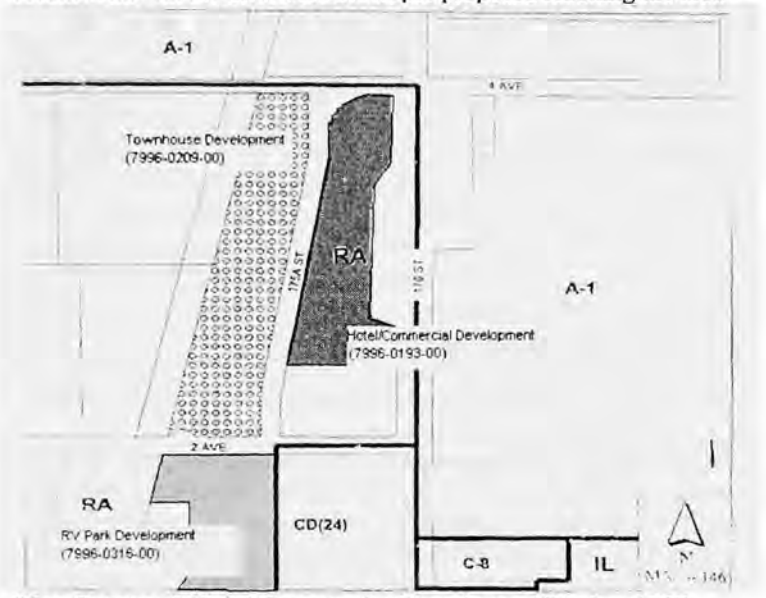
[14] Administration Decision Record March 20, 1990 p.1.

[15] City of Surrey Planning and Development Report File 7996-0316-00.

[16] City of Surrey Regular Council Minutes Monday, July 28, 1997.

[17] Personal communication with City of Surrey Planning and Development Staff. Temporary occupancy essentially grants approval for use of the building.

In July 1997 the City was approached by a developer requesting a change to the zoning of the area west of 176 Street between 2nd Avenue and 4th Avenue. The developer proposed rezoning the area



Map Source: Official Community Plan Annual Review 1997

between the proposed 175A street and 176 Street from Industrial to Commercial to permit development of a complex including a 90 to 100 room hotel, restaurant, and retail/factory outlet area.[18] This development was opposed by the City's Planning and Development department because of concerns that the development would attract customers from other communities in Surrey, compounding the traffic problems at the border and detracting from the viability of established commercial areas in nearby Town Centres.[19] The developer also requested rezoning of the area west of 175A Street from Commercial

[18] City of Surrey Planning and Development Report File 7996-0193-00.

[19] Official Community Plan Annual Review 1997 Corporate Report C393 Application no. 7997-0193-00.

to Urban to allow development of a 72 unit seniors-oriented townhouse complex.[20] This project was opposed by the City's Planning and Development department due to concerns that "the noise and activity from truck traffic along the proposed 175A Street will compromise the livability of residential development on this site"[21]. Both proposals were granted Approval to Proceed by Council. The developer subsequently requested that the City reduce the road dedication for the proposed 175A Street from 20 metres to 16 metres between 2nd Avenue and 4th Avenue, to preclude the use of this section as a truck route.[22] A study was subsequently commissioned to examine options for border traffic improvements.[23] The study identified expansion of parking facilities at the border as "the most critical measure to reduce congestion at the border".[24] It endorsed construction of 175A Street as an express lane to a separate Customs booth for pre-cleared, empty, and CVO/ITS trucks as a means of reducing the demand for parking at the border. It suggested that a southbound lane extending to 4th Avenue would be preferable, while noting potential conflicts with proposed development projects.[25]
The Ministry did not support the construction of an alternate truck route along 175A Street. A City of Surrey staff report notes that at a March 16, 1999 meeting the Ministry representative "reiterated their position with respect to maintaining truck traffic on Highway 15, and stated that they would not be in favour of dedicating 175A Street as a truck route and would not permit erection of signing on Highway 15

---

[20] City of Surrey Planning and Development Report File 7996-0209-00.

[21] Official Community Plan Annual Review 1997 Corporate Report C393 Map Index No.F Non-City initiated OCP Amendment Application no. 7997-0209-00.

[22] 175A Street and 0 Avenue, Truck Routing Update; Report to Mayor and Council. Corporate Report No. R1844, File 4000-502, April 22, 1999, p.2.

[23] Pacific Highway Border Crossing Study Surrey, British Columbia G.D. Hamilton Associates Consulting Ltd., October 1999.

[24] Hamilton Associates, p. ES-11.

[25] Hamilton Associates, p.45.

indicating 175A to be a truck route ..."[26] A subsequent letter to the City from the Ministry indicated that

"The Ministry does not require as a condition of land development the construction of 175A Street as a future alternative truck route to the Douglas border crossing. The Ministry does support any local improvements to the road infrastructure of the border crossing including optional truck routes but the Ministry is not prepared to designate 175A Street as an Arterial Highway (provincial jurisdiction) pursuant to the Highway Act. Should the City decide to provide 175A Street as an alternative truck route to the border, the Ministry would be prepared to allow erection of signing indicating the alternative route".[27]

City staff noted that it would be possible for Surrey to construct 175A Street as a truck route, but in the absence of Ministry cooperation, trucks using the alternate route and attempting to rejoin the queue on Highway 15 at 0A Avenue would cause serious traffic conflicts, and potentially create truck queues on 175A Street which would negatively impact access to the proposed development projects.[28] On April 26, 1999 Council directed that staff consider a reduced road allowance for 175A Street (i.e. 16 metres), as requested by the developer.[29] A request from the landowner responsible for construction of the portion of 175A Street between 2 Avenue and 0 Avenue for a similar reduction in road width was unopposed by the City's Engineering Department.[30] 175A Street was constructed to the reduced standard (16 metres) by the property owners in 1999-2000.

---

[26] Report to Mayor and Council from Manager, Land Development and Transportation 175 Street and 0 Avenue, Truck Routing Corporate Report No: C422 File 4000-502, April 12, 1999, p.2.

[27] Letter from Supervisor, Planning and Development, Lower Mainland South Coast Region, Ministry of Transportation and Highways, to Ken Zondervan, City of Surrey, April 6, 1999.

[28] Corporate Report No: C422, p. 3.

[29] City of Surrey Regular Council Minutes, Item No. R1844, April 26, 1999.

[30] City of Surrey Planning and Development Report Files 7996-0316-00 and 7999-0143-00.

The townhouse complex (named Douglas Pointe) was built in 8 phases. Building permits were issued between October 1, 2002 and July 8, 2003. Temporary Occupancies were granted between July 10, 2003 and November 1, 2004.[31] The hotel complex has not been developed – a subsequent application to amend the development permit to include a banquet hall was submitted in April 2003.[32] This application was rejected. The property owner is currently pursuing an application to subdivide the property.

Capital Projects at Pacific Highway – Southbound

On November 19, 2002 Transport Canada and the BC Ministry of Transportation announced funding of $30.4 million for the Lower Mainland border area under the Government of Canada's Strategic Highway Infrastructure Program (SHIP). Projects included widening of 8th Avenue between Highway 99 and Highway 15; improvements to the Highway 99/8th Avenue interchange; additional southbound lanes at the border for expedited traffic; an Advanced Traveller Information System to enable automobiles to choose which crossing to use, based on estimated delays; and improvements to the intersection at Highway 15/88th Avenue. Funding for additional federal/provincial projects related to the Pacific Highway crossing totaling $210 million was announced on March 5, 2003 under the federal Border Infrastructure Fund program, bringing the total of projects directly linked to the Pacific Highway crossing to $240 million.

The projects immediately adjacent to the Pacific Highway crossing included a vehicle staging area; an additional lane from approximately 4th Avenue to the border for expedited trucks (under the FAST program)33 and an additional lane for expedited automobiles (under the Nexus program) from approximately 2nd

---

[31] Personal communication with City of Surrey Planning and Development Staff.

[32] City of Surrey Planning and Development Report Files 7903-0136-00.

[33] British Columbia Lower Mainland Trade Corridor Border Projects Business Case Evaluation Final Report; prepared for British Columbia Ministry of Transportation by UMA Engineering (UMA), April 2002, p. 22.

Avenue to the border.[34] The initial design proposal for extending NEXUS contemplated reducing the number of northbound lanes between 1st and 2nd Avenue to one lane. The FAST lane proposal envisaged a single lane to accommodate FAST, empty, and non-broker (i.e. those not needing to access the parking area) trucks, with another lane accessing the parking area. This proposal was opposed by stakeholders due to concerns over potential northbound congestion, and concern with impeding the FAST truck flow. The Ministry therefore identified an alternative concept.
It included a combined and dedicated FAST/NEXUS/bus lane from 4th Avenue to 1st Avenue, diverging into three separate lanes to the U.S. Border; designated travel lanes for all other trucks along Highway 15 south of 4th Avenue, 2nd Avenue and a new two-lane roadway parallel to 175A Street behind the Duty Free store; and a parking area that is reconfigured to accommodate broker trucks as well as a staging area in which all trucks – except FAST trucks – would line-up to proceed to U.S. Customs.
On July 8, 2003 the Ministry of Transportation hosted an open house in the Douglas neighbourhood to solicit public input on the NEXUS and FAST lane proposals.[35] Approximately 70 individuals attended, and the majority favoured the revised proposal. The new route is essentially that favoured by the City of Surrey in their previous discussions with the Ministry, though extending only to 2nd Avenue instead of 4th Avenue as was originally envisioned. The total cost of the revised project was $3 million, and the new truck routing was implemented in October 2004.

Future Challenges for the Pacific Highway Crossing

Evolving land use patterns in the Douglas neighbourhood may pose challenges to the ability of the border crossing to accommodate future growth in commercial vehicle traffic, and necessitate further capital expenditures. Community resistance to the rerouting of trucks via 2nd Avenue and 175B Street is an example.

---

[34] UMA, p. 43.

[35] Information on project design and public consultation is taken from BC Ministry of Transportation project documents.

Residents of the recently completed Douglas Pointe townhouse complex (west of 175A Street between 2nd Avenue and 4th Avenue) have filed numerous complaints with the Ministry of Transportation regarding increased noise, vibration and air pollution from trucks queued along the new alignment. The new route brings the trucks within 100 metres of the townhouses at the south end of the complex. The Ministry commissioned a study on the noise impact which concluded that noise levels were below the threshold for mitigation under both CMHC and BC Ministry of Transportation guidelines. [36] The original project environmental assessment did not consider local air quality effects, and residents are pressing for local testing. The Ministry of Transportation has met with residents and offered mitigation measures including construction of a sound barrier, and signalization which will queue trucks on Highway 15 rather than adjacent to the townhouses. The issue has not been resolved to date (as of February 2005).

Additional traffic conflicts and noise concerns may be generated if the commercial development (hotel and/or some alternate project) on Highway 15 between 2nd Avenue and 4th Avenue is constructed. With the rezoning of the land on the west side of Highway 15, the City has received requests for development permits for border-related industrial uses for parcels of land in the ALR east of Highway 15 adjacent to the Canada Customs facility. City staff has been instructed to undertake a market research study to evaluate the demand for industrial land in the border area.[37]

In the longer term, residential developments within the Douglas neighbourhood may cause traffic conflicts at the two major access points to the neighbourhood, 172 Street at 8th Avenue, and 4th Avenue at Highway 15. Both of these locations are subject to heavy border traffic. Signals are being installed as part of the Border SHIP projects, which should improve service levels for residents. However,

---

[36] Highway 15 FAST/NEXUS Pacific Border Project Community Noise Mitigation Report Report No. 04-999-1 prepared for BC Ministry of Transportation by Wakefield Acoustics November 2004; p.7.

[37] Regular Council – Land Use Minutes City of Surrey Tuesday, September 7, 2004 and personal communications with City staff.

the Douglas Neighbourhood Concept Plan[38] envisions construction of Single Family subdivisions and multiple residential developments (i.e. apartments and townhouses) which would result in an increase of approximately 2800 persons to the existing community population of 900 (estimated 1999), and an increase of 954 housing units to the existing stock of 270 units. As the Douglas neighbourhood currently has no commercial services and no transit service, increased population converts directly into increased vehicle trips.
Additional investments in port of entry facilities may also be required to handle anticipated increases in border traffic. A recent study conducted for the Ministry of Transportation analyzed southbound queue lengths under forecast truck traffic levels and concluded "Preliminary runs of the queuing model indicated that with the current number of booths at the border, the truck queues would extend far beyond 8th Avenue for the future horizon years and that ... additional booths would be needed in the future".[39] This may be problematic due to the physical constraints of the existing US port of entry site. The existing port of entry was constructed in three stages starting in 1997 and ending in 2000. It has three booths for processing commercial traffic, and there is little or no room for adding additional booths. By way of comparison, current plans for rebuilding the U.S. port of entry at Champlain–Lacolle, New York call for construction of 7 commercial processing booths to handle a similar level of commercial traffic.

Policy Implications for the Federal Government

The problems encountered at the Pacific Highway border crossing provide a compelling lesson on the impact of land use decisions on capital costs, and on the potential for accommodating future growth at key transportation facilities. The major lesson to be drawn is that where local land developments encroach on major transportation

---

[38] Douglas Neighbourhood Concept Plan Stage 2 Report City of Surrey Planning and Development, July 7, 1999.

[39] Updated Business Case Extension of Highway 15 Fast Truck Lane Technical memorandum prepared for BC Ministry of Transportation by ND LEA Consultants, August 20, 2004 p.2.

assets, the mitigation costs which must be incurred to satisfy local concerns are not necessarily limited to those which would be thought appropriate to ensure an acceptable level of service. Encroaching residential and commercial development also limits options for further capacity expansions which may be required to accommodate traffic growth.

This is an issue of importance to the federal government, which has recognized the importance of these facilities to the national economy and provided hundreds of millions of dollars to fund improvements. In the case of the Pacific Highway, recent federal investments in highway-related projects directly linked to the border crossing have totaled $102 million. Controlling land use is a difficult exercise for the federal government because it generally lies in the jurisdiction of local governments. However they are subject to overriding provincial jurisdiction. Three potential options for controlling land use for transportation facilities of national importance are given below.

1. The federal government could pursue agreements with local government agencies where federal funding is provided for projects. This not very practical because local governments are seldom signatories to federal funding agreements (bilateral agreements between the federal and provincial governments are the primary mechanism for federal highway funding). Municipal agreements are subject to overriding provincial jurisdiction, and local councils tend to be preoccupied with local developments and issues.
2. Inclusion of provincially-applied land use restrictions adjacent to transportation facilities where federal project funding is provided. Under the provincial Transportation Act, the Province has the authority to influence zoning of land within 800 metres of any controlled access highway.[40] However this power is not exercised directly by the Ministry of Transportation, but delegated to individual approval officers. Implementing this proposal would require a change to centralize authorization of developments within the Ministry proper. The federal government might encourage this change by providing funding for the purchase of land dedicated to transportation use, which is currently an ineligible expense for federal cost-sharing.

[40] Transportation Act RSBC 2004 section 52.

This may be effective in preventing future encroachment but (as the Pacific Highway example has shown) problems may be due to decisions taken prior to involvement of the federal authorities in funding projects.

3. Creation (by federal-provincial agreement) of a Transportation Land Reserve similar to BC's Agricultural Land Reserve to protect key national transportation facilities from encroachment. This could be used to protect key ports and airports, and rail and highway corridors, from encroachment. The concept of overriding restrictions on land use in the vicinity of transportation facilities is not new. The Province of Alberta imposes special zoning restrictions adjacent to the province's two major airports – Edmonton International and Calgary International – under section 693 of its Municipal Government Act[41]. A B.C. study on port competitiveness suggested in 2001 the creation of a "Port Land Reserve".[42] However, the concept of a comprehensive Transportation Land Reserve established by Federal-Provincial agreement has not to my knowledge been discussed. It could provide a mechanism for the federal government to ensure that essential components of the national transportation system continue to function efficiently, and that federal project funding provides lasting benefits.

The author would like to thank all of the many people who provided information which was used in this paper, particularly William Garing of David Evans and Associates; Todd Harrison of Washington State Department of Transportation; Lor Pellegrino, Gary Vlieg, Ken Zondervan, Mike Dickinson, Jaime Boan and Tamara Wallace of the City of Surrey; Lina Halwani, Max Walker and Frank Blasetti of BC Ministry of Transportation; and Hugh Conroy, Melissa Miller, and Gordon Rogers of the IMTC Project (Whatcom Council of Governments). Responsibility for all conclusions reached and opinions expressed in this paper rests solely with the author.

---

[41] Municipal Government Act, R.S.A. 2000, c. M-26

[42] Options to Improve the Competitiveness of Ports in British Columbia Volume I – Final Report prepared for BC Ministry of Employment and Investment by Perrin, Thorau and Associates Ltd. March 2001; p. 45.

"Too Many Trucks on the Road"?

Gordon G. Baldwin, Transportation Division
Transportation Division
Statistics Canada, Ottawa ON K1A 0T6

Introduction

Since the late 1990's, there appears to be a perception that congestion, especially with the competition for road space between cars and trucks, is growing in Canada. There are "public concerns that there are too many trucks on the road".[1] What may be the cause? "As a result of deregulation, there are far too many trucks on the road..." is one explanation put forward.[2] This paper will provide a view of what are the characteristics of that truck population on the roads in Canada.

This paper is based primarily on the registration and performance data available from the Canadian Vehicle Survey. This is a voluntary vehicle-based survey started in 1999, conducted by Statistics Canada with funding from Transport Canada.

How Many Trucks Are There?

In Canada, in 2002, there were 366,969 registered small trucks (in the 4.5t to 15t range), large trucks (those 15t or more) had a further 277,347 registered. This compared to 17,543,664 registered cars (on-road vehicles up to 4.5t). This means that there were 47.8 cars to every small truck and 63 cars to every large truck.[3] During 2003, the number of light trucks registered increased by 11,289 and the number of large trucks increased by 4,845. The number of registered cars,

however, increased by 211,425 which means that the proportion of trucks to cars actually decreased slightly in 2004 to 46.9 cars to every small truck and 62.9 cars to every large truck.[4]

Vehicle Body Types

The vehicle body types, not surprising, are quite different for the small trucks versus the large trucks. This description also helps give a better picture of the vehicles in the respective groups. Small trucks, in 2003, were made up of:

- 4% vans;
- 11% pickups;
- 81% straight trucks;
- 3% tractor trailers;
- 2% others.

Large trucks, in 2003, were made up of:

- 41% straight trucks;
- 58% tractor trailers;
- 1% others.[5]

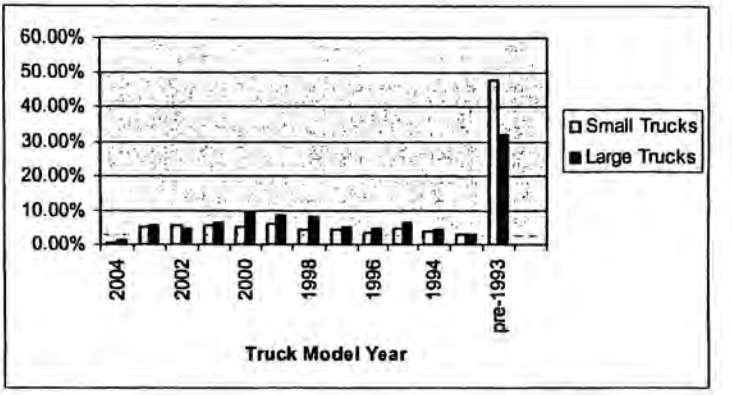
The small truck fleet is made up of 40% gasoline users, 58% diesel and 2% other fuels. This compares with the large truck fleet where only 7% use gasoline while 93% use diesel.[6]

Age of the Fleet

What other differences were there between the small and large truck fleets on the roads in Canada? Smaller trucks appear to have been kept longer as almost half of the registered fleet in 2003 was pre-1993 model year (see Chart 1). For the larger trucks, only 32% were of this vintage.

For the larger trucks, 28% were 5 years old or younger while only 22% of the smaller trucks were this young.

Chart 1 – Age of the Trucking Fleet in Canada, 2003



Source: Statistics Canada, Catalogue 53-223-XIE, Tables 3 and 4, Canadian Vehicle Survey, Annual 2003

Vehicle-Kilometres Performed

While obviously the number of trucks would affect one's perception of how many trucks are on the road, so also would how much they travel. While we saw earlier that there were more small trucks than large trucks, the larger trucks are driven more. In 2003, smaller trucks were driven 6.2 billion kilometres while the large truck fleet, used more for long-haul trucking, drove 18.6 billion kilometres.[7] The average number of kilometres driven per day (see Table 1) was slightly higher for small trucks than for cars. For large trucks the average annual kilometres was three to four times higher, once again reflecting their use for long-haul trucking.

Table 1 – Vehicle-Kilometres Driven, 2003

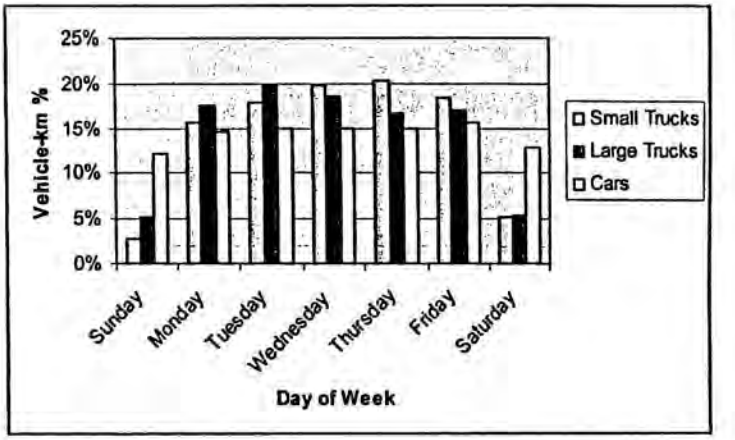
	Total Vehicle-Kilometres (millions)	Average Annual Vehicle-Kilometres per vehicle	Average Daily Vehicle-Kilometres per vehicle
Small trucks	6 163.5	19 144	52
Large trucks	18 567.9	66 640	183
Cars	286 271.2	16 333	45

Source: Statistics Canada, Canadian Vehicle Survey—Annual 2003, June 2004, pages 26 and 30

Day of Week Differences

Are cars and trucks sharing the road at the same time? This could also contribute to one's perception on the volume of truck traffic. Truck traffic is more concentrated during weekdays (see Chart 2) with only 8% of the vehicle-kilometres driven on the weekend for small trucks and 11% for large trucks. Car traffic is much more balanced with between 12-16% of vehicle-kilometres driven daily.

Chart 2 Vehicle-kilometres by Day of the Week



Source: Statistics Canada, Catalogue 53-223-XIE, Table 18, Canadian Vehicle Survey, Annual 2003

Time of Day Differences

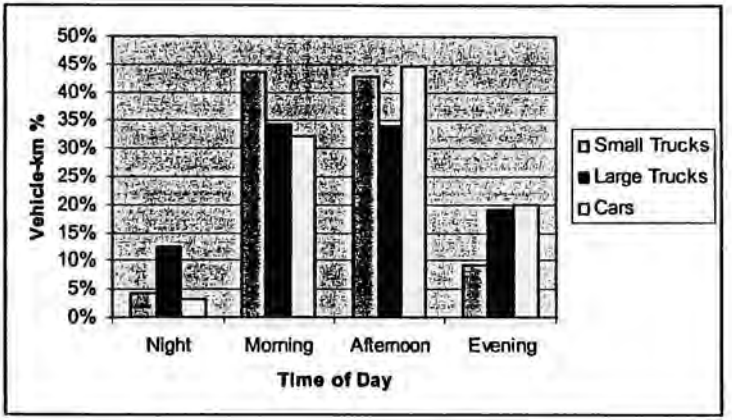
It is not just the day of week but also the time of day that vehicles are driven that could affect one's perception of the volume of truck traffic. If one divides the 24-hour day into four periods:

- Night 00:00 – 05:59
- Morning 06:00 – 11:59
- Afternoon 12:00 – 17:59
- Evening 18:00 – 23:59

and examines the percentage of vehicle-kilometres driven during these periods, it can be seen that cars, small trucks and large trucks are all competing for road space primarily during the morning and afternoon periods (see Chart 3). Small trucks drive 87% of their vehicle-kilometres during these two periods, cars 77% and large trucks 68%. Both large trucks and cars drive approximately one-fifth during the evening while large trucks are those that put the largest

proportion of their travelling between midnight and 6:00 a.m. with 12% of their vehicle-kilometre registered overnight as contrasted with only 4% for small trucks and 3% for cars.[8] The Ontario Trucking Association has stated that "If more shippers were willing to have freight delivered at night when traffic volumes are at their lowest, some of the truck traffic could be taken out of the road during peak congestion periods".[9]

Chart 3 Vehicle-kilometres by Time of Day



Source: Statistics Canada, Catalogue 53-223-XIE, Table 24, Canadian Vehicle Survey, Annual 2003

Concentrations at the Border

The data examined so far have been at the level of Canada totals. Not only can truck activity be concentrated into certain days of the week and certain times of the day but truck traffic and car traffic can be concentrated in specific areas. One area of concentration and the resulting congestion that has garnered much attention is at the Canada-United States border. The data in Table 2 show why this is

so. While, earlier it was noted that in terms of vehicles registered in Canada, there were 47.8 cars to every small truck and 63 cars to very large truck, the numbers are very different are select border crossings.

Based on U.S. Customs border crossing counts, at all land crossings in 2003, for crossings into the United States from Canada there were 4.5 cars (personal vehicles) per truck crossing. At some busy commercial crossing such as Alexandria Bay NY and Port Huron NY, the figures approached only 2 cars for every truck crossing. (Note: the data in Table 2 would include both Canadian and American registered vehicles.)

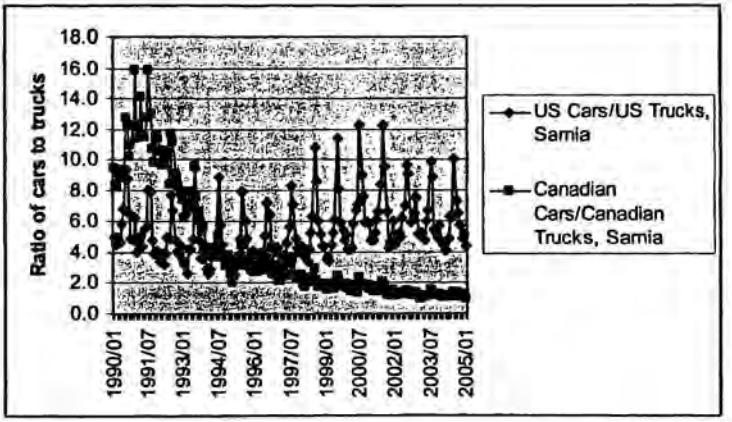
Table 2 – Vehicle Crossings, U.S.-Canadian Border, 2003[10]

U.S. Border Crossing	Trucks (Canada to U.S.)	Personal Vehicles (Canada to U.S.)	Ratio of Personal Vehicles to Trucks at Border Crossing
Detroit MI	1,634,319	6,315,590	3.9
Alexandria Bay NY	297,220	649,161	2.2
Buffalo-Niagara NY	1,162,961	6,414,415	5.5
Port Huron MI	928,074	1,965,011	2.1
Blaine WA	365,089	2,299,636	6.3
U.S.-Canada Border Total	6,728,228	30,220,184	4.5

Source: Bureau of Transportation Statistics, 'Border Crossing/Entry data" based on data from U.S. Customs Service, www.BTS.gov

If one looks at the Canadian border crossing data, one can see the same trend for northbound traffic with the proportion of truck traffic increasing at many border crossings. For all land crossings in 2003, for crossings into Canada from the United States, there were 4.5 cars per truck crossing.[11] The Canadian crossing data are released showing nationality of the vehicle. This shows a ratio of 7.2 U.S. cars per U.S. truck and 3.4 Canadian cars per Canadian truck.

Chart 4 Ratio of Cars to Trucks at Sarnia Border Crossing, Monthly, January 1990-January 2005[12]



Source: Statistics Canada CANSIM Table 427-00002: Number of vehicles travelling between Canada and the United States, monthly

Over time the number of Canadian cars relative to trucks has greatly changed at many border points. For example, at the Sarnia ON/Port Huron NY crossing (Chart 4) in 1990, the monthly ratio of U.S. cars to U.S. trucks ranged from a low of 4.5 to a high of 9.4 while in 2004 the range was from 4.1 to 10.1. For Canadian vehicles the monthly ratio ranged from a low of 8.2 to a high of 15.8 in 1990 to a range of 1.0 to 1.4 in 2004. While actual queuing at the border may well be in

separate lines for commercial versus personal vehicles, approaching many border crossings one could see a much higher proportion of trucks than the 47.8 cars to every small truck and 63 cars to very large truck shown in the national registration data.

Conclusion

It is apparent that there are large numbers of cars and trucks in Canada and the fleets are growing. These vehicles are competing for road space in growing numbers, often during the same day and the same period of the day. There also are areas of concentration such as borders where this competition is even more pronounced. With economic growth and growing populations, we are likely to see this competition increase even more in the future.

Note

The views and opinions expressed in this paper are those of the authors and do not necessarily reflect those of Statistics Canada. I would like to thank Tim Davis and Wendy Christoff of Statistics Canada and Terry Ganton, Transport Canada for their valuable comments, corrections and criticisms while retaining all responsibility for any remaining errors.

Endnotes

[1] *Moving Forward A Guide on the Importance of Transportation in Canada* page 21; December, 1999, http://www.westac.com/pdfs/mffreight.pdf

See also "There certainly is a perception that there are too many trucks on the road and they're impeding others' ability to be mobile" Mr. Paul Szabo (Mississauga South, Lib.), 37th PARLIAMENT, 1st Session, Standing Committee on Transport and Government Operations, Tuesday, April 30, 2002 http://www.parl.gc.ca/InfoComDoc/37/1/TRGO/Meetings/Evidence/trgoev62-e.htm

[2]) Ms. Cathy Walker, National Health and Safety Director, Canadian Automobile Workers; Proceedings of the Subcommittee on Transportation Safety, Standing Senate Committee on Transport and Communications; Issue 6 - Evidence, OTTAWA, Thursday, February 13, 1997
http://www.parl.gc.ca/english/senate/com-e/safe-e/06ev-e.htm

[3] Statistics Canada, Canadian Vehicle Survey, Annual 2002, page 21

[4] Statistics Canada, Canadian Vehicle Survey, Annual 2003, page 21

[5] Statistics Canada, Canadian Vehicle Survey, Annual 2003, page 28

[6] Statistics Canada, Canadian Vehicle Survey, Annual 2003, page 29

[7] Statistics Canada, Canadian Vehicle Survey, Annual 2003, page 30
[8] Statistics Canada, Canadian Vehicle Survey, Annual 2003, page 44

[9] Ontario Trucking Association, Relieving Traffic Congestion: A Discussion Paper, Spring 2001, page 8,
www.ontruck.org/issues/docs/pdf/relievingtrafficcongestion.pdf

[10] Note: the data in Table 2 are different from the other data in this report in the following:

- The data are not from Statistics Canada's Canadian Vehicle Survey but from the U.S. Bureau of Transportation Statistics;
- These data would include vehicles of all nationalities crossing the border. The CVS data are Canadian registered vehicles only.
- The U.S. border crossing data for personal vehicles includes motorcycles and snowmobiles which are not included in the Canadian Vehicle Survey counts.
- The U.S. border crossing administrative data can group several crossings. For example, the Buffalo-Niagara data

include the Lewiston Bridge, Peace Bridge, Rainbow Bridge and Whirlpool Bridge data.

[11] Statistics Canada, CANSIM Table 427-0002. The frontier count is done using the information collected about the entrants into Canada recorded on forms by Canada Customs and Revenue Agency (CCRA) officials. Each port of entry sends in its administrative data according to an understanding signed by Statistics Canada and CCRA.

At all ports of entry across Canada, a count is done to determine the number the number of vehicles (cars, trucks, motorcycles, snowmobiles and bicycles) in the case of highway and ferry points.

[12] Statistics Canada, CANSIM Table 427-0002.

Un portrait du marché du travail des conducteurs dans le secteur du transport routier des marchandises au Québec

Yvon Bigras, Professeur, École des sciences de la gestion
Université du Québec à Montréal
et
Guy Normandeau, Directeur général, Camo-route inc.

Introduction

Cette communication présente quelques résultats de trois enquêtes réalisées pour Camo-route et dont le but était de dresser un portrait de la situation du marché du travail dans le secteur du transport des marchandises au Québec. Les enquêtes ont été réalisées auprès de trois populations différentes :

- les détenteurs de permis de conduire de classes 1 et 3 dont l'emploi consiste principalement à conduire un camion nécessitant ces classes de permis et qui travaillent pour un employeur dont l'activité principale est le transport routier de marchandises. Sur une population totale estimée à 57 109, il y a eu 514 conducteurs qui ont répondu au questionnaire par le biais d'entrevues téléphoniques réalisées en juin et juillet 2003. Le taux de réponse s'est chiffré à 69,3 %, avec un taux de non-réponse estimé de 21,7 % et un taux de refus de 9,0 %.
- les propriétaires-exploitants possédant un ou deux véhicules servant au transport de marchandises, dont la totalité des activités sont réalisées pour autrui et dont une bonne partie de leur travail consiste à conduire les véhicules qu'ils exploitent (PEVL1&2). Sur une population totale estimée à 6 031, il y a eu 702 propriétaires-exploitants qui ont répondu au questionnaire par le

biais d'entrevues téléphoniques réalisées en mai, juin et juillet 2003. Le taux de réponse s'est chiffré à 64,3 %, avec un taux de non-réponse de 24,2 % et un taux de refus de 11,4 %.

- les entreprises possédant au moins trois véhicules servant au transport de marchandises (PEVL3+). La population totale de ce groupe a été estimée à 4 873, et 1 707 entreprises ont répondu au questionnaire par le biais d'entrevues téléphoniques réalisées entre décembre 2002 et février 2003. Le taux de réponse s'est chiffré à 65,3 %, avec un taux de non-réponse estimé de 19,9 % et un taux de refus de 14,8 %.

Ces enquêtes nous ont d'abord permis de constater que le transport routier est un secteur caractérisé par la très grande diversité des entreprises et des ressources humaines qui y œuvrent. Le type de marchandises transportées, le rayon d'action, les équipements utilisés, les ressources mises en place et la taille sont autant de variables qui peuvent expliquer des différences fondamentales entre les entreprises et les propriétaires-exploitants qui offrent leurs services. Il y a par ailleurs plusieurs indications qui confirment que l'industrie du transport routier demeure très concurrentielle, mais les plus grands transporteurs routiers occupent quand même une place prépondérante dans le marché.

Dans ces conditions, il est certes difficile et risqué de vouloir généraliser certaines observations sur le secteur. Malgré ces contraintes, la richesse des enquêtes réalisées pour Camo-route et le grand nombre de répondants qu'elles ont permis de rejoindre, nous donnent la possibilité de tirer un portrait nuancé de ce secteur qui reflètent la situation des entreprises et des ressources humaines qui s'y trouvent. Nous avons en effet retenu ici uniquement des résultats qui sont significatifs statistiquement. Ainsi, lorsqu'on présente des résultats pour des sous-groupes de répondants, les différences sont significatives à un seuil de signification de 5 % ou moins. Le plus souvent, le test utilisé est le test de khi-carré de Pearson qui permet de conclure de façon significative à l'existence d'un lien statistique entre les deux variables d'un tableau croisé. À quelques occasions, on

utilise le test de Fisher pour vérifier que les moyennes sur des variables continues sont différentes entre les différents sous-groupes.

Les informations que l'on retrouve dans cette communication sont par ailleurs tirées d'un rapport disponible sur le site de Camo-route[1]. Nous nous limitons ici à présenter quelques résultats particulièrement intéressants qui traitent de la question du recrutement des conducteurs, de la mobilité en emploi de ceux-ci et de leur satisfaction au travail.

Recrutement des conducteurs

Le manque de conducteurs est certes une grande préoccupation dans l'industrie, d'autant plus que la population des conducteurs est relativement âgée. L'enquête auprès des PEVL3+ confirme l'ampleur de ce problème. D'abord, au tableau 1 on indique le pourcentage de transporteurs ayant de la difficulté à recruter des conducteurs dans le secteur qui est leur principal secteur d'activité. Les résultats sont plus marqués pour les entreprises dont le principal secteur d'activité est le transport international, puisque 88,6 % d'entre elles ont de la difficulté à recruter des chauffeurs pour ce secteur d'activité. Cependant, la proportion des entreprises spécialisées dans le transport local qui rencontre la même difficulté atteint pour sa part 74,6 %, ce qui est un résultat assez surprenant puisqu'on associe souvent le manque de conducteurs au transport routier longue distance.

Le tableau 2 montre par ailleurs que la difficulté de recruter croît avec la taille des entreprises, mais que l'écart entre celles qui ont entre 11 et 49 véhicules (67,6 %) et celles qui en ont plus de 50 (70,3 %) est relativement faible. Dans le même tableau on peut aussi noter une importante différence entre les transporteurs pour compte d'autrui

[1] Camo-route inc. (2005), *Diagnostic sectoriel - Transport routier de marchandises 2004 : Analyse d'enquêtes portant sur la main-d'oeuvre de l'industie*, 78 p. (disponible à l'adresse www.camo-route.com)

qui, avec un pourcentage de 66,5 %, ont beaucoup plus de difficulté à recruter que les transporteurs pour compte propre (45,9 %).

Tableau 1 Entreprises ayant de la difficulté à recruter des conducteurs pour leur principal secteur d'activité (en %)	
Transport local (n= 515)	74,6 %
Transport interrégional (n=192)	57,8 %
Transport interprovincial (n=92)	57,6 %
Transport international (n=132)	88,6 %

Tableau 2 Entreprises ayant de la difficulté à recruter des conducteurs selon la taille de l'entreprise et le type d'entreprise (en %)	
Ensemble des PEVL3+ (n = 1678)	55,5 %
Taille des entreprises	
3 à 4 véhicules (n = 515)	44,3 %
5 à 10 véhicules (n = 631)	54,2 %
11 à 49 véhicules (n = 441)	67,6 %
50 et plus (n = 91)	70,3 %
Type d'entreprise	
Compte propre (n = 945)	45,9 %
Compte d'autrui (n = 731)	66,5 %

Le problème de recrutement rencontré par un nombre croissant de transporteurs a aussi provoqué une évolution des perceptions quant à la principale menace qui pèse sur le secteur. En effet, l'enquête réalisée pour Camo-route en 1997 donnait la concurrence sur les tarifs comme première menace pour 61 % des entreprises du secteur, puis venait le coût du carburant avec 13 % et le manque de personnel qualifié avec 6 %. Or dans la présente enquête, 22,6 % considère la concurrence sur les tarifs comme la première menace, ce qui est à peine plus que le 21,7 % qui considère plutôt le manque de personnel qualifié comme la première menace. Le coût du carburant recueille quant à lui 17,9 %. Si on regarde plus en détail selon le principal secteur d'activité, on constate que le manque de personnel est

considéré comme la principale menace par les transporteurs dont le principal secteur d'activité est le transport local (23,2 %) et le transport international (23,4 %). Pour ceux dont le principal secteur d'activité est le transport interrégional, la principale menace est la concurrence sur les tarifs (26,9 %), et pour ceux dont c'est le transport interprovincial, la principale menace est le coût du carburant (24,6 %).

Globalement, on peut conclure que ces données rejoignent celles du CCRHC[2] sur les difficultés rencontrées dans l'industrie au niveau du recrutement des conducteurs. Dans cette étude on demandait aux entreprises si la pénurie de chauffeurs figurait parmi l'une des deux principales préoccupations et l'on obtenait environ 50 % de réponses positives à cette question. Par ailleurs, environ 75 % des entreprises considéraient que la situation allait empirer au cours des cinq prochaines années. L'étude du CCRHC confirme également que la situation est plus préoccupante chez les plus grands transporteurs, chez les transporteurs pour compte d'autrui et chez les transporteurs qui font du transport transfrontalier. Cette étude ajoute par ailleurs qu'environ 30 % des gestionnaires de parcs pensaient que le plus gros problème, ce n'est pas la pénurie de chauffeurs mais plutôt le trop grand nombre de camions sur les routes.

L'enquête auprès des PEVL3+ présente par ailleurs les trois principales attentes des conducteurs mentionnées lors de leur recrutement par des transporteurs. Pour l'ensemble, le salaire ressort nettement puisque 67,9 % des répondants l'ont mentionné, la stabilité des horaires arrivant deuxième avec 41,3 % et le nombre d'heures travaillées arrivant troisième avec 39,3 %. Selon le principal secteur d'activité, pour cinq des sept attentes proposées il y a une différence significative à un seuil de signification de 1 %. On constate par ailleurs que la présence au domicile est un enjeu qui prend de l'importance au fur et à mesure que l'on s'éloigne, et qu'à l'inverse la

[2] Conseil canadien des ressources humaines en camionnage, *Profil de la pénurie de routiers, du taux de roulement et de la demande anticipée*, CCRHC, Ottawa, 2004, p. 13 à 16.

stabilité des horaires et le nombre d'heures travaillées sont plus importants pour les conducteurs opérant sur des marchés plus rapprochés. On peut finalement noter que pour le salaire qui est l'attente la plus souvent citée, il n'y a pas de différence significative selon le principal secteur d'activité. Lorsqu'on demande directement aux conducteurs quelle est leur principale attente lorsqu'ils postulent pour un emploi[3], le salaire ressort nettement comme raison principale puisqu'elle est évoquée par 45,3 % d'entre eux, ce qui est très supérieur à la stabilité des horaires (6,9 %) ou à la présence au domicile (5,5 %).

On peut mettre ces résultats en parallèle avec ceux de l'étude du CCRHC[4] qui a demandé aux gestionnaires de parc quelles étaient les raisons de la pénurie de routiers. La raison qui est le plus souvent évoquée est que «les routiers ne sont pas assez bien payés», selon 42,4 % des répondants, puis vient «la piètre qualité de vie et les piètres conditions de travail», avec 35,8 %, et «le métier de routier n'attire pas les jeunes» avec 25,0 %. Ces résultats vont dans le même sens que les résultats de la présente étude et confirment que les salaires arrivent en tête des préoccupations des candidats et que cette préoccupation est cependant suivie de près par les conditions de travail. Par contre, l'enquête menée par le CCRHC ne parvient pas à établir de différence selon le principal secteur d'activité, le type d'entreprise ou la taille de l'entreprise.

La mobilité en emploi

Dans cette section, nous tenterons d'évaluer la mobilité des conducteurs à partir d'un certain nombre d'indicateurs tels que la durée de leur lien d'emploi avec un transporteur et les raisons de leurs

---

[3] Dans cette enquête, les conducteurs ne donnaient qu'un seul choix de réponse.

[4] Conseil canadien des ressources humaines en camionnage, *Profil de la pénurie de routiers, du taux de roulement et de la demande anticipée*, CCRHC, Ottawa, 2004, p. 21 à 24.

départs. Le tableau 3 indique qu'une bonne majorité des conducteurs semblent relativement stables puisque 43,6 % des conducteurs ont occupé un seul emploi au cours des 10 dernières années. On peut parler d'une minorité qui ont occupé quatre emplois (6,1 %) ou cinq emplois et plus (8,2 %) au cours de cette période. On a par ailleurs calculé les pourcentages pour les conducteurs ayant plus de 10 ans d'expérience. Or, curieusement, il n'y a pas vraiment de différence, ce qui nous amène à conclure que parmi les plus jeunes conducteurs, un certain nombre change assez souvent d'emploi mais par la suite, soit qu'ils se stabilisent ou qu'ils abandonnent le métier de conducteur de camions après quelques essais plus ou moins fructueux chez différents transporteurs. La relative stabilité en emploi pour une grande majorité des conducteurs est par ailleurs confirmée par le fait que 71,4 % des conducteurs ayant plus de 10 ans d'expérience occupent le même poste depuis six ans et plus, comme on peut le constater dans le tableau 4.

Tableau 3 Nombre d'entreprises pour lesquelles le conducteur a travaillé au cours des 10 dernières années (en %)

	Ensemble des conducteurs (n = 512)	Conducteurs ayant 10 ans et plus d'expérience (n = 334)
Une	43,6 %	44,6 %
Deux	26,0 %	26,6 %
Trois	16,2 %	15,3 %
Quatre	6,1 %	6,0 %
Cinq ou plus	8,2 %	7,5 %

Selon le principal secteur d'activité, l'enquête auprès des conducteurs révèle que les conducteurs qui se concentrent sur le transport international se démarquent ici des trois autres marchés avec une plus grande mobilité qui les amène à travailler dans une entreprise de plus en moyenne que les conducteurs sur le marché local et sur le marché interrégional, les conducteurs sur le marché interprovincial se situant à mi-chemin. Dans le même sens, il y a environ deux fois moins de conducteurs sur le marché international qui occupent le même emploi depuis six ans et plus que pour les conducteurs des autres marchés.

Tableau 4 Depuis combien de temps le conducteur occupe l'emploi actuel (en %)		
	Ensemble des conducteurs (n = 514)	**Conducteurs ayant plus de 10 ans d'expérience (n = 336)**
Moins d'un an	14,0 %	9,2 %
De 1 à 3 ans	20,0 %	12,2 %
De 4 à 5 ans	11,7 %	7,1 %
6 ans et plus	54,3 %	71,4 %

On a aussi regardé la principale raison d'un départ lorsqu'un conducteur quitte un emploi, en présentant les résultats pour les deux derniers emplois, s'il y a lieu, occupés par les conducteurs, et pour les trois derniers emplois, s'il y a lieu, occupés par ceux qui sont maintenant propriétaires-exploitants. Les conditions de travail sont citées en premier (24,6 %) par les conducteurs puis viennent la fermeture, la vente ou la fin du contrat (20,0 %) ou d'autres raisons (20,0 %). Le salaire n'est cité comme principale raison que par 14,4 % des conducteurs. Pour les PEVL1&2 la première raison est le désir de partir sa propre entreprise (32,0 %), puis viennent les conditions de travail (16,2 %), d'autres raisons (14,2 %) et le salaire qui est cité comme principale raison dans seulement 10,7 % des cas. Contrairement à ce qui a été observé au niveau du recrutement, les conditions de travail sont donc plus importantes que le salaire lorsqu'il s'agit d'expliquer le départ d'un emploi.

On a par ailleurs analysé la relation qui peut exister entre les moyens mis en place ou non pour garder ses conducteurs et la difficulté pour l'entreprise de recruter des conducteurs. On peut ici faire deux hypothèses contraires : soit d'une part que la mise en place de certains de ces moyens diminue les difficultés rencontrées par le transporteur ou, d'autre part, que ces moyens sont mis en place lorsque l'entreprise rencontre des difficultés de recruter des conducteurs. Il semble que ce soit la deuxième explication qui prime

et que les moyens mis en place le sont plutôt en réaction aux difficultés rencontrées lors des tentatives de recrutement. En effet, parmi les neuf moyens possibles, quatre montrent une différence significative statistiquement et ce sont toujours les entreprises qui ont de la difficulté à embaucher des conducteurs qui sont plus nombreuses à utiliser ces moyens. Ainsi, pour les augmentations de salaire, ceux qui ont de la difficulté à recruter des conducteurs sont 40,9 % à utiliser ce moyen, alors que ceux qui n'ont pas cette difficulté représentent 33,6 %. Les autres résultats significatifs sont :

- diminution du nombre d'heures travaillées : 5,1 % vs 3,0 %
- affectation exclusive du véhicule : 14,3 % vs 10,1 %
- prime au rendement : 9,7 % vs 6,2 %.

Finalement, le tableau 5 étudie la relation entre les moyens mis en place par les transporteurs et les attentes mentionnées par les conducteurs qui se présentent chez eux. On a inclus dans ce tableau uniquement les différences significatives dans l'utilisation des différents moyens entre ceux qui ont identifié ou non chacune des attentes mentionnées par les conducteurs. Ainsi, la régularité des horaires est un moyen utilisé par 31,7 % des entreprises qui ont identifié la stabilité des horaires comme l'une des attentes mentionnées par les conducteurs, alors que ce moyen n'est utilisé que par 19,2 % des entreprises qui n'ont pas identifié cette attente. Il ressort de ce tableau la cohérence entre les moyens utilisés par les entreprises et les attentes mentionnées par les conducteurs. Ainsi, par rapport aux deux principales attentes que sont le salaire et la stabilité des horaires, les moyens utilisés sont, respectivement, l'augmentation du salaire et la régularité des horaires. Au niveau du nombre d'heures travaillées et des avantages sociaux, les moyens utilisés sont tout aussi cohérents. Lorsque la présence au domicile est une attente des conducteurs recrutés, la réponse des employeurs comprend plusieurs moyens utilisés en combinaison. On trouve alors trois moyens où il y a une différence significative : affectation exclusive du véhicule, prime au rendement et autres moyens.

Tableau 5 Moyens mis en place afin de maintenir au travail les conducteurs et attentes mentionnées par les conducteurs (en %)

	Moyens mis en place (n = 1409)	
	Oui	**Non**
Attente : stabilité des horaires		
Moyens : - régularité des horaires[1]	31,7	19,2
- changement de poste[2]	2,1	0,7
Attente : salaire		
Moyen : - augmentation du salaire[1]	43,7	30,1
Attente : technologie à bord du véhicule		
Moyens : - augmentation du nombre d'heures travaillées[2]	8,0	3,9
- affectation exclusive du véhicule[1]	23,0	12,3
- augmentation du salaire[2]	48,7	38,5
- aucun moyen[1]	10,6	21,5
Attente : présence au domicile		
Moyens : - affectation exclusive du véhicule[1]	23,0	10,8
- prime au rendement[1]	15,3	7,6
- autres moyens[1]	55,1	46,4
- aucun moyen[1]	13,5	22,3
Attente : nombre d'heures travaillées		
Moyens : - diminution du nombre d'heures travaillées[1]	6,5	3,5
- augmentation du nombre d'heures travaillées[1]	6,3	2,9
- affectation exclusive du véhicule[2]	10,5	15,0
Attente : avantages sociaux		
Moyens : - bonification de retraite si report de celle-ci[1]	6,1	1,9
- prime au rendement[1]	14,7	8,2

[1] Différence significative à un niveau de 1 %.
[2] Différence significative à un niveau de 5 %.

La satisfaction au travail

Le tableau 6 présente le taux de satisfaction au travail des conducteurs par rapport à 12 dimensions de cette satisfaction. Ce taux est mesuré sur une échelle de 10, 1 signifiant pas du tout satisfait et 10 signifiant tout à fait satisfait. Le score moyen pour les douze dimensions est de 7,83. Les trois facteurs qui dominent ce classement sont reliés à la dimension temporelle (exigences de l'employeur quant au respect des lois de la route, temps alloué pour compléter les voyages et le nombre de soirs passés au domicile) et ont des scores supérieurs à 8,42, alors que les conditions salariales se classent nettement moins bien, en fait 7e, avec un score de 7,88. On peut aussi noter que le plus grand niveau d'insatisfaction, et avec une bonne marge sur les autres dimensions, est la façon dont on consulte le conducteur avant l'achat de nouveaux équipements.

Une analyse par principal secteur d'activité du conducteur nous apprend que pour les conducteurs qui sont surtout actifs dans le transport local et le transport interprovincial, leur taux de satisfaction moyen est près d'un demi-point supérieur au taux des conducteurs oeuvrant dans le transport interprovincial et le transport international. Par contre, il faut nuancer ces résultats car les différences ne sont significatives qu'à un seuil de signification de 5 % et plus pour trois de ces dimensions : le nombre de soirs que l'employeur permet de passer au domicile, la façon dont on consulte le conducteur sur l'achat de nouveaux équipements et les conditions salariales. Pour la première dimension, on note que la satisfaction décroît avec l'éloignement du principal secteur d'activité, alors que pour les deux autres dimensions, les conducteurs actifs sur les marchés local et interrégional sont plus satisfaits que ceux actifs sur les marchés interprovincal et international.

Plusieurs autres variables ont été examinées pour tenter d'expliquer le degré de satisfaction des conducteurs. La variable qui a un effet sur le plus grand nombre de dimensions de la satisfaction est le type de rémunération. Il ressort clairement qu'une rémunération fixe à chaque semaine améliore la satisfaction et que cela se vérifie pour quatre

dimensions de cette satisfaction : le nombre de soirs passés au domicile, la fréquence des rencontres avec l'employeur, la qualité des rencontres avec l'employeur et la façon dont le conducteur est consulté avant l'achat de nouveaux équipements. À l'inverse, la rémunération à la distance diminue la satisfaction sur les quatre mêmes dimensions et sur une cinquième, soit les conditions salariales. La rémunération à l'heure diminue aussi la satisfaction sur trois dimensions, mais elle l'augmente sur une autre dimension.

Tableau 6 Satisfaction au travail des conducteurs face à leur employeur (score moyen) (n entre 486 et 514)	
La qualité de l'équipement qu'il met à votre disposition	8,41
Le temps qu'il vous alloue pour compléter votre voyage	8,54
Le temps de repos qu'il vous alloue entre les voyages	8,37
Le nombre de soirs qu'il vous permet de passer à votre domicile	8,42
La formation qu'il vous offre	7,02
La sécurité d'emploi qu'il vous offre	8,07
La fréquence des rencontres avec lui	7,43
La qualité des rencontres avec lui	7,73
Les exigences de votre employeur au sujet du respect des lois de la route	8,82
La façon dont on vous consulte avant l'achat de nouveaux équipements	5,58
Son ouverture face aux nouvelles technologies	7,69
Les conditions salariales qu'il vous offre	7,88
Moyenne	7,83
Note : sur une échelle de 10, 1 signifiant pas du tout satisfait et 10 signifiant tout à fait satisfait.	

Le fait d'être syndiqué augmente aussi de façon significative la satisfaction par rapport aux conditions salariales, mais cela diminue la satisfaction sur trois autres dimensions : la fréquence des rencontres avec l'employeur, la qualité des rencontres avec l'employeur et la

façon dont on consulte le conducteur avant l'achat de nouveaux équipements. Les chauffeurs les moins expérimentés sont moins satisfaits de la qualité de l'équipement que l'on met à leur disposition. Les chauffeurs les plus jeunes et les plus âgés sont par ailleurs plus satisfaits que les autres conducteurs de la formation que leur employeur leur offre. Finalement, le nombre d'heures de travail par semaine, lorsqu'il dépasse le seuil de 50 heures par semaine, diminue la satisfaction par rapport au temps de repos accordé entre les voyages, au nombre de soirs passés au domicile, aux exigences de l'employeur relatives au respect des lois de la route et aux conditions salariales.

La question a aussi été posée dans l'optique d'une comparaison avec d'autres groupes de chauffeurs, en leur demandant une évaluation de leurs conditions de travail sur une échelle de 10, 1 signifiant des conditions beaucoup moins avantageuses et 10 signifiant des conditions beaucoup plus avantageuses (voir le tableau 7). L'évaluation que font les conducteurs par rapport aux autres conducteurs est relativement positive avec un score de 7,99 pour l'ensemble des conducteurs. Les propriétaires-exploitants ayant 1 ou 2 véhicules sont beaucoup moins positifs avec un score de 5,98 lorsqu'ils se comparent aux conducteurs employés et de 6,21, ce qui n'est guère mieux, lorsqu'ils se comparent aux autres propriétaires-exploitants comme eux.

Tableau 7 Évaluation des conditions de travail par rapport aux autres conducteurs et propriétaires-exploitants (score moyen sur une échelle de 10)	
Évaluation des conducteurs par rapport aux conditions de travail des autres conducteurs	7,99 (n=499)
Évaluation des PEVL 1&2 par rapport aux conditions de travail des autres conducteurs	5,98 (n=645)
Évaluation des PEVL 1&2 par rapport aux conditions de travail des autres PEVL1&2	6,21 (n=615)

On a également analysé les relations entre cette satisfaction par rapport aux autres conducteurs et d'autres variables. Ainsi, au niveau

des conducteurs, leur âge semble avoir un effet, les plus jeunes (20 à 29 ans) étant beaucoup plus satisfaits que les autres, les plus insatisfaits étant les conducteurs entre 30 et 49 ans. Au niveau de la rémunération, un seul type a un effet statistiquement significatif, soit la rémunération selon la distance parcourue qui amène une baisse de cette satisfaction avec un score moyen de 7,50 pour ceux qui sont rémunérés ainsi, versus 8,15 pour ceux qui ne sont pas rémunérés de cette façon. L'augmentation du nombre d'heures travaillées par semaine a aussi un effet marqué puisque le fait d'avoir de plus longues semaines de travail amène une baisse d'environ un point sur l'échelle de satisfaction entre ceux qui font 51 heures et plus et ceux qui font 50 heures et moins. Par contre, le nombre de nuits par mois passées au domicile n'a pas un effet statistiquement significatif sur cette satisfaction relative. Si on fait une distinction sur la base de l'appartenance à un syndicat ou une association, cela se traduit pour les conducteurs appartenant à un syndicat par un taux de satisfaction par rapport aux autres conducteurs qui est plus élevé, soit 8,20 versus 7,86. Pour les propriétaires-exploitants, la relation est inversée puisque ceux qui ne sont pas membres d'une association évaluent plus positivement leurs conditions que ce soit face aux conducteurs ou aux autres propriétaires-exploitants.

Au niveau des propriétaires-exploitants, il y a un plus grand nombre de variables qui ressortent, surtout lorsqu'il s'agit de comparer leurs conditions de travail avec les conducteurs salariés. En effet, dans ce cas, la rémunération à la semaine augmente la satisfaction alors que la rémunération à la distance parcourue et la rémunération au poids diminuent la satisfaction de façon statistiquement significative. Le nombre d'heures travaillées a aussi un effet analogue à celui mesuré chez les conducteurs, le seuil de 50 heures par semaine marquant aussi une limite qui provoque une baisse de la satisfaction, quoique dans leur cas ceux qui travaillent entre 51 et 60 heures se trouvent dans une situation intermédiaire. Le nombre de nuits par mois passées au domicile n'a pas ici non plus d'effet significatif sur la satisfaction, et leur âge non plus. Si on regarde maintenant leur satisfaction relative aux autres propriétaires-exploitants, il ressort que la rémunération à la semaine augmente la satisfaction alors que les

rémunérations au poids et à la longueur la diminuent. Le nombre d'heures travaillées par semaine n'a pas ici d'effet marqué statistiquement sur la satisfaction mais, par contre, le nombre de nuits passées au domicile a un effet marqué, particulièrement chez ceux qui passent 8 nuits ou moins à la maison dont la satisfaction perd un point sur l'échelle quand on les compare aux autres répondants.

Conclusion

Le problème de recrutement de conducteurs est un problème réel qui a une acuité qui peut certes varier selon le principal secteur d'activité, la taille de l'entreprise ou le type d'entreprise, mais qui semble somme toute assez généralisé. Cela ne veut évidemment pas dire qu'il s'agit-là de l'unique problème rencontré par l'industrie, la concurrence sur les tarifs reliée au fort niveau de concurrence qui caractérisent le secteur, demeurant encore une grande préoccupation. Les entreprises ont commencé à agir face au problème de recrutement, mais il semble que les moyens d'actions sont d'abord mis en place en réaction aux problèmes rencontrés plutôt que de façon proactive, ce qui pourrait éviter ces problèmes.

On constate aussi que le problème de recrutement et de maintien des conducteurs en emploi demande des interventions portant sur plusieurs dimensions, ce qui rend plus complexe la gestion des ressources humaines. Les entreprises, et plus particulièrement les petits et moyens transporteurs, semblent alors particulièrement vulnérables à une détérioration de l'offre de conducteurs qualifiés. Dans ce sens, les procédures et techniques de recrutement de nouveaux conducteurs reposent encore très souvent sur des processus informels qui cadrent mal avec les exigences accrues dans la profession. Il faut aussi souligner la problématique particulière des jeunes conducteurs qui ont besoin de supports spécifiques compte tenu de leur manque d'expérience et de la nécessité de diminuer la migration de jeunes conducteurs vers d'autres emplois. Les jeunes conducteurs qui oeuvrent sur le marché international devraient recevoir une attention particulière.

Encore sur la route : profil des camionneurs au Canada*

Par Vincent Dubé *et* Denis Pilon

L'industrie du camionnage est actuellement confrontée à un besoin croissant de camionneurs qualifiés. Parmi les conséquences négatives pouvant découler de cette situation, on retrouve une perte de compétitivité découlant des coûts excessifs de recrutement et de formation, des interruptions dans les horaires de livraison et de la sous-utilisation de l'équipement (Min et Lambert, 2002).

La capacité à recruter et à retenir des camionneurs qualifiés est devenue un défi de premier plan pour l'industrie canadienne du camionnage. Cet article vise ainsi à dresser un portrait récent de la profession de camionneur au Canada à partir de différentes sources de données. La première partie propose une analyse de l'offre et la demande de camionneurs. La deuxième partie traite plus explicitement de leurs caractéristiques socio-économiques et des caractéristiques reflétant le marché du travail.

* : Notez qu'il s'agit d'une version réduite d'un article à paraître.

L'offre de camionneurs

Selon les données de l'enquête sur la population active (EPA), environ 271 000 personnes travaillaient en tant que camionneurs en 2004, soit 12 % de plus qu'en 1998. Globalement, cela représente une croissance de l'emploi inférieure à celle de l'ensemble des autres professions qui a été de 14 % pour la même période.

Croissance de la part du camionnage pour compte d'autrui

Il y avait environ 166 000 camionneurs dans le segment du camionnage pour compte d'autrui en 2004, ce qui représente une augmentation de 21 % avec 1998. D'autre part, on retrouvait 106 000 camionneurs dans le segment du camionnage privé, une augmentation

de tout juste 1 % pour la même période. C'est donc dire qu'avec 61 % des camionneurs en 2004, la part du camionnage pour compte d'autrui a vraisemblablement prit de l'importance au détriment du camionnage privé au cours des dernières années. Ce résultat semble ainsi refléter une hausse dans le recours à l'impartition amorcé à la suite de la récession du début des années 1990.

Environ quatre camionneurs sur cinq étaient chauffeurs de compagnie en 2004

En 2004, les chauffeurs de compagnie représentaient près de 80 % de tous les camionneurs, contre près de 20 % pour les camionneurs artisans. Au sein du camionnage privé, on retrouvait 97 % de chauffeurs de compagnie, alors que cette proportion était de 67 % pour le camionnage pour compte d'autrui.

Le nombre des chauffeurs d'agence est en croissance chez le camionnage pour compte d'autrui

Selon l'Enquête trimestrielle sur les transporteurs routiers de marchandises pour compte d'autrui (ETTRM), le nombre de chauffeurs d'agence était estimé à environ 5 400 en 2003. Bien que leur nombre soit en progression constante depuis près de 10 ans (augmentation de 58 % entre 1994 et 2003), ils représentent une proportion relativement stable, avec 7 % du total des chauffeurs du segment.

Le transport de marchandises générales représente le principal champ d'activité des camionneurs pour compte d'autrui

Globalement, environ 63 % des camionneurs pour compte d'autrui transportaient principalement de la marchandise générale en 2003. Celle-ci était ensuite suivi du transport des autres marchandises spécialisées (voitures, animaux d'élevage, etc.) avec 17 %, le transport des liquides en vrac et le transport des produits secs en vrac avec respectivement 7 % des camionneurs.

Prédominance du camionnage sur de longues distances chez le camionnage pour compte d'autrui

On observe également que près de quatre camionneurs sur dix en 2003 se spécialisaient principalement dans le camionnage sur de longues distances (78 % contre 22% pour le camionnage local)[1].

La demande de camionneurs

La croissance économique générale de même que la croissance des ventes du commerce de gros et de détail agissent comme catalyseurs de la demande des services de frets. L'augmentation du volume de biens produits et consommés se traduit par une demande accrue des services de transports de fret, de même que par l'augmentation du nombre des livraisons et des distances parcourues. En fait, de 1998 à 2003, les entreprises de camionnage pour compte d'autrui domiciliées au Canada ont vu leurs recettes progressées de plus de 50 % pour atteindre 16,8 milliards de dollars en 2003. Le tonnage transporté par ces entreprises a augmenté de 30,5 %, le nombre d'expéditions de 19,1 % et la distance moyenne par livraison de 2,3 %.

L'observation de la croissance comparative de plusieurs secteurs économiques durant la période de référence témoigne de cette concordance, puisque le secteur de la fabrication enregistre une croissance de 17,1 %, le commerce de gros 31,3 %, le commerce de détail 27,4 %, alors que le camionnage a cru de 28,3 %.

Sous l'angle des recettes, la croissance du secteur du camionnage se traduit surtout par la croissance du transport transfrontalier. Le camionnage étant le principal moyen de transport pour le transport des marchandises entre le Canada et les États-Unis. En 2003 l'industrie du camionnage a acheminé 53 % des exportations canadiennes aux États-Unis et 79 % des importations canadiennes en provenance des États-Unis. Pour ce qui est des livraisons, les

[1] Notez qu'un transporteur routier peut opérer à la fois des activités se rapportant au camionnage local et sur de longues distances.

entreprises de transport routier pour compte d'autrui spécialisées dans le transport sur de longues distances ont effectué au total 31,6 millions d'expéditions intérieures représentant 219,8 millions de tonnes et 89,0 milliards de tonnes-kilomètres[2]..

La ventilation des activités selon l'origine et la destination des livraisons révèle également l'impact grandissant du trafic transfrontalier. Les recettes du trafic transfrontalier sont passées de 42,1 % en 1998 à 47,5 % en 2003. Pour cette même période, la croissance du fret transfrontalier s'élève à plus de 71 %, soit plus du double de celle enregistrée par les activités de transport domestique.

Caractéristiques socio-économiques

Une profession typiquement masculine

Selon l'EPA, la profession de camionneur était occupée par une proportion de 98 % d'homme en 2004 (voir tableau 1 en annexe). Elle demeurait également la profession la plus populaire chez ce groupe selon les données du Recensement de 2001.

Moins de familles avec deux soutiens économiques chez les camionneurs

Selon les données de l'Enquête sur la dynamique du travail et le revenu (EDTR), environ 70 % des conjoints issus des familles époux-épouse dont la profession de camionneurs était le principal soutien économique avaient un travail rémunéré en 2002. Cette proportion est

[2] Camionnage sur longues distances seulement (ETTRM). Ces chiffres ne sont cependant pas révélateurs de l'activité de l'industrie du camionnage puisque les transporteurs du segment du camionnage privé, les petits transporteurs (ayant un revenu d'exploitation annuel de moins d'un million de dollars) et les transporteurs locaux pour compte d'autrui sont exclus de ces données

en deçà de celle que l'on retrouvait chez l'ensemble des professions (82 %).

De faible niveau de scolarité

Selon l'EPA, les camionneurs présentaient une plus faible scolarité que les travailleurs de l'ensemble des professions en 2004. Par exemple, plus du tiers d'entre eux n'avaient pas obtenu leur diplôme secondaire, contre seulement 15 % pour l'ensemble des professions. Le fait que la profession ne nécessitait aucun diplôme spécifique, outre le fait d'obtenir le permis de conduire correspondant au type de camion utilisé, est sans doute l'une des raisons expliquant cette situation[3]. Or, tel que le souligne Nix (2003), les formations que suivent les camionneurs ne se reflètent pas toujours au niveau de la scolarité. De fait, bien que certains cours puissent donner droit à des accréditations reconnues, ce n'est pas toujours le cas. Les formations offertes à l'interne par l'employeur en sont des exemples types.

Une population vieillissante et peu de relève chez les jeunes

Globalement, les camionneurs présentent une population qui est relativement plus âgés que l'ensemble des professions. En effet, environ 46 % des camionneurs avaient plus de 45 ans en 2004, contre 37 % pour l'ensemble des professions.

L'industrie du camionnage risque donc à terme de perdre un nombre important de camionneurs à un rythme plus rapide que les autres industries. Toutefois, une étude américaine a démontré que le taux de roulement des camionneurs plus âgés et/ou plus expérimentés tendent à être plus faibles, compte tenu qu'un changement d'emploi implique de plus grands risques (Min et Lambert, 2002). Les résultats d'une enquête auprès des conducteurs de camion sans emplois vont un peu dans le même sens puisqu'il est observé que les jeunes conducteurs

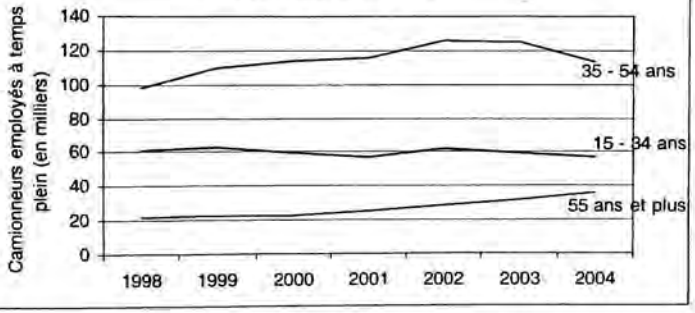
[3] Il existe toutefois des exigences supplémentaires pour le transport des marchandises dangereuses.

sont plus susceptibles de démissionner que les conducteurs plus âgés (CCRHC, 2003).

Une question toute aussi importante est celle du peu de relève observé chez les jeunes travailleurs au sein de la profession. De fait, on retrouvait seulement 5 % du total des travailleurs qui étaient âgés de moins de 25 ans en 2004, soit une proportion environ trois fois plus faible que ce que celle observée chez les travailleurs de l'ensemble des professions. De plus, le groupe des travailleurs âgés entre 15 et 34 ans représentait à peine 26 % des camionneurs alors que cette proportion était de 37 % chez l'ensemble des professions.

Il semble aussi que ce phénomène s'est progressivement amplifié depuis 5 ans. Par exemple, bien que le nombre total de camionneurs employés à temps plein ait connu une augmentation de 16 % entre 1998 et 2004, le nombre de camionneurs employés à temps plein âgés entre 15 et 34 ans n'a pas réellement augmenté durant cette même période (voir *figure 1*).

Figure 1: Le nombre de camionneurs âgés entre 15 et 34 ans employés à temps plein est en stagnation depuis plusieurs années au Canada (source: EPA)



Il existe plusieurs raisons pouvant expliquer cette situation. D'une part, la législation reliée à l'âge pour conduire un camion joue sans

aucun doute un certain rôle. Au Canada, l'âge minimum pour obtenir un permis de conducteur de véhicules commercial varie entre 18 et 20 ans selon les provinces, alors qu'aux États-Unis il est de 21 ans dans une majorité d'États.

D'autre part, ce n'est pas parce qu'une personne respecte tout juste l'âge minimum requis qu'elle peut exercer du jour au lendemain la profession. Dans la pratique, l'âge minimum demandé par les compagnies de transport est souvent plus élevé que celui déterminé par la législation. Bien que cette situation peut également être reliée aux coûts beaucoup plus prohibitifs en terme d'assurances, elle résulte principalement de l'importance accordée à l'expérience pratique de conduite dans la profession, une dimension fortement corrélée avec l'âge. De fait, l'expérience est sans aucun doute l'élément majeur dans la prise de décision d'embauche d'un camionneur[4]. Globalement, le nombre d'années d'expérience minimal exigé par les gestionnaires lorsqu'ils embauchent un camionneur pour conduire un camion semi-remorque est de trois ans (CCRHC, 2003).

Croissance du nombre de camionneurs venant de l'extérieur du Canada

Selon le Recensement de 2001, on retrouvait une proportion de 13 % de travailleurs immigrants parmi les camionneurs inclus dans la population active, en hausse de 88 % avec le Recensements de 1991. De même, environ le tiers de ces camionneurs demeuraient au Canada depuis moins de 10 ans en 2000, alors que cette même proportion était de seulement 19 % en 1990.

Caractéristiques du marché du travail

Temps plein, faible taux de syndicalisation et PME

[4] Le premier facteur pris en compte par les gestionnaires de parc lors de l'embauche de chauffeurs est l'expérience, suivi du dossier de conduite CCRHC (2003).

On retrouve généralement très peu de camionneurs qui travaillent à temps partiel dans l'industrie. De fait, selon l'EPA, environ 97 % des camionneurs travaillaient à temps plein en 2004.

Contrairement à ce que l'opinion populaire peut laisser croire, seulement le quart des chauffeurs de compagnie étaient syndiqués en 2004, un taux inférieur à celui observé chez l'ensemble des employés en 2004 (32 %). De même, les chauffeurs de compagnie issus du segment du camionnage pour compte d'autrui présentaient un taux de syndicalisation de près de 10 % inférieur à leurs confrères travaillant pour le camionnage privé.

Environ 78 % des chauffeurs de compagnie travaillaient pour des entreprises de moins de 100 employés (PME) en 2004. Cette proportion était encore plus forte chez le segment du camionnage pour compte d'autrui où quatre chauffeurs sur cinq étaient employés d'une PME. À titre de comparaison, cette part était plutôt des deux tiers chez les employés de l'ensemble des professions.

Un salaire légèrement supérieur à la moyenne

Selon les chiffres de l'EPA, un chauffeur de compagnie travaillant à temps plein gagnait un salaire hebdomadaire moyen de 791 $ en 2004, ce qui représente environ 41 400 $ par an en supposant qu'il travaille pendant toute l'année. Il gagnait ainsi un salaire légèrement supérieur à l'ensemble des salariés qui était de 778 $ par semaine (ou 40 500 $ par an selon la même hypothèse). Par ailleurs, ceux qui étaient syndiqués gagnaient en moyenne 11 % de plus que leurs confrères qui ne l'étaient pas.

Le salaire hebdomadaire moyen était légèrement plus élevé chez le segment du camionnage pour compte d'autrui avec 829 $ par semaine (43 100 $ par an). Bien que l'on ne dispose pas de donnée sur le salaire moyen pour le camionnage privé, on en déduit donc que les camionneurs de ce segment gagnaient en moyenne un salaire légèrement inférieur.

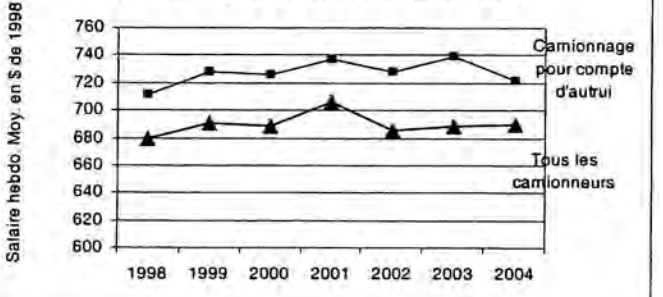
Les salaires en fonction des heures travaillées

L'un des principaux déterminants des salaires des camionneurs est le nombre d'heures travaillées. Par exemple, les chauffeurs de compagnie qui travaillaient 70 heures de travail et plus par semaine gagnaient en moyenne un salaire de 242 $ par semaine de plus que le salaire moyen. Or, malgré le fait que les heures supplémentaires sont souvent rémunérées à un taux plus élevé (notamment dans le transport local de marchandise), on note plutôt que les gains marginaux découlant des heures supplémentaires sont généralement décroissants. À ce sujet, mentionnons qu'une récente étude américaine (Monaco et Willmert, 2003), a observé que de plus faibles taux de salaire hebdomadaire des camionneurs sont associés à des probabilités plus élevées d'enfreindre le maximum d'heures hebdomadaires permis selon les règles sur les heures de services.

Le salaire réel des chauffeurs de compagnie en stagnation

Lorsque l'on observe l'évolution du salaire hebdomadaire moyen des camionneurs depuis 1998, on est à même de constater que celui-ci n'a pas réellement évolué depuis cette date. De fait, celui-ci a seulement augmenté de 2 % en 6 ans (voir *Figure* 2).

Figure 2: Les salaires réels des camionneurs employés à temps plein ont stagnés de 1998 et 2004 (source: EPA)



Par ailleurs, la séniorité semble jouer un rôle moins important que chez d'autres professions. En effet, s'il est vrai qu'il existe certaines barrières à l'entrée pour les travailleurs sans expérience dans la profession, leur salaire semble connaître un plafonnement relativement rapide. Par exemple, en comparant le groupe de travailleurs les moins expérimenté (les 20-24 ans) avec celui des travailleurs les plus expérimentés (les 55 ans et plus) chez le camionnage pour compte d'autrui, on y note une différence de salaire hebdomadaire de moins de 100 $, contre près de 350 $ chez l'ensemble des salariés.

Les chauffeurs de compagnie bénéficient de moins d'avantages sociaux

Les avantages sociaux, comme les régimes d'assurance et de retraite parrainés par l'employeur, constituent d'autres indicateurs de la qualité des emplois (Marshall, 2003). Selon les données de 2002 de l'EDTR, les chauffeurs de compagnie bénéficiaient en plus faible proportion d'avantages sociaux que les travailleurs de l'ensemble des professions. Cette situation est particulièrement apparente au niveau du régime de retraite où seulement 28 % des camionneurs participaient à un régime de retraite lié à leur emploi, contre 44 % chez les travailleurs de l'ensemble des professions (voir tableau 2).

Tableau 2: Avantages sociaux, 2002	Ensemble des professions	Chauffeurs de compagnie
Avantages sociaux offerts par l'employeur	%	
Régime de soins dentaire	60.0	54.3
Régime d'assurance vie/invalidité	59.8	55.3
Régime d'assurance médical	65.0	60.8
Régime de soins dentaire ou vie/invalidité ou médical	67.4	65.2
Régime de retraite	43.5	28.1

Source: Enquête sur la dynamique du travail et du revenu

Nota: Personnes ayant un emploi rémunéré. Lorsque les personnes occupent plus d'un emploi; on se sert de l'emploi principal.

Peu de différences selon les catégories de travailleurs

Selon les données du Recensement de 2001, il ne semble pas y avoir d'importantes différences entre les catégories de travailleurs. Cela dit, on remarque malgré tout que les camionneurs artisans gagnaient en 2000 un revenu annuel d'emploi légèrement plus élevé que les chauffeurs de compagnie (39 200 $ contre 38 800 $)[5]. Rappelons toutefois que ces travailleurs ne peuvent bénéficier d'avantages non-salariaux contrairement à leurs confrères salariés. Conséquemment, les travailleurs autonomes (camionneurs artisans) sont moins susceptibles d'être couverts par des régimes complémentaires d'assurance de soins médicaux, de soins dentaires et d'invalidité. De même, n'étant pas admissibles aux prestations de maladie, ils sont plus susceptibles de subir des pertes financières plus grandes s'ils cessent de travailler pour cette raison (Akyeampong et Sussman, 2003). De surcroît, étant donné qu'ils ne peuvent profiter d'un régime de pension offert par leur employeur, ils se doivent également d'économiser davantage pour leur retraite.

De longues semaines de travail

Selon l'EPA, les chauffeurs de compagnies travaillaient en moyenne 47 heures par semaines en 2004 et 39 % d'entre eux travaillaient plus de 50 heures hebdomadairement. De même, la semaine de travail des camionneurs issus du camionnage pour compte d'autrui était encore plus longue. Celle-ci était en moyenne de 50 heures et elle était de 60 heures et plus pour le quart des chauffeurs. Grosso modo, ce résultat s'explique par le fait que l'on retrouve une plus grande proportion de camionneurs spécialisés dans le camionnage sur de longues distances au sein de ce segment.

Or, comme d'autres l'ont fait remarquer (Best 1999, Nix 2003), ces chiffres peuvent représenter une sous-estimation des heures réelles de travail des camionneurs de longues distances. Ainsi, compte tenu

[5] Camionneurs ayant travaillé à temps plein toute l'année.

qu'ils représentent les heures habituellement rémunérées ou à contrats, ils reflètent peut-être davantage les heures officielles inscrites dans les carnets de route, qui excluent souvent une bonne partie du temps de non conduite qui est malgré tout travaillé. Les périodes d'attente, les inspections mécaniques et les vérifications aux douanes sont des exemples typiques d'activités souvent comptabilisées aux carnets de route comme des périodes de repos. Cette façon de faire permet aux camionneurs, qui sont majoritairement payés en fonction de la distance parcourue, de préserver leurs heures de services pour la conduite et ainsi, maximiser leurs gains (Nix, 2003).

Une profession pouvant être stressante

Étant donné la relation qui peut exister entre de longues heures de travail, les horaires de travail irrégulier et le niveau de stress, il est assez intuitif de croire que les camionneurs peuvent s'avérer particulièrement vulnérables à ce niveau. De fait, selon les chiffres de l'EDTR, on observe qu'un peu plus de sept camionneurs sur dix ont déclaré que leur vie professionnelle et personnelle leur créait beaucoup ou un peu de stress en 2002[6].

De plus, de longues heures de travail auraient un effet néfaste sur la santé. Ils entraîneraient ainsi des modifications du mode de vie malsaines pour la santé, tel le manque d'exercice physique, le manque de sommeil, les mauvaises habitudes alimentaires, etc. (Shields, 2000). Il n'est donc pas surprenant de noter que les camionneurs présentaient l'un des taux d'inactivité pour des raisons de santé ou d'incapacité reliés à la maladie le plus élevé chez les hommes avec 3.9 %, contre 2.6 % pour l'ensemble des professions en 2004[7]. De même, chaque camionneur avait perdu l'équivalent de 10

[6] Personnes de 16 ans et plus dont l'emploi principal rémunéré était camionneur (EDTR).

[7] Le taux d'inactivité est la proportion d'heures perdues par rapport au nombre total d'heures hebdomadaires habituellement travaillées par

jours pour ces mêmes raisons dans l'année, comparativement à 6 jours chez l'ensemble des travailleurs[8].

Résumé

Les camionneurs jouent un rôle de premier plan au sein de l'économie canadienne. Or, l'industrie du camionnage est désormais confrontée à un besoin croissant de camionneurs qualifiés suite à l'évolution et la forte croissance économique du secteur, mais aussi en raison de certaines caractéristiques de la profession. Par exemple, lorsque que l'on compare les 271 000 personnes qui exerçaient la profession de camionneurs au Canada en 2004 avec les travailleurs de l'ensemble des professions, on observe qu'ils présentent une proportion plus élevée de personnes n'ayant pas terminé leurs études secondaires et de personnes plus âgées, mais aussi une proportion plus faible de jeunes travailleurs. Par ailleurs, bien qu'ils obtenaient un salaire légèrement plus élevé que les travailleurs de l'ensemble des professions, ils travaillent en moyenne un nombre d'heures relativement plus élevé.

Sources de données

Il n'existe pas de source d'information complète unique disponible sur les camionneurs. Ainsi, pour les besoins de la présente étude, plusieurs sources de données ont été utilisées. Les données sont tirées de l'Enquête sur la population active (EPA), du Recensement de 2001, de l'Enquête dynamique du travail et le revenu (EDTR), de l'Enquête trimestrielle sur les transporteurs routiers de marchandises pour compte d'autrui (ETTRM). Notez que le fait que ses sources ne

---

tous les employés à temps plein. EPA, hommes employés à temps pleins seulement, exclusion du congé de maternité.

[8] On calcule le nombre de jours perdus par travailleur en multipliant le taux d'inactivité par le nombre estimatif de jours ouvrables dans l'année (250). EPA, hommes employés à temps pleins seulement, exclusion du congé de maternité.

visent pas exactement la même population entraîne une certaine imprécision. De même, puisqu'elles n'ont pas nécessairement la même année de référence, l'utilisation des données les plus récentes disponibles a été privilégiée dans le but de fournir une analyse la plus actuelle possible.

Documents cités

Akyeampong, Ernest B. et Deborah Sussman, « Régimes de santé des travailleurs indépendants», *L'emploi et le revenu en perspective*, n° 75-001-XPF au cat., vol. 15, n° 2, Statistique Canada, été 2003.

Bess, Irwin, « Le régime de travail des camionneurs », *L'emploi et le revenu en perspective*, n° 75-001-XPF au catalogue, vol. 11, n° 4, Statistique Canada, hiver 1999.

Conseil Canadien des Ressources Humaines en Camionnage, « Le moteur de l'économie canadienne », Ottawa, 2003.

Marshall, Katherine, « Les avantages de l'emploi », *L'emploi et le revenu en perspective*, n° 75-001-XIF au cat., vol. 04, n° 5, Statistique Canada, édition en ligne, mai 2003.

Min, Hockey et Thomas Lambert, "Truck Driver Shortage Revisited, *Transportation journal*, 22décembre 2002.

Monaco, Kristen et Joshua Willmert, « Wages and hours of truck drivers: an analysis across data sets », *Journal of the Transportation Research Forum*, published in *Transportation Quartely*, vol. 57, No. 3, Summer 2003 (167-177), Washington, D.C., 2003.

Nix, Fred P., « Le camionnage au Canada, le profil de l'industrie », Direction de la politique des transports routiers, Transport Canada, Ottawa, mars 2003.

Shields, Margot, « Les longues heures de travail et la santé », *L'emploi et le revenu en perspective*, n° 75-001-XPF au cat., vol. 12, n° 1, Statistique Canada, printemps 2000.

Annexe

Tableau 1: Caractéristiques des camionneurs	Toutes les professions	Camionneurs		
		Total	compte d'autrui	Privé
		milliers		
Emploi total	**15949.7**	**271.3**	**165.6**	**105.7**
		%		
Type de camionneur	100.0	100.0	100.0	100.0
Employé	84.6	78.7	67.2	96.6
Travailleur indépendant	15.4	21.3	32.8	3.4
Age	100.0	100.0	100.0	100.0
5 - 24 ans	15.4	5.4	3.8	7.9
25 - 54 ans	71.5	76.0	77.5	73.7
25 - 34 ans	21.7	20.1	21.2	18.4
35 - 44 ans	25.9	28.9	29.3	28.3
45 - 54 ans	23.9	27.0	27.0	27.1
55 ans et plus	13.1	18.5	18.3	18.9
Instruction	100.0	100.0	100.0	100.0
Études secondaires non complétées	14.4	33.9	34.5	32.9
Études secondaires complétées	20.3	27.0	24.5	30.9
Études postsecondaires partielles	10.0	8.9	9.5	7.9
Diplôme d'études postsecondaires	34.1	27.2	28.6	25.1
Grade universitaire	21.2	3.0	3.0	3.1
		milliers		
Employé à temps plein - Total	**11053.5**	**206.2**	**108.1**	**98.1**
		%		
Heures habituelles travaillées	100.0	100.0	100.0	100.0
Moins de 30.0 heures	15.6	10.1	10.4	9.9
30.0 à 39.9 heures	32.2	11.5	9.3	14.1
40.0 à 49.9 heures	41.6	39.9	33.0	47.4
50.0 à 59.9 heures	6.8	19.3	22.1	16.1
60.0 à 69.9 heures	2.5	12.0	15.6	8.0
70.0 heures et plus	1.3	7.2	9.5	4.7
Moyenne	39.5	47.3	50.0	n.d.
Salaire hebdomadaire		$		
Total	777.7	791.4	828.6	n.d.
Heures habituelles travaillées		$		
Moins de 30.0 heures	766.06	775.71	800.03	n.d.
30.0 à 39.9 heures	709.09	696.55	709.75	n.d.
40.0 à 49.9 heures	768.32	722.74	754.1	n.d.
50.0 à 59.9 heures	1002.22	830.16	832.73	n.d.
60.0 à 69.9 heures	1078.49	917.19	939.01	n.d.
70.0 heures et plus	1176.54	1032.63	1042.26	n.d.

Source: Enquête sur la population active, 2004

COMPARING HIGHWAY SPEED CONSISTENCY MEASURES USING ANALYTICAL METHOD

by Young-Jin Park
Research Assistant, Ph.D. Candidate
Department of Civil Engineering
University of Waterloo
200 University Avenue West
Waterloo ON Canada, N2L 3G1
Tel: 519-885-1211 ext.6596
Fax: 519-888-6197
Email: y2park@engmail.uwaterloo.ca

Submitted to the 40th Annual Conference of the Canadian Transportation Research Forum

ABSTRACT

The V85 speed reflects the 85th percentile speed of vehicles in a traffic stream passing a given element on the highway section. The large amount of speed differential in V85 between the two successive elements along a highway reflects lack of speed consistency in the highway section. This lack of speed consistency believes to be translated into increased crash risk. Recent research points to a critical problem associated with the conventional method of obtaining ΔV_{85} between successive highway elements: Failure to account for inter-dependence in individual vehicle speeds between successive elements. Vehicle speeds in the first element might be associated with vehicle speed in the second element. This problem has serious implications for safety treatment based on the conventional ΔV_{85} measure. Some researchers have suggested that the estimated speed differential based on individual vehicle speed profiles in successive elements is significantly higher than that obtained using the conventional ΔV_{85}. In this study, we assess the safety implications in ΔV_{85} using analytical method. Findings lead to important implications for establishing engineering treatments to improve safety in two-lane rural highways.

Keywords: Operating speed profile; Speed consistency; Road Safety.

INTRODUCTION

Operating speed profile models are used to establish consistency of speeds over successive elements of a highway. Consistency of speed is normally measured as the difference in the 85th percentile speeds (ΔV_{85}) observed between successive elements of the highway. In principle, safe highways are those highways that maintain consistency of vehicle speeds between the upstream and the downstream elements (e.g. a tangent/curved section).

Current operating speed models for estimating ΔV_{85} do not require prior specification of individual vehicle speed profiles over successive highway elements. Conventional ΔV_{85} estimation procedures assume that V85 at each element applies equally to all vehicles in the constituent traffic stream (Fitzpatrick et al. 2000; Krammes et al. 1995; Lamm et al. 1988; Leisch and Leisch 1977). In reality, however, we know that speeds will vary from vehicle to vehicle, and this leads to an under-estimation of the ΔV_{85} measure over successive elements (Hirsh 1987).

The practical implications of this problem as reflected in the conventional ΔV_{85} measure can be quite serious. On the basis of this measure, highway designers and managers may accept inconsistent and potentially unsafe highway elements, as being consistent and safe and hence not in need of any treatment. As a result, potentially hazardous problems and locations are unattended or ignored.

McFadden and Elefteriadou (2000) tried to validate Hirsh's hypothesis through the analysis of field data and the introduction of a new speed differential measure called 85MSR. This reflects the 85th percentile maximum speed reduction between two successive highway elements as experienced by the same vehicle or driver. McFadden and Elefteriadou found that the speed differential captured by the new measure is on average about two times greater than the speed differential obtained from conventional ΔV_{85} measure. Misaghi (2003) reached similar conclusions using practically the same measure as McFadden and Elefteriadou. Notwithstanding these findings, however, researchers have failed to address the issues as to why speed differential as measured by 85MSR is greater than the conventional ΔV_{85} measure.

This study conducted to compare differential speed profiles using ΔV_{85} and 85MSR measures for successive elements (tangent and curve) of a highway and assess the safety implications of using either procedure for establishing speed consistency along the highway.

ANALYTICAL COMPARISION OF ΔV_{85} AND 85MSR MEASURES

The aforementioned two speed consistency measures were defined by McFadden and Elefteriadou (2000) as follows, and Table 1 shows how we can obtain both measures.

- ΔV_{85} represents speed differential between the 85th percentile speed at the approach tangent and the 85th percentile speed at the mid-point of the horizontal curve (V85@approach tangent – V85@mid-point in curved section).
- 85MSR represents the 85th percentile of individual speed differentials and can be calculated by using the individual vehicle speed profile (i.e. paired speed dataset).

Table 1

Observed vehicle speed and speed differential along a tangent-curved section

Vehicle number	Tangent speed; *V1* (km/h)	Curve speed; *V2* (km/h)	Speed differential per each vehicle; ΔVi (km/h)
1	88	68	20
2	90	85	5
3	75	67	8
4	93	78	15
.	.	.	.
.	.	.	.
.	.	.	.
the 85th percentile value (km/h)	*V85* @ approach tangent section	*V85* @ mid-point in curved section	85MSR from individual vehicles' speed differential

Generally, previous researchers assumed that a speed data collected in a roadway follows a normal distribution. Therefore, a speed distribution from a tangent section (*V1*) should be;

$$V_1 \sim N(\mu_1, \sigma_1^2) \qquad \backslash(1)$$

where, μ_1 = mean speed at a tangent section

σ_1^2 = variance of speed at a tangent section

Also, a speed distribution from a curve section ($V2$) should be;

$$V_2 \sim N(\mu_2, \sigma_2^2) \qquad (2)$$

where, μ_2 =mean speed at a curve section.

σ_2^2 =variance of speed at a curve section

Let us presume that the $Z_{[0.85]}$ as of the value of standardized normal distribution corresponding to 85 percentile value. Then,

$$V_{1[0.85]} = \mu_1 + Z_{[0.85]} \times \sigma_1 \text{, and} \qquad (3)$$

$$V_{2[0.85]} = \mu_2 + Z_{[0.85]} \times \sigma_2 \qquad (4)$$

We want to calculate the difference in the 85th percentile speed between two locations. Assuming independence of vehicle speeds between successive elements, conventional ΔV_{85} can be expressed as:

$$\Delta V_{85} = V_{1[0.85]} - V_{2[0.85]} = (\mu_1 - \mu_2) + Z_{[0.85]} \times (\sigma_1 - \sigma_2) \qquad (5)$$

On the other hand, if we presume that the speeds between two consecutive sections are not independent, then the 85th percentile speed difference by the matched speed data (i.e. 85MSR) would be derived as follows;

First, the variance would be;

$$Var(V_1 - V_2) = \sigma_1^2 + \sigma_2^2 - 2Cov(V_1, V_2) \quad (6)$$

Therefore,

$$\Delta V = (V_1 - V_2) \sim N(\mu_1 - \mu_2, \sigma_1^2 + \sigma_2^2 - 2Cov(V_1, V_2)) \quad (7)$$

Let 85th percentile value in distribution of ΔV as 85MSR, then;

$$P(\Delta V < 85MSR) = 0.85 \quad (8)$$

If we standardize eq. (8), then;

$$P\left(\frac{\Delta V - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2Cov(V_1, V_2)}} < \frac{85MSR - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2Cov(V_1, V_2)}} \right) = 0.85 \quad (9)$$

$$P\left(Z < \frac{85MSR - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2Cov(V_1, V_2)}} \right) = 0.85 \quad (10)$$

$$\frac{85MSR - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2 + \sigma_2^2 - 2Cov(V_1, V_2)}} = Z_{[0.85]} \quad (11)$$

This is the 85 percentile value in the standardized normal distribution, Therefore;

$$85MSR = (\mu_1 - \mu_2) + Z_{[0.85]} \sqrt{\sigma_1^2 + \sigma_2^2 - 2Cov(V_1, V_2)} \quad (12)$$

Now we can investigate the analytical relationship between ΔV_{85} and 85MSR.

In this discussion, we wish to show;

$$\sqrt{\sigma_1^2 + \sigma_2^2 - 2Cov(V_1, V_2)} \geq \sigma_1 - \sigma_2 \tag{13}$$

The left side of the eq.13 is a component of the 85MSR expression [eg. (12)] and the right side of the eq.13 is the component of the conventional ΔV_{85} expression [eq. (5)]. If we can show that the eq. (13) is true, then we can conclude the 85MSR is always greater than or equal to the conventional ΔV_{85} measure.

By definition, we know that the correlation coefficient (ρ) has the following property:

$$-1 \leq \left[\rho = \frac{Cov(V_1, V_2)}{\sigma_1 \sigma_2} \right] \leq 1 \tag{14}$$

It follows therefore that the covariance term can be bounded as follows:

$$-\sigma_1 \sigma_2 \leq Cov(V_1, V_2) \leq \sigma_1 \sigma_2 \tag{15}$$

1) Case 1: If $Cov(V_1, V_2) = \sigma_1 \sigma_2$, then; $\sqrt{\sigma_1^2 + \sigma_2^2 - 2\sigma_1 \sigma_2} \geq \sigma_1 - \sigma_2$

2) Case 2: If $Cov(V_1, V_2) = -\sigma_1 \sigma_2$, then; $\sqrt{\sigma_1^2 + \sigma_2^2 + 2\sigma_1 \sigma_2} \geq \sigma_1 - \sigma_2$

Consequently, eq. (13) is always true in either case. Hence we can say that the speed differential obtained using 85MSR is always greater than or equal to the differential using the conventional ΔV_{85} measure.

As mentioned previously, the conventional ΔV_{85} measure assumes independent sampling, that the speed data samples collected at the tangent and curve elements are drawn from independent populations that follow a normal distribution. On the other hand, the expression for 85MSR accounts for dependency between V_1 and V_2 through the $Cov(V_1,V_2)$ term. Basically, 85MSR measure does not consider the two different samples but rather a single sampling population of speed differences.

Generally, a driver's personal characteristics (e.g., age, gender, trip purpose, fatigue, etc.) are assumed to comprise one of the dominant factors for choosing operating speed at any element along a given highway. It is reasonable to assume that these characteristics would not vary along the tangent-curved section. Resultantly, the independence assumption between two consecutive elements in a section in the conventional speed differential measure (ΔV_{85}) is too strong an assumption to be generalized for all drivers. In situations where individuals are members of a group (a tangent/curved section), this independence assumption can be and is often violated.

EMPIRICAL COMPARISION OF ΔV_{85} AND 85MSR MEASURES

Based on empirical study, as stated, McFadden and Elefteriadou found that the speed differential captured by 85MSR is on average about 2.0 times greater than the speed differential obtained from conventional ΔV_{85}. What is the meaning of magnitude 2.0 in their relationship?

Fundamentally, case 1 in eq. (15) assumes a perfect positive correlation between the speeds achieved in two successive highway elements. This means that the fastest vehicle in the tangent element must be the fastest vehicle in the corresponding curved element. Also, the slowest vehicle in a tangent element must be the slowest vehicle in the corresponding curved element. On the other hand, case 2 in eq. (15) assumes a perfect negative correlation between the speeds at two highway elements, such that, the fastest vehicle in the tangent element becomes the slowest vehicle in the corresponding curved element, and vise versa. Figure 1 and 2 describes the individual vehicles speed profiles in both cases. However, the assumption of a perfect correlation between speeds in two consecutive highway elements is unrealistic regardless of its direction (i.e. signs). In public highways, the speed profiles would be somewhere in between case 1 and case 2.

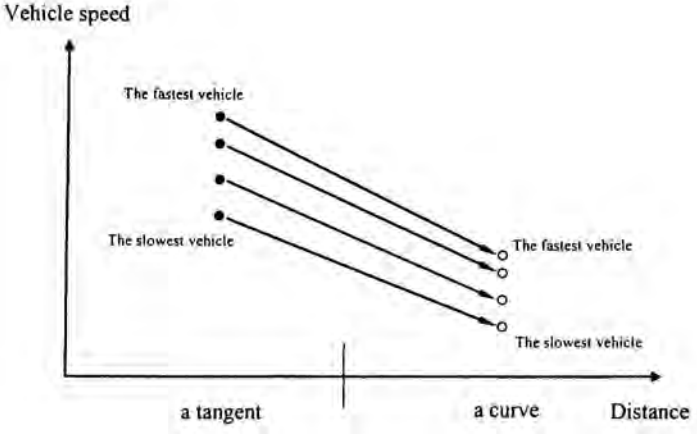


Figure 1. Speed change profiles when ρ approaches +1.0.

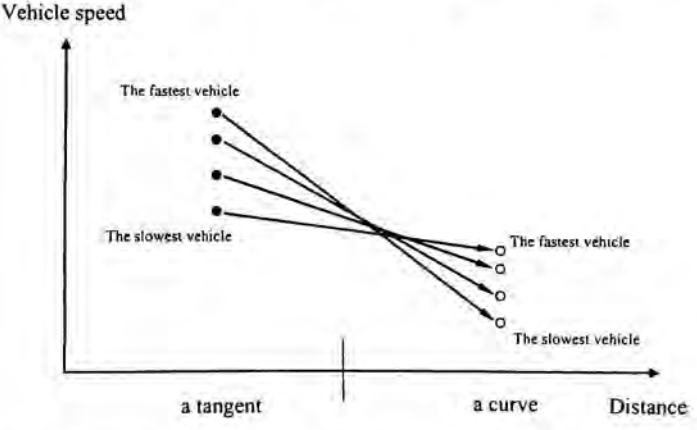


Figure 2. Speed change profiles when ρ approaches -1.0.

We can imagine that the magnitude of the relationship between the two measures represents the intensity of fluctuations in individual vehicle speeds along a given highway. The larger the values of the estimated relationship, the closer we get to perfect negative correlation (case 2). The magnitude of 2.0 provided by McFadden and Elefteriadou reflects the fluctuations of individual vehicles speeds in their study samples. Consequently, where the fluctuation of individual vehicles speed is more severe than McFadden and Elefteriadou's study dataset, the magnitude should be greater than 2.0, and vise versa.

Let us presume that Table 2 is the defensible criteria for evaluating speed consistency between two successive highway elements. Inasmuch as the conventional speed consistency measure (ΔV_{85}) significantly underestimate the amount of speed variation experienced by the individual vehicle/driver. Resultantly, the conventional measure is prone to accept inconsistent highway sections (e.g. poor safety condition) as consistent (i.e. fair safety condition).

Table 2.
Safety evaluation criteria for a tangent/curved section (Lamm 1988)

Safety condition	Speed consistency criterion
Good	$\Delta V = \left\|V_{i[0.85]} - V_{i+1[0.85]}\right\| \leq 10 km/h$
Fair	$10 < \Delta V \leq 20 km/h$
Poor	$\Delta V > 20 km/h$

DISCUSSION AND CONCLUSION

It might be impractical to track individual vehicle speed profiles along successive elements of a highway. Since the conventional $\Delta V85$ procedure uses aggregate speed profiles, it does not require the tracking of individual vehicle speeds, and hence is relatively easy to apply. However, we will need to understand the safety deficiencies of using this $\Delta V85$ measure as opposed to 85MSR.

Based on the comparison results from an analytical perspective, we can conclude that the 85MSR measure reflects two distinct enhancements over the conventional $\Delta V85$ measure. They are;

1) It consistently yields higher speed differentials between successive elements of a highway section, hence providing a more conservative indication of potential crash risk for the highway.

2) It allows for dependency in speeds attainted by individual vehicles/drivers traversing the same two element highway section, hence providing a more reasonable explanation of how drivers choose their operating speeds when faced with changing highway conditions.

These two enhancements suggest that 85MSR measure should be used to establish speed consistency and safety along a highway. In adopting 85MSR as an input into highway design and traffic management we would better able to develop strategies that lead to significant safety gains.

REFERENCES

Fitzpatrick. K., L. Elefteriadou, D. W. Harwood, J. M. Collins, J. McFadden, I. B. Anderson, R. A. Krammes, N. Irizarry, K. D. Parma, K. M. Bauer, and K. Passetti. (2000) *Speed Prediction for Two-Lane Rural Highways*. Report FHWA-RD-99-171. Federal Highway Administration.

Hirsh, M. (1987) Probabilistic approach to consistency in geometric design. *Journal of Transportation Engineering*, 113, No.3, 268-276.

Krammes, R. A., R. Q. Brackett, M. A. Shafer, J. L. Ottesen, I. B. Anderson, K. L. Fink, K. M. Collins, O. J. Pendleton, and C. J.

Messer. (1995) *Horizontal Alignment Design Consistency for Rural Two-Lane Highways*. Report RD-94-034. Federal Highway Administration.

Lamm, R., E. M. Choueiri, J. C. Hayward, and A. Paluri. (1988) Possible Design Procedure to Promote Design Consistency in Highway Geometric Design on Two-Lane Rural Road. National Research Council. Washington, D.C., *Transportation Research Record* 1195, 111-122.

Leisch, J.E. and Leisch, J.P. (1977) New concepts in design-speed application. National Research Council. Washington, D.C., *Transportation Research Record* 631, 4-14

McFadden, J. and Elefteriadou, L. (2000) Evaluating horizontal alignment design consistency of two-lane rural highways: development of new procedure. National Research Council. Washington, D.C., *Transportation Research Record* 1737, 9-17.

Misaghi, P. (2003) Modelling operating speed and speed differential for design consistency evaluation. M.Sc. Thesis, Department of Civil and Environmental Engineering, Carleton University, Ottawa, Ontario, Canada.

ESTIMATION OF VEHICLE-TO-VEHICLE ACCIDENT RISK AT SIGNALIZED INTERSECTIONS

Dewan Masud Karim, Department of Civil Engineering, Ryerson University, Toronto, Ontario, Canada

ABSTRACT: Conventional approaches of intersection safety management are not always effective enough to reduce the black spots in a congested city. Driver safety support systems integrated with navigation system could be a valuable means of eliminating several major types of intersection accident where reconstructing the infrastructures would require a large-scale investment. Based on the relationship between the distribution of disturbance near intersections and drivers reactions, this study illustrates a potential integration of microscopic accident models with the application of intelligent transport systems to predict the risk involve for major types of accident frequency and describes some explanatory variables that represent both traffic flow and intersection design characteristics. Negative Binomial distribution applied to counter overdispersion issue that usually discerned in accident data. Coefficients of the parameters were estimated using maximum likelihood method. The upshots of the model exhibits the influence of certain variables on four major types intersection accident which fruitfully explain the mechanism accident and assists to take certain meticulous countermeasures to reduce complex judgments at intersection especially for seniors as well as decrease driving mistakes at intersections in general. This methodology provides tools for both management agencies and individual intersection user.

1. INTRODUCTION

The severity of vehicle-to-vehicle intersection accidents elevates their importance to traffic safety professionals. According to a study

conducted by the Institute for Traffic Accident Research and Data Analysis (ITARDA 1998), vehicle-to-vehicle accidents at signalized intersections account for 27-32% of total accident in Japan. Approximately 58.4% of accident and 47.4% of fatality at intersection elevate the importance of intersection accident. These striking features point to a need for undertsanding the microscopic nature of acident occurrence mechanism in order to develop advanced and intelligent countermeasures to increase safety at signalized intersections.

Highly congested city, for instance Tokyo and Osaka, suffers a large number of life and property losses at intersection blackspots mostly due to impaired visibility, inadequate sight distance, complex intersection shape, geometric design and road infrastructure facilities (Hamaoka 1999 and Karim 2001). On the other hand, traffic safety was not major concern during the design and construction of transport infrastructure in 1970's. However, although poor geometric designs are one the major issues of safety performance, a complete redesign of intersections is not possible due to massive reconstruction cost. Conversely, intelligent transport systems, for example Driving Safety Support Systems, DSSS (Figure 1), Intersection Collision Warning System (ICWS), dynamic message display equipment and compact text-based information assists driver as well as educate intersection users to remain alert about the cause of accident when they arrive at intersection. The intent of the research is to describe a new concept of estimating intersection accident probabilities and potential application of ITS tools that would provide the path to minimize risks of crashes by the harmony of safety-enhancing functions within the intelligent vehicles and advanced transport infrastructure.

Hauer et al. (1988) classified the intersection vehicle-to-vehicle accidents into 15 types according to vehicle movements before collision. This classification provides microscopic details of vehicle-to-vehicle accident frequencies. This study adopted aforementioned defition of accident classification and describes safety countermeasures for four major types of intersection accident: Rear-end (RE) and three types of angle-accident - Rigth-turn (AG1), Right-

angle (AG2) and Left-turn (AG3). Head-on collisions are not consided in this paper. Detail mathematical derivation of microspic accident model of each type were described elsewhere (Wang 2003, Karim et. al. 2001, and Karim 2005).

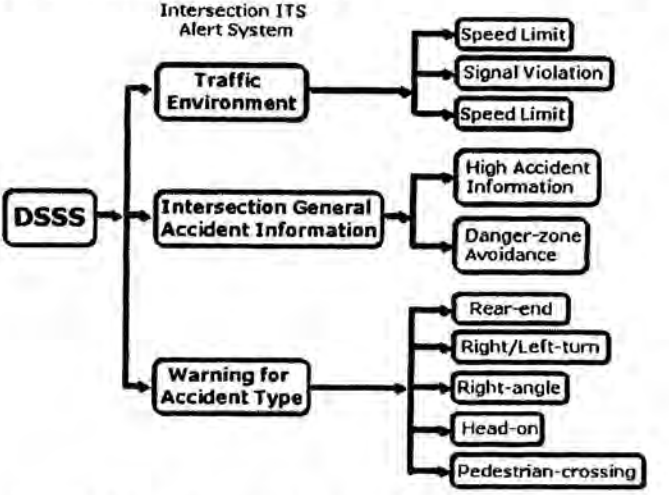


Figure 1: Components of Driving Safety Support Systems
Note: Right-hand driving in Japan

This study aims to develop a new concept of advanced accident countermeasure technique to aid the traffic safety management and regulatory agencies using histories and probability estimation of four major types of vehicle-to-vehicle accidents at four-legged signalized intersection. Several insights will be investigated, establishing a relationship between the frequency of collisions and various geometric and traffic related environment and considering the interaction of the perception and reaction behavior of the two involved drivers. The results reveal some important categorized factors causing traffic accident and recommend new cost-effective strategies in terms of driver and road user information to reduce the probability of severe intersection accidents at four-legged signalized intersections.

2. CHARACTERISTICS OF MAJOR ACCIDENT TYPE AT SIGNALIZED INTERSECTIONS

For the analysis purpose, a database of 457 four-legged signalized intersections in six prefectures was developed for the period of 1992 to 1995. Table 1 reveals some important characteristics vehicle-to-vehicle accident at intersections in Japan (IATSS 1999). Most frequent accident type is rear-end accounting 29% to total followed by right-angle accident (27%). Angle accidents, for instance, right-angle accident suffers the most devastating records of having 39% fatality and left-turn accident occupies 45% of accident with motorcycle at signalized intersections in Tokyo prefecture. For this reason, a new approach was considered in this research to overcome the shortcoming of conventional safety countermeasures taking into consideration that human factors play a very important role in accident occurrence, a serious specification error conducted by previous researcher in absence of psychological factors.

Accident Type	Percent of Total Accident	Percent of Total Fatality
Head-on Collision	3.9	13.3
Rear-end Collision	29.6	6.1
Right-Angle Collision	27.8	17.5
Right-turn Collision	11.1	5.5
Left-turn Collision	6.2	1.5
Other Vehicle-to-vehicle accident	6.6	3.5
Total	**85.2**	**47.4**

Table 1: Basic Facts of Vehicle-to-Accident at Signalized Intersections in Japan (1999)

Aging society in developing nations brings significant change of trend in intersection accident in recent years. Study results (Oda et.al. 1999) showed that senior citizens occupy 32% of fatality of total 10,000 total fatalities each year in Japan. Figure 2 illustrates high accident rate at or near the intersection mainly due to driving mistakes at complex situations and lower reaction time experienced

by the senior drivers over 65 years. Therefore, in the view of future traffic society, it is indispensable to consider senior friendly traffic signal systems. Results of this research expected to reduce complex judgments at intersection especially for seniors, decrease driving mistakes happen at intersections, particularly right-turn and rear-end collisions.

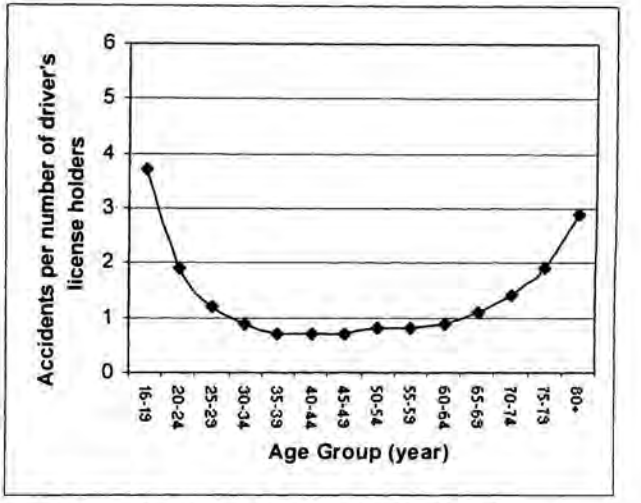


Figure 2: Intersection Accident by Age Group

3. RESEARCH BACKGROUND

Recent projects of intersection collision warning systems and improvement of potential application of advanced intelligent transport systems receiced new attention from the traffic engineers. Research indicates that Japan and several European countries are making significant advances in the areas of motorist information and system integration (FHWA 2004). National Police Agency which governs Japanese traffic management has been promoting to introduce Universal Traffic Management (UTMS 21), an advanced traffic management system for the twenty-first century. UTMS 21 consists of eleven sub-systems including DSSS (Aoyama 1999) which has the

special abilities of interactive two-way communication between infrared beacons and in-vehicle units to collect vehicle idenfication number and trasmit information to the driver.

Driving Safety Support Systems ensures the safe and free mobility for senior drivers and discourages dangerous driving practice. The key concept is to detect drivers at the intersection using infrared vehicle detector in an existing infrastructure facilities and two-way communication and global positioning facilities (Oda et. al. 1998). The central idea behind the system is to estimate the risk of specific types of accident when drivers failed to understand their possible collision risks. Similar initiative taken by Federal Highway Administration, USA (Ferlis 2005) also emphasizes the importance identification of intersection characteristics using actual crash information to estimate the risk of different types of accident and categorize the principle cause of accidents (Ferlis 2000).

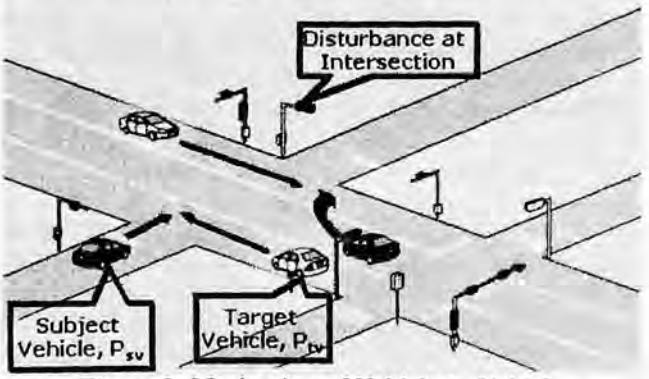


Figure 3: Mechanism of Vehicle-to-Vehicle Accidents at Intersection

This study explicitly considers two unique facts to estimate the risk of major collision risk at the signalized intersections. Firstly, experiment underlying the mechanism of visual search (Miura 1992) shows that in demanding situations, especially in crowded intersections, a driver's functional field of view becomes narrower and hence reaction time for detecting relevant objects becomes longer.

Therefore, complex visual tasks while approaching to intersection change the range of perception-reaction time (PRT) depending on the complexity of the solution and driver's expectancy of the hazard (Bates 1995). Secondly, a vehicle, which violated the traffic rules or faces an obstacle first (say 'subject or obstacle vehicle') always possess higher probability of accident involvement compared to the 'following or traget vehicle'. Thus, different set of variables are responsible for the different vehicles involved in vehicle to-vehcile accidents. Figure 3 depicts the conceptual diagram of the vehicle-to-vehicle accident occurrence mechanism at signalized intersections.

4. METHODOLOGY

The concept of the development of models is that each vehicle carries a certain portion of the 'total risk' of vehicle-to-vehicle accident. This key consideration was not the focus of previous research since, from the police accident investigation to the prosecution of the offence, each system pertaining to a crash makes an effort to identify the 'main' offender i.e. subject vehicle. With the "proportional risk" concept, multiplying the individual probabilities, which are assumed independent of each other, provides the total accident risk. To simplify the task of modeling, only two vehicle crashes are considered for the analysis.

Derivation of the probability of subject vehicle depends on the distribution of disturbance faced by the driver's traveling towards the intersection. It is reasonable to assume that occurrences of such 'disturbances' follow a Poisson process. Manner and Kilareski (1998) suggest that time intervals between acceleration or deceleration-inducing disturbances are exponentially distributed. After simplifying the exponential density function, the probability of target vehicles encountering a subject vehicle that corresponds to at least one disturbance occurs within specified time period becomes

[1] $$P_{sv} = 1 - e^{-e^{\alpha X}}$$

where α and X represent estimation coefficients and explanatory variables of disturbance frequency respectively.

From the following vehicle driver's point of view, perceiving unexpected obstacles depends on the number of vehicles approaching a red or yellow light from the main approach (for rear-end and right-angle accidents) or the quality of their response while waiting for gaps (for turning accidents). Available gaps represent required perception-reaction time that can be assumed as Weibull distribution due to its good approximation of Normal distribution (Abernethy 1996). Conversely, required perception-reaction time varies with driver characteristics, intersection geometry and traffic conditions (Bates 1995) but can again can be assumed as Weibull distribution. Integrating both perception-reaction distribution and related explanatory variables using an exponential distribution generates the probability of target vehicles avoiding an accident

$$[2] \qquad P_{tv} = \frac{1}{1 + e^{-\beta Y}}$$

where β and Y represent estimation coefficients and explanatory variables respectively.

Finally, the total accident probability can be achieved by multiplying the probability of target vehicle encountering subject vehicle (Equation 1) and probability of target vehicles avoiding the accident (Equation 2). It can be assumed that both events are independent of each other. The expected number of accident can then be estimated multiplying the accident probability by the corresponding turning movement traffic volume of the approaching street. The Negative Binomial distribution was used in this study since the Poisson, even if it were valid, is a special case of this distribution. The models were estimated by standard maximum likelihood methods.

5. MODEL RESULTS

Tables 2 and 3 show the model estimation results. After considering 94 variables that were collected during the site survey, only 37 were found relevant in terms of statistical significance tests, logic of model development and the knowledge from the literature review. Of these,

21 variables had an impact on the probability of subject vehicles being involved in an accident and 16 variables affected the same probability for target vehicles. The average probability for the cross-street vehicle is lower than for the main street vehicle as explained earlier. A positive sign corresponds to an increasing effect on vehicle-to-vehicle accident probability whereas a negative represents a decreasing effect. All parameters are statistically significant at the 10% level. Below is a discussion of the model relationships for individual variables.

The effect of traffic flow: Some accidents, for instance, AG2 shows a unique relationship with traffic flow compared to other types of accidents at signalized intersection. Baguley (1988) hypothesized that the red-light running rate is not a simple function of major flow. Higher right-turns seemed to be associated with increased for RE1, AG1 accidents and reduced AG2 accidents since as a major part of right-turn flow diverts from the through flow, thereby decreasing the chance of becoming an obstacle for the cross-street traffic flow. On the contrary, higher cross-street flow (cross-street left and right traffic flow) results in lower headway and hence higher chance of encountering a subject vehicle. Other researchers reported the identical results regarding cross-street flows.

The effect of traffic control: Signal progression clearly has a smoothing effect on traffic that reflected in reduced probability of encountering a RE1 subject vehicle (Wang et. al. 2003). Existence of multiphase signal control reduced the risk of RE1 and AG1 accident as traffic flow segregated by the different phase of the signal. A negative sign of intergreen period variables in this analysis confirmed the similar conclusion of previous research that longer change intervals can reduce AG2 accident risk (Retting and Greene 1997).

The effect of Driver and vehicle characteristics: The coefficient sign indicates that functional field of view reduces all types angle accident frequency (Miura T., 1992). The variable, large vehicle ratio

Parameters in Model	Coefficients			
	RE1[6]	AG1	AG2	AG3
Constant	4.236	-3.13	-1.924	-3.745
Combined through-right, right turn drop lanes and two or more lanes on the approach[2]	-	-	0.981	-
Intersection in commercial areas[1]	-0.487	-	-	0.32
Right-turn restriction[1]	-	-	1.101	-
Reciprocal of lanes no of leg	-7.821	-	-	-
Total lane no. of left approach	-0.09*	-	-	0.11
Angle of road in between two adjacent legs[1]	-	0.112	1.977	-
Absolute displacement between two apposite legs[1]	-	0.973	-	-
Curvature on approach leg[1]	-	0.438	0.207	-
Speed limit of entering approach	-0.06*	0.026*	-	0.01*
Pedestrian overpass leg at corner[1]	-	-	0.964	-
Intersection sheltered by elevated road[1]	-3.65*	0.179	1.81	1.18
Existence of fence in median of entering approach[1]	-0.022	-	-	-0.42
Central road median[1]	-	-	0.524	
Signal Progression[1]	-0.254	-	-	0.549
Signal control pattern[4]	-.0431	-0.368	-	-
Intergreen period on current leg[3]	-	-	-1.352	-
Right approach road width (m)	-	-	-0.048	-
Right-turn traffic volume (4 years) of the current approach[5]	0.249	0.014	-0.077	-
Through traffic volume (4 yrs) of the cross-street approach[5]	-	-	0.075	0.03
Through traffic volume (4 years) of current approach[5]	0.053	-	0.076	-
Right-turn traffic volume (4 years) of the opposite approach[5]	0.073	-	-0.059	-1.41

Table 2: Variables Affecting the Probability of Encountering Subject Vehicle

Notes: *: Parameter significant also for target vehicle; 1. If exist 1, 0 otherwise; 2. If there are two or more combined through-right or right-turn drop lanes exist 1, 0 otherwise; 3. If period longer than 3 sec then 1, 0 otherwise; 4. For 4 phase control 1, 0 otherwise; 5. Traffic volume in thousand; 6. Results cited in Wang et. al 2003.

Parameters in Model	Coefficients			
	RE1[7]	AG1	AG2	AG3
Constant	-19.46	-12.68	-15.57	-20.17
Existence of driveway[1]	-	-	0.138	0.630*
Angle between entering approach and opposing approach[2]	0.566*	0.059*	0.591	0.180
Exclusive right lane[1,*]	-	-0.274	-0.342	-
Permissive right turn[1]	-		-2.115	-
Right-turns not aligned and not single lane approach[1]	-	-	-0.164	-
Existence of more than two right-turn lanes[1]	-	0.43	-	-
Through lane no. of left approach	-	-	-	0.11
Total no. of approach lanes	-0.441	-0.268	-	0.42
Slope of entering approach	0.622	-	-	0.30
Local street approach[3]	-	0.904*	0.881	2.09
Large vehicle ratio of current leg	-	0.02	0.053	0.03
Functional field of view[4]	-	-0.812	-0.702	-1.49
Visual noise (from level 0 to 4)	0.130	-		-
Night-to-day traffic flow ratio	1.588	-	1.382	-
Left-turn volume (4 years) of current approach[6]	0.137*	-	0.009	-
Motorcycle ratio of opposite through traffic	-	1.686	-	1.07*
Overall Estimation Results (Sample Number 1146)				
Likelihood ratio Index	0.698	0.51	0.68	0.47
Avg. probability of encountering a subject vehicle (P_{sv})	0.339	0.133	0.159	0.295
Average probability of the failure of target vehicle driver (P_{tv})	$2.0*10^{-6}$	$1.28*10^{-6}$	$5.25*10^{-6}$	$3.8*10^{-6}$

Table 3: Variables Affecting Target Vehicle Failure Probability

Notes: *: Parameter significant for subject vehicle; 1. If exist 1, 0 otherwise; 2. If the angle is less than 15^0 then 1, 0 otherwise; 3. If local street with traffic less than 10,000 approach and long red time 1, 0 otherwise; 4. If angle of clear view is less than 10^0 or any kind of disturbances exist in opposite side of current leg then 1, 0 otherwise; 5. If greater than ±3% then 1, 0 otherwise; 6. Traffic volume in thousand; 7. Results cited in Wang et. al 2003.

represents greater perception and reaction time required by the large vehicle drivers to avoid an angle accident risk. It takes heavy braking within insufficient time to avoid the subject vehicle.

The effect of lane configuration: The effects of eight types of lane configuration were explored. As the right turning maneuvering takes significant time, the right turning vehicle approaching at intersection from the combined through and right lane will have slowed down; consequently a following vehicle proceeding through the intersection will visualize the right turning vehicle as an obstacle and would move to an exclusive through lane if available, increasing accident risk. On the other hand, the existence of an exclusive right lane reduces right angle accident likelihood as a result of the smooth maneuvering of right-turning vehicle. Separation of entering legs more than 30m with opposite approach increases the risk of AG1 accident. Right-turn pocket also reduces the probability of conflict at the intersection approach for through traffic vehicle by decreasing through traffic volume in individual through lanes (Datta 1991). Results show that if the major approach has protected right turn, AG2 accident risk tends to increase due to high resulting trough traffic. Similarly, if there are more right-turn lanes, the existence of the right-turn permission may be difficult to ascertain by opposite through vehicle drivers; thus the average available PRT should be shorter, thereby reducing the chance to avoid failure.

The effect of geometric variables: The angle between the adjacent right or left approach is the best predictor of all types intersection accident risk. After studying several signalized intersection types, Hall (1986) concluded that intersection accidents were likely to increase with angles greater 105 degrees between the adjacent legs. Four geometric variables capture obstruction of vision effects: the pedestrian overpass legs at the corner, the elevated structures, the approach curvature and central road median. Pedestrian legs exactly at the corner restrict driver vision requiring more time to decide whether to stop or go at the onset on amber. Sight-distance restriction could also occur due to horizontal or vertical curvature across the intersection. Due to land resource constraints in Tokyo, most of the

elevated circular expressways are built over the urban streets. If an elevated road shelters an intersection, the darkness delays the perception of through vehicles in the intersection and hence increases the likelihood of angle accident frequency. Existence of driveways near the intersection approach creates disturbances that affect safety positively for the cross-street vehicle driver.

6. SUMMARY AND CONCLUSIONS

This paper describes the establishment of a new concept of intersection safety countermeasure technique by identifying the important factors involved for major types of vehicle-to-vehicle accident at signalized intersections. Microscopic view of intersection accidents illustrates that involved vehicles possess different risk probabilities that are affected by different sets of variables. Results of the models demonstrated that poor signal design coupled with complex geometric design is a principal contributor to the right-angle accident risk at signalized intersections. Psychological and design parameters combined with drivers license IC card could reveals the nature of divers behavior at the intersection, their historical record and corresponding response from the integrated traffic control center directly interacting with drivers through voices. Depending on data availability of explanatory variables, this model may be transferable to any other location at signalized intersections. In addition, the microscopic mechanism presented could be an important theoretical foundation for exploring the effects of innovative measures such as intelligent traffic safety warning systems for improving vehicle-to-vehicle collision safety at urban intersections.

7. ACKNOWLEDGEMENTS

The Ministry of Land and Infrastructure, Japan and Tokyo Metropolitan Police Department (TMPD) sponsored this research project. The author is also grateful to Traffic Accident Research and Data Analysis (ITARDA) by providing the valuable data and information for the model development. The writing of the paper is part of on-going work on this topic that is supported by a grant from

the Canadian Government sponsored Network Centers of Excellence on the Automobile and the 21st Century (Auto21).

8. REFERENCES

Abernethy R. B. (1996) The New Weibull Handbook, 2nd ed., *Society of Automotive Engineering International*, North Palm Beach, FL, USA.

Aoyama K. (1994) Universal Traffic Management System (UTMS) in Japan, Proceedings of Vehicle Navigation and Information Systems, IEEE, pp. 619-622.

Baguley C. J. (1988) 'Running the Red' at Signals on High-Speed Roads, *Traffic Engineering and Control*, July/August, pp.415-420.

Bates J. T. (1995) Perception Reaction Time, *ITE Journal*, Vol. 65-2, pp. 35-36.

Datta T.K. (1991) Head-on, Left-Turn Accidents at Intersections with Newly Installed Traffic Signals, *Transportation Research Record*, National Research Council, Washington D.C., 1318, pp. 58-63.

Federal Highway Administration, FHWA, (2004) Intersection Collision Warning System, Publication No. FHWA-RD-99-103.

Ferlis B. (2000) Deployment Strategies – FHWA Perspective, Workshop on Intersection Collision Avoidance Research, Oct. 25.

Ferlis R. A. (2005) Infrastructure Systems for Intersection Collision Avoidance, Unpublished Paper.

Hall R.D. (1986) Accidents at Four-arm Single carriageway Urban Traffic Signals, *Transport and Road Research Laboratory*, Department of Transport, Crowthorne, CR65.

Hamaoka, H., Nagashima, H., and Morichi, S. (1999) An Analysis of the Cause of Traffic Accidents at the Black Spots, *Proceedings of the 8th World Conference on Transport Research, WCTR*, Vol. 2, pp. 111-122.

Hauer E., Jerry N. and Lovell J. (1988) Estimation of Safety at Signalized Intersections, *Transportation Research Record*, 1185, pp. 48-60.

ITARDA (Institute for Traffic Accident Research and Data Analysis, 1998), Statistics on Black Spot Accident, Tokyo, Japan.

IATSS Statistics (1999) Road Accidents in Japan, *International Association of Traffic and Safety Sciences*, Traffic Bureau, National Police Agency, Japan.

Karim D. M., Ieda H. and Terabe S. (2001) Angle Accident Model to Evaluate the Risk at Four-Legged Signalized Intersections, *Journal of Eastern Asia Society for Transportation Studies, EASTS*, Vol. 4-5, pp. 343-358.

Karim D. M. (2005) A Model to Estimate Right angle accidents at Signalized Intersections, paper submitted to *Canadian Society Civil Engineering (CSCE) Annual Conference*, Toronto, Ontario.

Mannering F. and Lilareski W. (1998) Principles of Highway Engineering and Traffic Analysis, 2nd ed., Wiley, New York, USA.

Miura T. (1992) Visual Search in Intersections-An Underlying Mechanism, *IATSS Research*, Vol. 16-1, pp. 42-49.

Oda T. and Takeuchi K. (1998) Driving Safety Support Systems in UTMS 21, 5th Annual World Congress on ITS, Seoul.

Oda T. and Tajima T. (1999) Driving Safety Support Systems for Senior Society, 6th Annual World Congress on ITS, Toronto.

Retting R.A., Williams A. F. and Greene M. A. (1998) Red-light Running and Sensible Countermeasures, Summary of Research Findings. *Transportation Research Record*, 1640, pp. 23-26.

Retting R.A. and Greene M. A. (1997) Influence of Traffic Signal Timing on Red-light Running and Potential Vehicle Conflicts at Urban Intersections, *Transportation Research Record*, 1595, pp.1-7.

Traffic Volume Statistics (1992, 1993, 1994), *Tokyo Metropolitan police Department*, Tokyo (in Japanese).

Wang, Y., Ieda H., Mannering F., (2003) Estimating Rear-End Accident Probabilities at Signalized Intersections: Occurrence Mechanism Approach, Journal of Transportation Engineering, ASCE, Vol. 129-4, pp. 377-384.

On a necessity of sensor redundancy for safety integrity improvement in GNSS/INS based train location

Introduction

For a longer time the satellite navigation has been playing an important role in many applications where position determination requirement exists. Many tasks regarding the improvement of accuracy and availability of position determination based on satellite navigation have been solved by means of e.g. additional inertial system (INS) and/or usage of GNSS receivers (i.e. GPS and Glonass - satellite navigation dual system receivers). Today, great effort focused on employment of these advantageous technologies in safety-related applications and the corresponding fast progress in this field is evident. The European Geostationary Navigation Overlay Service (EGNOS) or the European Navigation System Galileo (should be in operation from 2008) seems to be sufficient examples.

In the railways environment these technologies have already been successfully used for non-safety railway applications (e.g. passenger information system, managerial dispatch system for locomotives etc.) However, in spite of the fact that these technologies meet only non-safety railway application requirements, they cannot be applied in safety of life applications, such as signalling or track workers protection. The reason is that the train position locator (TPL) based on GNSS/INS and intended for safety applications should meet, excepting the high reliability and accuracy requirements, also high integrity design. It can be achieved among others with the active sensor redundancy, which should help to eliminate common cause sensor failures, detect and exclude both hardware and systematic sensor faults and improve position determination trustworthiness.

This paper is mainly focused on design and general analysis of TPL inertial odometry sensor (IOS) subsystem from the point of view of

hardware integrity and systematic safety integrity according to the Cenelec standard EN 61508 [1] (and also EN 50126, EN 50129). Both of these integrities must be properly considered during the design of any safety system in order to achieve the required safety integrity level (SIL).

The reason for this investigation is that the IOS subsystem will play a key role in cases when signal in space (SIS) is not temporarily available. It is most often on regional single lines (or parallel tracks without switches), especially in the sections, which are situated outside the railway stations. In these cases the train movement trajectory is also known with high accuracy and thus the odometry becomes very important one-dimensional relative train position determination technique.

1. TPL architecture description

The general architecture design of the whole TPL is depicted in Fig.1. The TPL consists of three parts: the sensor subsystem (which comprises INS, GNSS receiver and track-map database), the logic solver and the effector subsystem.

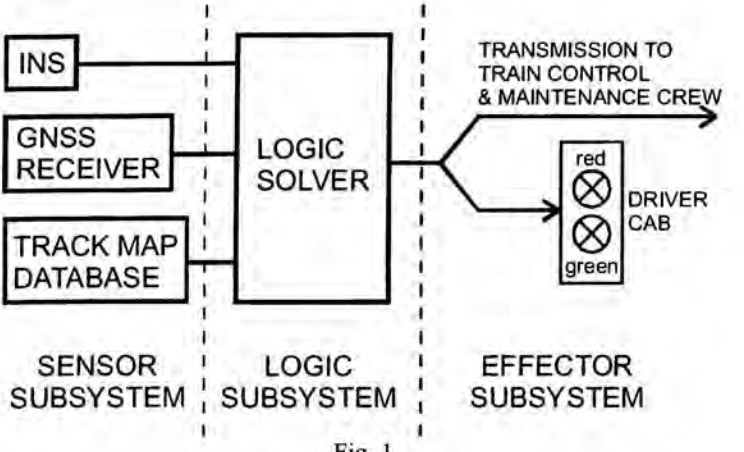


Fig. 1

1.1. Sensor subsystem

As it is shown in the Fig. 1 the sensor subsystem consists of INS, GNSS receiver(s) and track map database. This arrangement enables to set up two independent loops for odometry problem solution, where one loop involves GNSS receiver(s), track map database and gyroscope, and the second one consists of accelerometer, odometer and Doppler radar speedometer.

The original INS was only composed of the CXL01LF3 Crossbow 3-axis accelerometer (range of +/-1g), the LTV 14 4-channel opto-electronic axle odometer (100 pulses per revolution) and the KVH-RA2100 fibre optic gyroscope (rate of +/-100 degrees per second). The odometric sensor subsystem (consisted of the accelerometer and the odometer) as a part of the INS seems to be enough accurate and sufficient to detect short time wheel slips. The disadvantage of this system consists in the drift and thermal bias of the accelerometer. From this point of view such system is not sufficient for long time operation without GNSS SIS or long time slight wheel slips detection. An example is shown in chapter 4. Some experiments concerning the validation of the odometric system by means of the gyroscope, track maps and double heading differences method [2] have been already performed in our lab. However, this mentioned method has limited usage during the lower speeds (below 2 km/h) due to the gyro drift effect.

Thus, the primary task to make the IOS subsystem without this disadvantage has remained. The improvement of the IOS subsystem performance can be reached by the implementation of higher sensor redundancy as shown in the following chapters. In order to extend the IOS subsystem favourably, the requirement for another independent sensor with the acceleration drift reduction has to be fulfilled. Several sensors have been considered – an optical system with the spatial filter [3], an eddy current sensor [4] and the microwave Doppler speedometer. Finally, the Doppler speedometer has been chosen for the following reasons. Only the two-beam 24GHz Doppler radar DRS05 (Deuta - Werke GmbH., Germany) fulfilled the requirements regarding its robustness to heavy railway environment according to the EN 50155 standard, built-in diagnostics, good resolution and small error in wide velocity range in comparison with the eddy

current sensor. The DRS is not also affected by negative effects caused by dust particles, spraying water, snowflakes and sunshine as the optical system with the spatial filter is. The detailed architecture design of odometric sensor subsystem is described in chapter 2.

1.2. Logic solver

This subsystem is considered a black box with the 2oo3 architecture or the 2oo4 architecture into the future. These architectures are mostly recommended by safety logic solver manufacturers for their high integrity and availability. The logic solver performs especially the following tasks regarding the sensor subsystem.

a) Diagnostics

The logic solver primarily ensures the hardware safety integrity. It runs and evaluates diagnostic test(s) for single sensors (of course, if the sensors are accommodated to this task), or only monitors sensors diagnostic outputs (in case of the sensor diagnostics and self-test existence). If the hardware failure is detected, the logic solver performs the proper action to the effector(s) in dependence on the sensor subsystem architecture.

b) Failure detector

The logic solver also provides the fusion of the sensor data. The fusion is provided by means of the Probability Data Association Filtering (PDAF) method [5], i.e. robust estimation, which performs both hardware and systematic fault detection and exclusion of faulty measurements by down-weighting measurements that are distrusted.

c) Monitor integrity

Another important task is the monitor integrity (MI), which assesses whether the failure detector provides sufficient integrity to the estimated parameters. It means the MI will assess the detection power of the fault detection algorithm, as well as the influence that undetected faults might have on the position determination [6].

1.3. Effector section

The last subsystem is divided into two parts for the following reason. The message, which represents output from the logic solver, includes position information, estimated error of the position determination (accuracy), safety integrity flag (or level) and indication of potential sensor failure. This information could be used by train control (for evaluation of train movement and subsequently for back-to-train transmission of its next movement authorization), maintenance crew (for repair execution) and train driver (e.g. for train driving under bad conditions of visibility). Than the effector can be represented 1) by the train control or maintenance, and 2) by "can rely on" or "don't use" TPL position information for driver.

2. Odometric sensor subsystem architecture design

The usage of the TPL in safety related applications requires safety risk assessment and safety integrity level (SIL) determination. The architecture design of the IOS subsystem and its simplified analysis represent the first steps in solution of these tasks. In accordance with the EN 61508 the IOS subsystem architecture design should take into account two kinds of the sensor failures: 1) the hardware (physical) sensor failures, and 2) the systematic (non-physical) sensor failures. The design should also carefully distinguish the influence of chosen architecture on hardware and systematic fault detectability and also on both hardware and systematic safety integrity. However, as it has been already mentioned above, the diverse sensor redundancy is very useful and necessary. It can improve both hardware and systematic safety integrity. Thus, a brief summarization can be made. The architecture design and sensor redundancy have significant influence on reliability, availability, hardware and systematic safety integrity as it is shown in the next chapters.

3. IOS subsystem - hardware safety integrity

This chapter outlines some considerations and analysis of simple examples regarding the TPL IOS hardware safety integrity according to the EN 61508 standard. An influence of such parameters as the sensor failure rates or the factors representing rate of the detected/undetected common cause failure is discussed.

To assign the SIL to the system, the values of PFD_{avg} (Average Probability of Failure on Demand) or PFH_g (Probability of failure per hour for channel group with majority voter logic) should be determined. The probability of failure per hour for channel group with majority voter logic (PFH_g) is used for railway signalling systems (systems with high demand). If the sensor subsystem includes only one group of voter channels (in our considered case) then the PFH_g is equal to the PFH_s (PFH for sensor's subsystem). For next consideration the failure rate is assumed identical for all sensors.

3.1. Influence of architecture design on hardware safety integrity

Firstly, the original IOS subsystem consisting of accelerometer and odometer is considered. It can be arranged and analyzed according to the following configurations: the one-out-of-two (1oo2) architecture, the one-out-of-two with diagnostics (1oo2D) architecture and the two-out-of-two (2oo2) architecture.

The expressions for PFH_s computation of these IOS subsystem architectures are following:

$$PFH_{s1oo2} = 2[(1-\beta_D)\lambda_{DD} + (1-\beta)\lambda_{DU}]^2 t_{CE} + \\ + \beta_D\lambda_{DD} + \beta\lambda_{DU} \tag{1}$$

$$PFH_{s1oo2D} = 2(1-\beta)\lambda_{DU}[(1-\beta)\lambda_{DU} + \\ + (1-\beta_D)\lambda_{DD} + \lambda_{SD}]t_{CE}' + \beta_D\lambda_{DD} + \beta\lambda_{DU} \tag{2}$$

$$PFH_{s2oo2} = 2\lambda_{DU}, \tag{3}$$

where β is rate of undetected common cause failures (-), β_D is rate of detected common cause failures (-), λ_{DD} is rate of dangerous detected failures (per hour) in subsystem channel, λ_{DU} is rate of dangerous undetected failures (per hour) in subsystem channel, λ_{SD} is rate of safe

detected failures (per hour) in subsystem channel, t_{CE} is channel equivalent mean down time (in hours) for architectures of 1oo1, 1oo2, 2oo2, 2oo3 and t_{CE}' is channel equivalent mean down time (in hours) for 1oo2D architecture. The values of PFH_s are calculated on the base of these expressions as it is presented in the standard. It is possible to prove from the expressions (1)-(3) and is also resulting from values of PFH_s, the 1oo2D or 1oo2 architectures represent the best choice of design.

This order results from the fact that the 1oo2 architecture provides greater period between potential failures to respond to a hazard, but a higher probability of spurious responses. The 2oo2 architecture provides a great period between spurious responses, but higher chance of not responding when required. The 1oo2D architecture is able to address covert failures and help to redress the balance between failure to respond and spurious response [7].

The explanation of these facts for IOS subsystem is following. If the 1oo2 architecture is considered, the logic solver evaluates the hardware failure of either sensor as a failure of safe position determination. It is said in this case: "it's insufficient to obtain measured data only from one of sensors (e.g. odometer) for subsequent processing (systematic fault detection) and safe position determination " and "don't use" message is generated for the effector section. This approach secures higher hardware safety integrity. When the 2oo2 architecture is considered, the hardware safety integrity of this IOS subsystem architecture is lower, but resulting availability should be higher. This fact results from following reason. Under certain specific conditions, it can be sufficient to obtain measured data from only one sensor for the safe position determination.

The following architectures can be considered for the current IOS subsystem consisting of three sensors: the one-out-of-three (1oo3) architecture, the two-out-of-three (2oo3) architecture, the three-out-of-three (3oo3) architecture or similar architectures with diagnostics. It's possible to say in general the similar principles and rules will be valid for the three sensor architectures as for the two sensor architectures.

3.2. Influence of redundancy on hardware safety integrity

This chapter describes considerations and comparison of the two and the three-sensor odometric subsystems on the base of EN 61508 standard. The parameters as diagnostic coverage, common cause failure resistance and sensor redundancy will affect PFH_s values. The differences in magnitudes of these values are investigated by means of Tab. 1 and Tab. 2, which are overtaken from EN 61508 standard. The 1oo2 and the 2oo3 architectures have been chosen as examples.

As it is evident from Tab. 1 and Tab. 2 for low values of diagnostic coverage the 1oo2 architecture achieves slightly better values of PFH_s than the 2oo3 architecture. If the diagnostic coverage increases then the PFH_s values become the same. Excepting this fact, other important facts described in the EN 61508-6 standard have been considered. The first one, the E/E/PES (electrical/electronic/ programmable electronics systems) have capabilities to include protection against common cause failures. For example, the DRS provides limited diagnostics by means of a status byte and so it is possible to expect that β, β_D factors can be low. The second fact concerns the diversity of sensors. The higher sensor diversity is, the further lower β, β_D factors are. The third fact regards sensors redundancy when sensors diverse each other. According to the standard it can be expected that the probability of common cause failure for the three-sensor system (i.e. three sensors will be affected at the same time) is lower than the probability of common cause failure for the two-sensor system (i.e. two sensors will be affected at the same time). Thus, it is again possible to expect the reduction of β, β_D factors. On the other hand, an influence on failure rate in dependence on E/E/PES complexity must be taken into account. The higher complexity of E/E/PES is the more failure rate of the system increases. The following conclusion can be made on the above mentioned facts. The IOS subsystem with 2oo3 architecture reaches approximately the same PFH_s values as the 1oo2 architecture. Only small differences in PFH_s values can probably occur in dependence on the exactly known values of the failure rates λ and the common cause failure rates β, β_D for each sensor.

Tab.1 Probable failure rate per hour according to the EN 61508-6 for λ=1.0E-06

	DC	λ=1.0E-06		
		β=2% β_D=1%	β=10% β_D=5%	β=20% β_D=10%
1oo2	0%	1.1E-08	5.1E-08	1.0E-07
	60%	7.4E-09	3.5E-08	7.0E-08
	90%	5.6E-09	2.8E-08	5.5E-08
	99%	5.1E-09	2.5E-08	5.1E-08
2oo3	0%	1.3E-08	5.3E-08	1.0E-07
	60%	8.3E-09	3.6E-08	7.1E-08
	90%	5.8E-09	2.8E-08	5.5E-08
	99%	5.1E-09	2.5E-08	5.1E-08

Tab.2 Probable failure rate per hour according to the EN 61508-6 for λ=1.0E-05

	DC	λ=1.0E-05		
		β=2% β_D=1%	β=10% β_D=5%	β=20% β_D=10%
1oo2	0%	2.1E-07	5.9E-07	1.1E-06
	60%	1.1E-07	3.9E-07	7.3E-07
	90%	6.6E-08	2.9E-07	5.6E-07
	99%	5.2E-08	2.5E-07	5.1E-07
2oo3	0%	4.2E-08	7.7E-07	1.2E-06
	60%	2.0E-08	4.6E-07	8.0E-07
	90%	8.8E-08	3.1E-07	5.8E-07
	99%	5.5E-08	2.6E-07	5.1E-07

In spite of these facts, there is possible to make another remark regarding the architecture. The supplementation of two sensors of the IOS subsystem by the third diverse sensor can bring slight decrease in

hardware safety integrity in dependence on parameters described above, but on the other hand it significantly improves systematic safety integrity. Thus, the IOS subsystem with the diverse sensors and the 2oo3 architecture brings advantages in comparison with the IOS subsystem of the 1oo2 architecture, particularly, in the systematic fault detectability domain.

4. IOS subsystem - systematic safety integrity

A proper method must be chosen to detect and exclude systematic sensor faults. As it has been written in chapter 1.2., the PDAF method is used for the data processing instead of pure Kalman filtering due to its higher stability and efficiency. As it has been also presented in [5] the basic principle of this method consists in a fact that the method compares the measured sensor data with the estimated state vector by means of a Kalman filter with an appropriate dynamic model of the vehicle movement. In other words the method evaluates the innovations. If the innovations fall to the *a priori* defined validation gate (i.e. confidence interval), the measured data are employed for the estimated state vector update. In this case the associated probability of each event is computed, innovations are combined with the probability weights and the state estimate is updated by means of the „combined" innovation. If all the measured data are inconsistent or somehow corrupted at the given time step, then the measured data don't pass through the validation gate and the state vector is updated by means of the state vector and the corresponding covariances computed in the previous cycle. The sensor diagnostics in terms of systematic faults can be performed by means of this method.

A lot of measured data has been processed by means of this method. The data processing has been made separately for the 1oo2 and the 2oo3 architectures of IOS subsystems. Train speed versus time for the short time period wheel slip is depicted in the Fig. 2 and Fig. 3. Sensor data have been recorded from the measurements under the real time conditions with the 3kV DC electrical locomotive. Fig. 2 represents sensor behaviour of the IOS subsystem composed of the accelerometer and the axle odometer. Resulting speed curve has been obtained by means of the PDAF method.

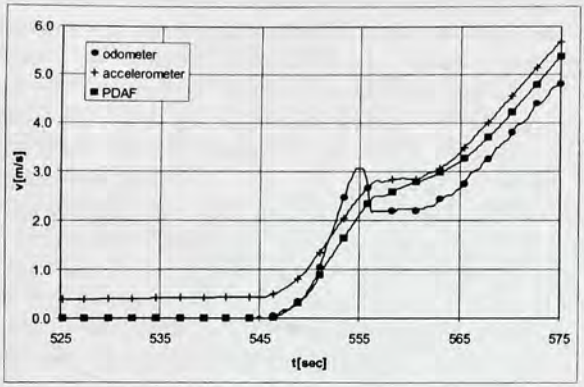


Fig. 2 - Wheel slip data from the IOS composed of the two sensors

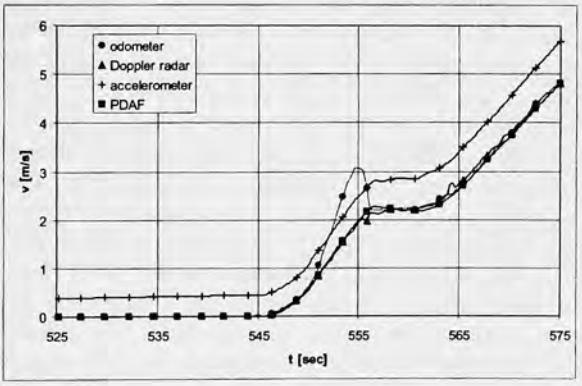


Fig. 3 - Wheel slip data from the IOS composed of the three sensors

The effect of the acceleration drift (the offset change) is evident in the picture. For this reason the PDAF method assigns high probability weight to the odometer and resulting fused velocity data are mainly

based on an application of the dynamic model of vehicle movement and the real odometer data. When strong wheel slip appears during the short time interval, the difference between the speed of the dynamic model and the odometer rapidly increases. Then the PDAF method uses only the dynamic model for the fault detection and the data fusion. After the wheel slip disappears, the method is not able to provide correct output data. The Fig. 3 outlines the sensor speed curves of the IOS extended with the DRS and resulting curve from the PDAF data processing method. The same speed curves are used for the accelerometer and the odometer in this picture as there are in Fig. 2. It is evident from comparison of both pictures that the three-sensor-odometric subsystem is not affected by the wheel slip and the PDAF method is able to provide correct output data when the wheel slip disappears. This system is able to detect not only fast short time period wheel slips, but also smooth long time wheel slips. So a better systematic fault detectability can be expected with IOS subsystem supplemented by the DRS.

Another important fact, which should be considered, is the availability of the IOS subsystem. The availability of the whole IOS subsystem also depends on the systematic safety integrity similarly as availability of the whole IOS subsystem depends on the hardware safety integrity. As it has been mentioned in chapter 1.2., in order to assess whether the fault detector provides sufficient integrity to the estimated parameters, the technique called monitor integrity is employed. This method is based on the computation of the probability of missed detection, probability of the false alarm and the time-to-alarm factor [6]. The Fig. 2 represents IOS subsystem in so called degraded mode, which appears during the whole time interval. It means that values of the safety integrity cannot be guaranteed and "don't use" message should be generated. In Fig. 3 this mode occurs only during the wheel slip duration, when both the odometer and the accelerometer are affected by the systematic failures.

The first computations indicates that the availability of the 2oo3 architecture of the IOS subsystem has been improved approximately with factor of 10^2 (from tens to thousands seconds) in comparison with the 1oo2 architecture of the IOS subsystem for the same level of the safety integrity.

5. Towards to own architecture design

Upon the facts and results described in previous chapters the IOS subsystem consisting of three diverse sensors is taken into account for the following consideration. In order to secure high safety integrity, here is an idea to use either safety related and non-fault tolerant 1oo3 architecture or safety related and fault tolerant 2oo3 architecture.

The 1oo3 architecture gives best results from viewpoint of hardware safety integrity, but a limitation of this architecture consists in lower availability. This fact results from a reason that this architecture doesn't allow possibility to determine safe position by using of only two remaining non-fault sensors during following processing (the fault detection and exclusion, the data fusion, MI). However, this is possible under certain specific conditions. In order to improve availability, there is an idea to duplicate each diverse sensor and so implement for whole IOS subsystem triplicated 1oo3 architecture with 2oo2 (or better 1oo2D, if the sensors will be suitable equipped with a diagnostics) sub-architectures. But this solution requires additional costs for the IOS subsystem realization.

The 2oo3 architecture provides good compromise between hardware safety integrity and availability, but in a case of the one-sensor hardware failure it can (but needn't) bring lower systematic failure detectability, i.e. lower systematic safety integrity (or lower availability, when the same level of safety integrity is hold).

The precise analysis based on computation of the hardware and the systematic safety integrity, the reliability and the availability for both architectures above mentioned remains the object of future research and development of the TPL.

6. Test platforms

In order to test the odometric system performance and the data processing methods over the whole range of the measured values (reached by train) the different kinds of tests have been employed. The reference trajectory has been mostly obtained by means of the GNSS or GPS L1/L2 receivers, which have been working in the RTK (Real Time Kynematic) mode (i.e. carrier phase differential measurement providing centimetre accuracy).

The first kind of the tests has been performed with the 3kV DC electrical locomotive of 130 series in daily operating train mode or special alone locomotive runs. The locomotive was fully equipped with GNSS/INS TPL as it was depicted in Fig.1. There was also utilized radio digital network dedicated only for the tests. This radio network covers track length of approximately 120 km and enables to transmit RTCM-104 correction message into the TPL, remotely configure TPL and transmit TPL data back to the laboratory. The trial speed maximum is about 100 km/h. A relatively high value of the reference position determination error, which could appear for very low speeds ranging (from zero up to few tents of centimetres per second) due to the limited accuracy of the GPS L1/L2 RTK measurement, could be a drawback of these test methods. On the other hand, the number of measured values at small speed range is very low. So, the method is suitable for long time testing of sensor performance over remaining wide speed range.

Another test method, also performed under real time conditions, has been carried out with the mobile robots. This measuring method is focused on the tests of the odometric system at very low speeds. The same error as in the first kind of tests can appear in the position determination due to low speeds. But this error can be sufficiently excluded by connection of odometer to the non-powered axle of wheel set. Thus no wheel slip can occur and measured data from the odometer can be considered as the precise reference. The mobile robots can be controlled through the onboard computer with a predefined speed profile data or through the joystick or remotely through the Wi-Fi radio network according to the 802.11b specification.

The last test platform is represented by laboratory simulation. This platform is dedicated for the evaluation of data processing method performance. Sensor's systematic errors are intentionally introduced into the measured data (this approach even allows to simulate faulty GNSS system performance). The TPL simulator consists of precise controlled servos and all of the sensors described in the first chapter. The acceleration is simulated by means of gravity acceleration fraction measurement. The moving belt principle is used to test DRS

performance. The GNSS and sensor data recorded during real tests on the track are used for simulation.

Conclusion

The paper presents the first steps towards fulfilment of the safety requirements in the GNSS/INS based train position determination. The simplified analysis of the safety integrity of IOS subsystem with respect to the sensor redundancy and the architecture of IOS subsystem was carried out. The findings resulting from this analysis will be utilized for overall design of the TPL.

Acknowledgement

This research is supported by the Czech Science Foundation under contract No. 102/03/0871.

References

[1] Cenelec standard EN 61508 (1-7), "Functional safety of electrical/electronic/programmable electronic safety-related system", 2002

[2] A. Filip, L. Bazant, H. Mocek, J. Taufer & V. Maixner, "Dynamic properties of GNSS/INS based train position locator for signalling applications", *Computer in Railways VIII*, pp. 1021-1030, 2002

[3] A. Prokeš, "Correctness of Velocity Evaluation of System Using Spatial Filter", *Radioengineering*, vol. 11, no. 3, pp.24-26, September 2002

[4] A. Geistler, "Train location with eddy current sensors", *Computers in Railways VIII*, pp. 1053-1062, 2002

[5] J. Taufer, A. Filip, H. Mocek, L. Bazant, V. Maixner, "Validation and Fusion of the Sensor's Data Used for GNSS/INS Train Position Locator", *Proceeding of Third Scientific Conference of Jan Perner Transport Faculty*, pp. 317-322, 2003

[6] P. B. Ober, "Integrity Prediction and Monitoring of Navigation Systems", Integricom Publishers, 2003

[7] ICS Triplex, "The 8000 Series TMR System", Safety manual, Doc. No. T8094, Issue 14 September 2003

THE RELATIONSHIP OF COLLISION RATES AND DYNAMIC PERFORMANCE OF HEAVY TRUCKS

Alfonso Corredor, Karl Groskopf and Ron Madill
Freight Policy Office, Ontario Ministry of Transportation

1. Introduction

Ontario introduced a substantial increase in the maximum allowable gross vehicle weight, when it began regulating truck axle group and gross weights by bridge formula in 1970. A wide variety of vehicles with widely spaced axles and liftable axles emerged to take advantage of the flexibility offered by the new rules.

The Ministry of Transportation (MTO) has identified that loaded heavy trucks, having raised or improperly adjusted liftable axles, cause significant road wear and increase the risk of bridge failure. Road wear is worst at intersections and ramps where vehicles have to raise axles for cornering. As a result, MTO commenced work to reform the Vehicle Weight and Dimension (VW&D) regulations in 1996.

In addition to addressing excessive infrastructure damage, the project is designed to:

- improve highway safety;
- maintain industry productivity;
- improve harmonization with neighbouring jurisdictions; and,
- simplify the current complexity of VW&D rules.

Carefully prescribed vehicles with superior dynamic performance characteristics, identified as Safe, Productive and Infrastructure-Friendly (SPIF), will gradually replace existing vehicles.

The reform project consists of four phases with each phase addressing a different group of vehicles. Implementation of Phase 1 addressing non-dump light semi-trailers (up to 3-axles) started in January 2001. Phase 2 addressing all dump semi-trailers started implementation in July 2002. Phase 3 is expected to be implemented in 2005 and affects all remaining semi-trailers as well as double trailers. Phase 4 affecting straight trucks and their trailers will commence upon completion of Phase 3.

Evaluation methodologies are readily available to estimate economic impact and industry productivity changes induced by the project.

Qualitative analysis indicates achievements on harmonization and regulatory complexity. However, previous attempts to relate safety performance to vehicle dynamic characteristics have had limited success.

The purpose of this paper is to present the results of the procedure developed within MTO to relate truck collision rates and performance measures. The relationship was used to estimate reductions in collisions and related social costs expected from the project.

2. Truck Dynamic Performance and Safety

While it is implicitly assumed that there is a strong correlation between heavy vehicle performance measures and collision rates, there are few successful attempts actually proving the association between the two.

The Canadian Council of Motor Transport Administrators (CCMTA) and the Roads and Transportation Association of Canada (RTAC) proposed a set of vehicle performance measures standards in the mid-1980's [1]. Threshold values for each performance measure were introduced. The presumption was that vehicle stability, controllability and space occupied on the road would be superior when vehicles match proposed threshold values. The proposed performance measures can be assessed through computer simulations, test demonstrations, and full-scale testing.

Performance measures threshold values are now accepted standards for vehicle operational worthiness in Canada.

Not long after the CCMTA/RTAC proposal, the evaluation of the Turner Proposal done by the Transport Research Board[2] included an estimation of fatality rates for prototype and baseline trucks. The proposal was based on research done by the University of Michigan Transportation Research Institute (UMTRI)[3] showing that there is a strong relationship between fatal collisions and low and high speed offtracking, as well as with static rollover and other performance measures.

---

[1] CCMTA/RTAC, Recommended Regulatory Principles for Interprovincial Heavy Vehicle Weights and Dimensions, Vehicle Weights and Dimensions Study, June 1987.

[2] TRB, New Trucks for Greater Productivity and Less Road Wear. Special Report 227, 1990.

[3] Francher P.S., et al. Turner Truck Handling and Stability Properties Affecting Safety: Final Report, Volume 1. UMTRI, 1989.

UMTRI and other previous and later related research[4] have had difficulties withstanding peer's review[5]. Most studies relied on limited data which usually hamper statistical conclusions. However, in practice, the relationship has become accepted and used for policy and design purposes.

The most recent conclusive evidence comes from the introduction, in 2003, of a minimum static rollover threshold requirement into regulation in New Zealand[6]. Supportive research showed that vehicles with low static rollover threshold, high load transfer ratio and low yaw damping ratio have a higher likelihood of being involved in a rollover crash.

This paper provides additional evidence on the relationship between truck dynamic performance measures and safety.

3. Heavy Truck Operations in Ontario

Ontario allows some of the heaviest axle and gross vehicle weight limits in North America.

The estimated number of commercial vehicles in Ontario in 2002 includes:

- 115,000 straight trucks;
- 150,000 tractors; and,
- 262,500 trailers.

It is estimated that 75 percent of all tractor-trailer combinations operating on Ontario highways currently meet SPIF standards. The remaining 25 percent cause excessive damage to highway infrastructure and /or do not meet national vehicles performance guidelines.

Commercial vehicles increased by 37% since 1990 while the total number of vehicles increased by only 22%. During the same period, fatalities in large truck collisions decreased by 13%; the fatalities per 100,000 trucks decreased by 36%. In fact, Ontario's roads are rated as one of the safest in North America.

---

[4] See for example: Clarke K.L., et al., Analysis of Accident Rates of Heavy-Duty Vehicles, UMTRI 1988; Jovanis P.P., et al., A Comparison of Accident Rates for Two Truck Configurations, TRB, Transport Research Record 1249, 1989; Clarke R.M. Safety Considerations Associated with Truck Size and Weight Policy Deliberations, SAE Technical Paper Series, 1998

[5] See for example: Hauer E. Large Truck Safety – A Chronological Reading of the Literature. Unpublished document, Sep 2000.

[6] De Pont, J., et.al., Including Performance Measures in Dimensions and Mass Regulations. Paper presented to the 7th International Symposium on Heavy Weights & Dimensions, June 2002.

Heavy truck operations in Ontario present an ideal scenario to test the relationship between dynamic performance and vehicle safety. Concurrence of heavy vehicles, low collision rates and the VW&D reform project provided the right conditions to search for additional evidence of the relationship.

Furthermore, availability of information makes the task even more viable. Ontario has been collecting information about operational characteristics of commercial vehicles since the 1980's. The latest commercial vehicle survey (CVS) performed in 1999-2000 provided a reliable basis for estimating exposure rates for desegregated truck classes.

As in most North American jurisdictions, Ontario requires production of collision reports on crashes involving fatalities, personal injuries or property damage (in Ontario if damage exceeds $1,000). Linkage to vehicle license registration database provided a way to fill the gaps encountered in police collision reports, as explained below.

Over the years, Ontario has also developed dynamic performance measurements for a number of vehicles in addition to the original configurations analyzed by the CCMTA/RTAC study.

4. Analysis Approach

The analysis relies on demonstrating that existing classes of tractor-trailers with good performance characteristics (groups of vehicles that are likely to meet SPIF standards) have lower collision rates than existing vehicles classes with poor performance characteristics (groups of vehicles that are unlikely to meet SPIF standards).

Information about the characteristics of tractor-trailers involved in collisions is limited. Police collision reports identify involved vehicles but are geared toward determining driver actions, mechanical failures and local conditions that led to the collision. However, determination of tractor-trailer sub-classes based on number of axles was possible by linking to MTO's vehicle licensing registration database.

With tractor-trailer classes and collisions identified, the following steps comprise the analysis:

- Determine collision rates by class;
- Determine typical performance measures by class;
- Determine correlation between performance measures and collisions rates; and,
- Propose expected collisions rates once migration to SPIF configurations is achieved through the VW&D reform measures

Analysis is limited to provincial roads since there is no information available to estimate vehicle kilometres of travel (VKT) by tractor-trailer class on municipal roads.

In addition, information from the CVS and the accident report database showed that:

- Drivers of tractors combined with 2-axle semi-trailers tend to have less experienced drivers than those of heavier vehicles; and,
- 2-axle semi-trailers have relatively more collisions in bad weather than heavier vehicles.

Both, drivers' lack of experience and bad weather are known to increase the likelihood of collisions. As reported later, collision rates of 2-axle semi-trailers are lower than of heavier vehicles. Consequently, if 2-axle semi-trailers were to have the same average driver experience and collision in bad weather as other tractor-trailers, associated collision rates would be even lower. By ignoring driver experience and effect of weather, results of the analysis are conservative.

Analysis of straight trucks and their trailers will be undertaken as part of phase 4 of the project.

5. Results

5.1 Collision Rates

Table 1 shows the estimated collision rates for the six tractor-trailer classes defined for the analysis. Collisions statistics from 1995 to 2002 were used in order to improve statistical reliability. However, VKT were estimated only for 2002. If the general vehicle trend to reduced collisions per million VKT also applies to tractor-trailers, the collisions rates used in this study would be on the low side, since it is known that tractor-trailer VKT grew over the period.

Not all vehicles involved in collisions were fully matched to licensing registration information, because of missing data or vehicles registered out of Ontario. Tractor-trailers with full information cover 50 to 96 percent of collisions depending on the vehicle class.

Collision data by class were proportionately adjusted to incorporate the total number of collisions. The procedure assumes that distribution of missing data among classes is the same as the distribution of collisions with known information. Use of partial information, allowed distribution within certain groups of vehicles. For example, semi-trailers missing information on the number of trailer axles allotted only within semi-trailers classes.

Table 1: Collision Rates By Tractor-Trailer Class

Classes	Collisions / VKT (million)
1&2-axle semi-trailer	0.45
3-axle semi-trailer	0.60
4-axle semi-trailer	0.66
5+axle semi-trailer	1.03
A & C-train double	1.33
B-train double	0.36
Average	0.51

Sources: MTO, Motor Vehicle Accident Report database, 1995 – 2002.
MTO, Commercial Vehicle Survey, 1999-2000.

Semi-trailers with 1 and 2-axles and B-train doubles have the lowest collision rates. Both groups currently have the highest composition of vehicles matching SPIF configurations.

5.2 Performance Measures

Vehicle dynamic performance data is usually determined through computer simulations. Various sources were used to obtain individual performance measures for typical tractor-trailer configurations in each class. The appendix compiles the data used. Vehicles loaded with axles properly adjusted, loaded with axles misadjusted, empty and vehicles operating at SPIF standards are included. The vehicle configurations in the appendix were considered typical within classes. Weighted average performance measures were obtained with the estimated percentages for each loaded condition. Current percentage of SPIF vehicles by class were determined from the CVS to be:

- 1 & 2-axle semi-trailers: 88.5%
- 3-axle semi-trailers: 31.1%
- 4-axle semi-trailers: 0%
- 5+axle semi-trailers: 0%
- A&C-train doubles: 48.8%
- B-train doubles: 73.4%

A high correlation between performance characteristics and collision rates per class was found. However, it was impossible to statistically determine the specific effect of each measure, since there is high collinearity among measures. As an alternative, a performance measure failure indicator was developed.

Each performance measure was given a value of 0, 1, or 2 according to the matching degree with the standard threshold for each measure. Zero meaning the measure is well within the standard; one meaning the measure is around the standard (considered as a recommended value and not as an absolute threshold); and two if the measure is well

above the standard. Low speed offtracking and rear outswing were only given values of 0 or 2 if matching or exceeding the standard. Table 2 depicts the threshold values used for each measure.

Table 2: Performance Measures Threshold Values

Performance Measure	No failure = 0	Around threshold = 1	Failure = 2
SRT	> 0.41	[0.41 – 0.39]	< 0.39
HSOT	< 0.44	[0.44 – 0.48]	> 0.48
LTR	< 0.59	[0.59 – 0.61]	> 0.61
TOT	< 0.78	[0.78 – 0.82]	> 0.82
LSOT	<= 5.688	na	> 5.688
RO	<= 0.20	na	> 0.20

Each performance measure received a weight according to perceived importance on vehicle stability and controllability, as well as on the individual correlation with collision rates. The weights used were:

- Static rollover threshold (SRT): 0.5
- High-speed offtracking (HSOT): 0.5
- Load transfer ratio (LTR): 1.0
- Transient high-speed offtracking (TOT): 1.0
- Low speed offtracking (LSOT): 0.3
- Rear outswing (RO): 0.2

The failure indicator was calculated for each configuration and loaded condition. The last column of the table in the Appendix presents the results.

In addition, an average failure indicator for each vehicle configuration in the Appendix was estimated by taking the indicator for each loaded condition (A, B, C, and SPIF) weighted by the percentage of trips under each condition, indicated by the CVS (second column in the Appendix).

The estimation was repeated for the future condition when full migration to SPIF configurations is expected with the VW&D reforms. Only loaded and empty percentages are applicable since SPIF vehicles will not be improperly adjusted/loaded (axle load equalization will be required).

The indicator for each vehicle class was estimated from the composition of vehicle configurations in the class. Composition percentages for existing classes were obtained from the CVS. Future class composition comes from migration expectations with the VW&D reform project.

The resulting performance measure failure indicator for each tractor-trailer class is shown in Table 3. Failure indicator improvement is higher for classes with current low composition of SPIF vehicles. The introduction of SPIF configurations will have a greater impact within classes with few or no SPIF vehicles in current operation.

Table 3: Performance Failure Indicators

Class	Existing	Proposed	Improvement
1&2-axle semi-trailer	0.00	0.00	0.0%
3-axle semi-trailer	0.45	0.33	25.9%
4-axle semi-trailer	1.52	0.35	77.0%
5+axle semi-trailer	2.97	1.03	65.3%
A & C-train double	3.72	3.23	13.3%
B-train double	1.27	1.20	5.3%

5.3 Future Proposed Collision Rates

There is a high correlation between performance failure indicators and collision rates (the correlation coefficient is 0.90, with a statistical significance of 99%). Based on the statistical relationship and using the expected future value for the failure indicator by class (Table 3), Table 4 proposes the expected collision rates and the resulting percentage improvement by vehicle class.

Table 4: Proposed Collision Rates and Improvement from Existing

Class	Existing 1/	Proposed	Improvement
1&2-axle semi-trailer	0.45	0.45	0.0%
3-axle semi-trailer	0.60	0.56	6.5%
4-axle semi-trailer	0.66	0.57	14.4%
5+axle semi-trailer	1.03	0.79	23.2%
A & C-train double	1.33	1.15	13.3%
B-train double	0.36	0.34	4.3%
Weighted Average	0.51	0.49	4.5%

1/ From table 1

The proposed future collision rates follow different patterns for semi-trailers and double trailers and reflect a conservative approach to the impact of SPIF vehicles on collision rates. Chart 1 shows proposed improvement to semi-trailers, while Chart 2 portrays double trailers. The proposal is based, although does not mathematically follow, the coefficients of change found with the regression analysis between collision rates and failure indicators by vehicle class. The arrows

show how the change in failure indicators are expected to draw collision rates as proposed in Table 4.

Collision rates on non-freeway provincial roads are 40% higher, on average, than rates observed on freeways. Further research should desegregate impact by type of road. However, the authors estimate that compound result on freeways and non-freeways would provide similar estimates as reported in this paper.

Chart 1: Expected Changes to Semi-Trailer Collision Rates and Failure Indicators

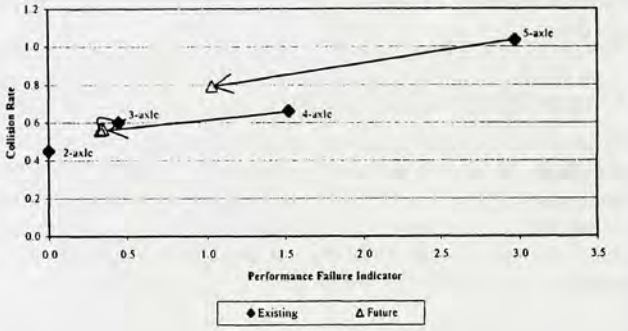
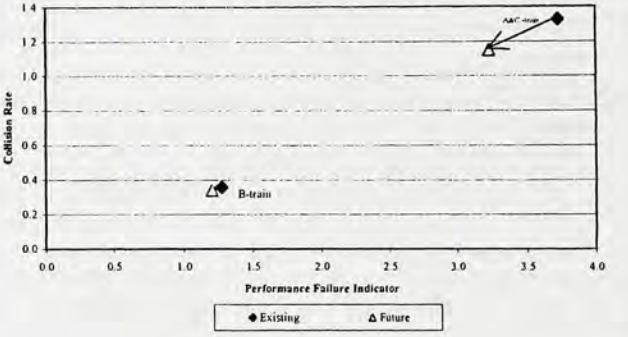


Chart 2: Expected Changes to Double Trailer Collision Rates and Failure Indicators



5.4 Reduction in Collisions

Using collisions statistics for 2002, it is then possible to estimate the potential reduction in collisions attributable to the VW&D program. In 2002, tractor-trailers were involved in 5,360 collisions (after adjusting for collisions where the vehicle was unknown).

The proposed improvement of 4.5% in collision rates would represent a reduction of about 240 collisions per year. Impact is assumed proportional to existing distribution of fatal, personal injury and property damage only collisions. The potential reduction would be achieved once full migration to SPIF vehicles is complete. Full migration is expected in about 20 years, as the VW&D reform grandfathered existing vehicles during their reasonable operating life.

Future measurement of the impact of SPIF vehicles on road safety would lack precision as there are other variables affecting road safety. However, the authors recommend a revision of collision rates by vehicle class in five-year intervals to validate results. As already mentioned, the number of truck collisions has been decreasing at least since 1993. Although the trend is expected to continue, there is no doubt that the introduction of vehicles with better dynamic performance measures will have a significant contribution in reducing tractor-trailer collision rates.

6. Conclusions

Using eight years of tractor-trailer collisions statistics in Ontario, this paper provides evidence of a strong correlation between performance measures and collision rates of tractor-trailers.

The analysis grouped vehicles with highly differentiated dynamic performance measures. Tractor-trailer classes with performance measures well within CCMTA/RTAC thresholds clearly perform better than classes with measures above thresholds.

A failure performance indicator developed by vehicle class was found highly correlated to collision rates. Correlation was then used to propose expected collision rates when migration to SPIF vehicles is completed, as part of the VW&D reform program.

Potential reduction in collisions would grow to around 240 with full migration to SPIF vehicles. This represents a 4.5% reduction in tractor-trailer collision rates.

Appendix
Tractor-Trailer Performance Measures

Configuration	Condition*	Note	GVW (kg)	SRT >0.400	HSOT <0.460	LTR <0.600	TOT <0.800	LSOT <6.0m	RO <0.2m	Failure Indicator
S2	A - 73%	1	39,500	0.437	0.303	0.384	0.298	5.906	ok	0
	B - 0%	1	39,500						ok	
	C - 27%	1	14,500	0.498	0.187	0.36	0.237	5.906	ok	0
	SPIF	1	39,500	0.437	0.303	0.384	0.298	5.906	ok	0
S12 (tanks/dumps)	A - 29%	2	48,500	**0.393**	0.34	0.447	0.372	5.483	ok	0.5
	B - 29%	2	48,500	**0.306**	0.36	0.433	0.384	5.483	ok	1
	C - 42%	2	15,500	0.498	0.187	0.36	0.237	5.483	ok	0
	SPIF	4	48,500	0.486	**0.478**	0.454	0.642	4.709	ok	0.5
S111 (tanks/dumps)	A - 25%	3	49,500	**0.373**	0.409	0.486	0.482	4.793	ok	1
	B - 25%	3	49,500	**0.351**	0.442	0.508	0.519	4.793	ok	1.5
	C - 50%	3	15,500	0.468	0.187	0.36	0.237	4.793	ok	0
S3 (other 3 axle)	A - 70%	4	48,500	0.402	0.321	0.42	0.335	5.577	ok	0.5
	B - 0%	4	48,500						ok	
	C - 30%	4	15,500	0.482	0.34	0.364	0.266	5.577	ok	0
	SPIF	4	48,500	0.402	0.321	0.42	0.335	5.577	ok	0.5
S12 (other 3 axle)	A - 35%	5	48,500	**0.393**	0.34	0.447	0.372	5.483	ok	0.5
	B - 35%	5	48,500	**0.306**	0.36	0.433	0.384	5.483	ok	1
	C - 30%	5	15,500	0.498	0.187	0.36	0.237	5.483	ok	0
	SPIF	4	48,500	0.486	**0.478**	0.454	0.642	4.709	ok	0.5
S13	A - 35%	6	55,500	0.446	0.345	0.534	0.601	5.577	ok	0
	B - 35%	6	55,500	**0.351**	0.442	0.508	0.519	5.577	ok	1.5
	C - 30%	6	16,200	0.482	0.34	0.364	0.266	5.577	ok	0
	SPIF 48'-35%	24	57,199	0.431	**0.498**	0.514	0.658	4.971	0.045	1
	SPIF 53'-35%	6	55,500	0.471	0.328	0.489	0.62	4.822	ok	0

Appendix (continued)

Configuration	Condition*	Note	GVW (kg)	SRT >0.400	HSOT <0.460	LTR <0.600	TOT <0.800	LSOT <6.0m	RO <0.2m	Failure Indicator
S22	A - 30%	7	57,500	**0.332**	**0.485**	**0.625**	0.772	4.166	ok	4
	B - 30%	7	57,500	**0.305**	**0.538**	**0.692**	**0.878**	4.166	ok	6
	C - 40%	7	16,200	0.498	0.187	0.36	0.237	4.166	ok	0
S1111	A - 38%	8	57,500	**0.332**	**0.485**	**0.625**	0.772	4.166	ok	4
	B - 38%	8	57,500	**0.305**	**0.538**	**0.692**	**0.878**	4.166	ok	6
	C - 24%	8	16,200	0.498	0.187	0.36	0.237	4.166	ok	0
S121	A - 39%	9	57,500	**0.332**	**0.485**	**0.625**	0.772	4.166	ok	4
	B - 39%	9	57,500	**0.305**	**0.538**	**0.692**	**0.878**	4.166	**fail**	6.4
	C - 22%	9	16,200	0.498	0.187	0.36	0.237	4.166	Ok	0
S131	A - 39%	13	61,735	**0.398**	**0.557**	0.58	0.793	4.177	**0.254**	2.9
	B - 38%	21	61,735	**0.367**	**0.582**	**0.643**	**0.852**	4.177	**0.254**	6.4
	C - 23%	23	21,600	0.482	0.34	0.364	0.266	4.177	**0.254**	0.4
	SPIF	14	61,375	0.427	**0.533**	0.569	0.746	3.979	0.173	1
S113	A - 34%	15	61,508	**0.392**	**0.529**	**0.601**	**0.813**	5.165	0.033	3.5
	B - 34%	21	61,508	**0.374**	**0.584**	0.586	0.785	5.165	0.033	3
	C - 32%	23	22,800	0.482	0.34	0.364	0.266	5.165	0.033	0
	SPIF	16	61,735	0.435	**0.561**	0.544	0.747	5.057	0.039	1
S141	A - 30%	17	59,466	0.429	**0.511**	**0.612**	**0.859**	4.059	**0.235**	5.4
	B - 30%	21	59,466	**0.398**	**0.530**	**0.639**	**0.858**	4.059	**0.235**	5.9
	C - 40%	23	23,500	0.482	0.34	0.364	0.266	4.059	**0.235**	0.4
	SPIF	18	61,346	0.404	0.452	**0.677**	0.786	4.003	**0.248**	4.4
S114	A - 25%	19	59,466	0.427	**0.48**	0.594	0.785	4.977	0.059	2.5
	B - 25%	21	59,466	0.403	**0.526**	0.567	0.738	4.977	0.059	1.5
	C - 50%	23	26,300	0.482	0.34	0.364	0.266	4.977	0.059	0
	SPIF	20	61,598	0.449	**0.49**	0.551	0.744	4.905	0.039	1

Appendix (continued)

Configuration	Condition*	Note	GVW (kg)	SRT >0.400	HSOT <0.460	LTR <0.600	TOT <0.800	LSOT <6.0m	RO <0.2m	Failure Indicator
A - 12S11-1	A - 40%	10	43,000	0.446	**0.478**	**0.785**	**0.852**	4.881	ok	4.5
11S1-11	B - 40%	22	43,000	0.401	**0.526**	**0.864**	**0.937**	4.881	ok	5.5
	C - 20%	10	15,800	**0.311**	0.375	**0.64**	0.575	4.881	ok	3
	SPIF	10	43,000	0.446	**0.478**	**0.785**	**0.852**	4.881	ok	4.5
A - 12S21-2	A - 30%	11	61,500	**0.355**	**0.552**	**0.97**	**1.132**	4.293	ok	6
12S2-12	B - 30%	22	61,500	**0.323**	**0.607**	**1.067**	**1.245**	4.293	ok	6
	C - 40%	11	17,900	0.498	0.362	0.484	0.559	4.293	ok	0
	SPIF	11	49,000	0.434	0.431	**0.795**	0.783	4.293	ok	3
B- 12S3S2	A - 35%	12	64,500	**0.359**	**0.488**	0.572	0.722	4.473	ok	2
	B - 25%	25	64,500	**0.341**	**0.512**	**0.601**	0.758	4.473	ok	3
	C - 40%	12	19,300	0.462	0.344	0.292	0.358	4.473	ok	0
	SPIF	12	64,500	**0.359**	**0.488**	0.572	0.722	4.473	ok	2

* Condition: A - loaded, lift axles deployed and properly adjusted
B - loaded, lift axles raised or improperly adjusted or improperly loaded
C – empty
SPIF – loaded configuration with load axle equalization and self-steer axles, if applicable

Bold numbers indicate performance measures over the recommended standard

Notes:

1 RTAC, "Influence of VW&D on Stability and Control - V2", July 1986, pp 29 Appendix D (Configuration 1.0, Variation 1.00)
2 RTAC, pp 38 Appendix D (Configuration 1.5, Variations 3.03 and 2. 03)
3 RTAC, pp 38 Appendix D (Configuration 1.3, Variations 3.01 and 2.02)

4 Billings, JR "Evaluation of the Dynamic Performance of 16.20 m Self-Steer Triaxle Semi-trailers", Ontario Trucking Association, May 2003, pp 13 (Average of 53' trailer - 100" + 72" spread with 38,000 kg and 52,000 kg GVW)
5 RTAC, pp 38 Appendix D (Configuration 1.5, Variations 3.03 and 2. 03)
6 Billings, JR "Evaluation of the Performance of 16.20 m Quad-axle Semi-trailers", Ontario Trucking Association, March 1999, pp 5 (48' trailer - 60" + 60" spread and RTAC Configuration 1.3, Variation 2.02)
7 RTAC, pp 37, Appendix D (Configuration 1.4, Variations 1.00 and 2. 02)
8 RTAC, pp 37, Appendix D (Configuration 1.4, Variations 1.00 and 2. 02)
9 RTAC, pp 37, Appendix D (Configuration 1.4, Variations 1.00 and 2. 02)
10 RTAC, pp 44, Appendix D (Configuration 2.2, Variation 1.00)
11 RTAC, pp 44, Appendix D (Configuration 2.1, Variation 4.02 and 5.10 and 1.00)
12 RTAC, pp 44, Appendix D (Configuration 3.1, Variation 4.01)
13 NRC, "Performance of Infrastructure - Friendly Vehicles", October, 2003. Table 6, pp 27 (CG High, Lift down for high speed measures and up for low speed measures)
14 NRC, Table 44, pp 77 for high speed measures and Table 40, pp 73 for low speed (CG High, Case 1b)
15 NRC, Table 4, pp 25 (CG High, Lift down for high speed measures and up for low speed measures)
16 NRC, Table 30, pp 61 (CG High, Case 2d)
17 NRC, Table 10, pp 31 (CG High, Lift down for high speed measures and up for low speed measures)
18 NRC, "Assessment of the Dynamic Performance SPIF Configurations for Hauling Municipal Waste" Draft Report, December, 2004. Table 22, pp 27 (Case 1b)
19 NRC, Table 8, pp 29 (CG High, Lift down for high speed measures and up for low speed measures)
20 NRC, Table 50, pp 84 (CG High, Case 1e)
21 High speed performance measures assumed as average between lift axles down and up
22 Performance measures assumed as 10% worst that alternative A (axles deployed and properly adjusted)
23 Performance measures assumed the same as empty tridem semi-trailers (S3)
24 NRC, Table 26, pp. 50 (CG High, average of cases 1a, 2a, and 3a)
25 Performance measures assumed as 5% worst that alternative A (axles deployed and properly adjusted)

A Framework For Winter Road Maintenance Optimisation

Mathieu Trudel, MASc Candidate
Department of Civil Engineering
University of Waterloo
mtrudel@uwaterloo.ca

Abstract

Advances in winter road maintenance practice have enabled new levels of service and safety to be reached while simultaneously reducing the financial and environmental impacts of maintenance operations. However, research in management and planning of vehicle routing and dispatching have generally not matched pace with these advances. We describe a new framework for optimising the timing and routing of vehicle dispatching during a storm event, subject to level of service and fleet size restrictions. This optimisation can seek to minimise cost of providing service, maximise effective service levels, or any other arbitrary function which meets certain basic criteria. This model is intended to be used in both decision support and simulation capacities, and can utilise various meteorological and service models in a modular way to enable future research or be tailored to specific deployments. Mathematically, the model is implemented as an integer linear program, the formulation of which is geared towards extensibility and computational efficiency. This paper provides an overview of the framework, structure and prototype implementation of the proposed model.

Introduction

Winter maintenance is a significant challenge to most transportation agencies in Canada. Transport Canada has estimated that provincial and municipal governments spend more than $CDN 1 billion dollars every year to keep roads and highways clear of snow and ice and to maintain safe travel conditions[1]. This work includes spreading over 5 million tonnes of salt annually onto Canadian roads – causing a growing concern over its environmental impact. Although modern research into alternate anti-icing strategies greatly helps to mitigate this, most research in the field focuses on low-level operational concerns such as alternative chemical agents, application rates, and pavement condition factors. Beyond providing improved weather data and vehicle instrumentation for supervisors, little has been done to improve winter road operations at a supervisory level, where great benefit can be expected.

Many of these supervisory decisions must be made with time and resource constraints, such as when to start sanding and plowing, what operation routes and sequence to follow, and how much chemical agent to apply. The situation becomes even more challenging when an anti-icing strategy is deployed, as the effectiveness of this strategy depends largely on the ability to make accurate prediction on weather and pavement conditions and to determine the precise timing and rate for applying treatment chemicals. Careful planning of such applications repays large dividends.

Conventional approaches to the scheduling and assessment of winter road maintenance activities are highly empirical in nature. Decisions pertaining to when and where to deploy service vehicles are typically made by a human supervisor based on static weather forecasts, first hand reports by deployed vehicles, and personal experience. However, despite these apparent shortcomings, such decisions are almost always correct (or at least adequate for the task at hand).

Despite the fact that there is nothing 'wrong' with current maintenance scheduling practice per se, there are still improvements to be

realised from a formal maintenance scheduling model. Specifically, such a model allows for knowledge to be quantified and transferred between supervisors and districts. It also allows for simulation of various scheduling policies and parameter changes without having to carry them out in real life (for example, the expected benefit of adding a plow vehicle to a fleet can be estimated in quantifiable terms, entirely on paper).

The Winter Road Maintenance Scheduling Program

The Winter Road Maintenance Scheduling Program (WRMSP) is a system currently being developed by the Civil Engineering Department at the University of Waterloo. It is designed to mathematically model the operations of a winter maintenance fleet, and provide a list of recommended vehicle dispatches based on optimising a given function (such as overall snowfall accumulation, operational cost minimisation, or virtually any other function which meets some basic mathematical criteria). The WRMSP can also be used in an analysis capacity to estimate the impact of a given treatment plan on a road network, or to evaluate the expected operational impact of a change in road topology, fleet size, or several other factors.

The WRMSP currently exists in a prototype form, and is the subject of ongoing research. As the ultimate goal of this research is the introduction of the WRMSP into a production environment, we are seeking input from practitioners and other researchers as to the eventual structure of the WRMSP in practice. Towards this end, this paper provides a general overview of the WRMSP, and does not discuss the inner workings of the model beyond the extent needed to understand its general structure. For more details of the WRMSP beyond those discussed here, please contact the author directly. Also, since the WRMSP is still very much a work in progress, this paper is primarily informational in nature, and does not present the findings of a study brought to its conclusion.

WRMSP Structure

The WRMSP can be formally defined as follows. Consider a road network consisting of a set of road segments. The network has been covered by a set of pre-defined maintenance routes with each representing an itinerary that a service vehicle may take through the road network. It is assumed that all routes are cyclic (that is, they begin and end at the same location, namely their home patrol yard). Note also that routes are not necessary mutually exclusive (that is, there could be overlapped portions with some road segments belonging to multiple routes). A fleet of vehicles is available to provide winter maintenance services such as plowing and salting during a snowstorm.

The problem solved by the WRMSP is to develop an operations plan for available service vehicles that specifies route assignment, service type, and the corresponding start time. The schedule proposed by the WRMSP should take into account the following requirements:

- At no time should the safety of the public be at avoidable risk. This corresponds to maintaining a minimal condition on all roads in accordance with the relevant level of service policy;
- At no time should there be more vehicles scheduled to be on the road than there are vehicles in the fleet;
- The given optimisation function (minimise storm impact, minimise cost of service, or other) should be used as the benchmark to determine an optimal solution.

The WRMSP models a road network for a fixed time window into the future. This window is broken up into discrete intervals of equal length. For example, a typical instance of the WRMSP may model operations for 3 hours into the future, breaking that window into eighteen 10 minute periods. These periods represent the finest granularity that can be modelled by the WRMSP.

The WRMSP deals with large amounts of input and output data representing many different quantities. To ease the management of this data, three principal indices are used to define data;

- A time index (t). This is the number of time intervals into the model that a quantity corresponds to.
- A road index (i). This is a physical segment of road which a quantity corresponds to.
- A route index (k). This is a particular pre-selected route which a quantity corresponds to.

Most input and output parameters are specified in terms of these indices. For example, weather forecasts can be specified on a road by road basis, and can vary over time, so weather information is given to the WRMSP as a matrix of values indexed by road (i) and time (t).

System Inputs

The WRMSP requires the following major input parameters;

- Weather forecasts for the road network being modelled. As previously indicated, weather forecasts are allowed to vary over space (on a road by road basis) and also by time (on a time period by time period basis). The actual value of the weather forecast is a single number for each (i,t) pair, which represents how weather affects road i during time interval t. These values are represented by $W_{i,t}$.

  The exact metric used for deriving these values from a weather forecast is largely immaterial, provided that they represent a useful estimate of road condition, and also represent a value which can be altered by treatment actions. For example, depth of snow on the road is a useful metric for illustration; weather impact for each time period is the amount of snow which falls on the road in question during the time period, and the condition of each road is represented by the depth of snow currently on that road. This value can

also easily be directly altered by treatment events, since plowing a road removes (most of) the snow on the road at the time of service.

As weather forecasts exist only as an input to the WRMSP, the particular package used to provide these forecasts is flexible. It is sufficient to assume that a 'black-box' forecasting model exists which can provide the WRMSP with time and road indexed forecasts in an efficient and automated manner. There is already a large body of excellent research into road weather forecasting (such as the weather component of the Maintenance Decision Support System being developed by the FHWA[2]), and we elect to build on top of that work rather than reinvent it.

- A list of routes, and the path each of them takes through the road network. This is necessary to correlate vehicle dispatches along a specific route to that vehicle's arrival time at each road section along the route.

- The total number of vehicles available for dispatch at any time. This ensures that we never indicate a schedule which requires more vehicles to be on the road than we have on hand.

- An upper limit on road condition for each road segment. This allows us to represent level of service obligations in a formal way, and allows for variation of these levels on a road by road basis to model different classes of roads.

Other parameters exist to indicate priority status for specific roads, already-deployed vehicles (to allow for hot-starting of the WRMSP), and other secondary parameters.

System Outputs

The outputs of the WRMSP are as follows;

- A list of recommended dispatches, termed X. Recommendations are indexed on route (k) and time (t), and indicate at which time a vehi-

cle should depart, and on which route (values of X are set to 1 if a vehicles should be dispatched, and 0 otherwise). Note that dispatches are made without respect to a particular vehicle; the WRMSP indicates that a vehicle should depart, but deciding which vehicle and operator to dispatch are outside its scope. This information is intended to be used by maintenance supervisors as a decision support system for dispatch timings.

- A list of estimated road conditions by time, given the weather forecast and recommended treatments. These estimates are indexed by road (i) and time (t), and are represented by $C_{i,t}$. This information is useful to entities such as emergency response dispatchers, as well as the general public.

A flowchart indicating the overall data flow of the WRMSP is shown in Figure 1.

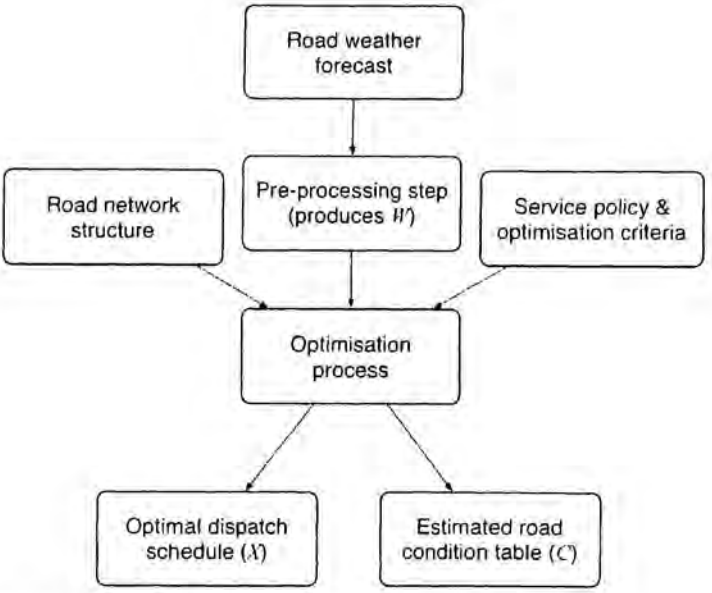


Figure 1 - Data flow within the WRMSP

A Brief Overview of the WRMSP's Inner Structure

Internally, the WRMSP maintains an estimate of the condition of each road section at each time period, based on weather forecasts W' and treatment events X, which is represented by C. In functional terms,

$$C_{i,t} = f(C_{i,t-1}, W_{i,t}, X)$$

that is, road condition at any time period is entirely determined by the road condition in the preceding time period, the weather forecast for the current time period, and any treatment events which may be scheduled to occur on the road section during the current time period.

Of all of these input variables, the only one which is independently controllable is X. Furthermore, since the layout of all routes in the model is fixed, the value of X for any road section is entirely determined by when vehicles leave the patrol yard (in other words, by when they are dispatched). Thus, given a fixed weather forecast, we can manipulate the vehicle dispatch schedule in order to minimise the overall impact of the storm (which is represented by the sum of all $C_{i,t}$ values). If desired, we can also optimise for the cost of providing treatment (since we know how often vehicles depart on routes of a known length, we can obtain the total vehicle distance travelled).

Integer Programming

As mentioned previously, the WRMSP utilises a mathematical structure known as an Integer Linear Program (ILP) to optimise its proposed schedule. Briefly, ILPs consist of a set of *variables* (such as $x,y,z,...$), an *objective function* based on those variables, and a set of *constraints* which restrict the values the variables may take. ILPs are then *solved* by finding the set of variable values which *optimise* the value of the objective function while simultaneously not violating any of the constraints imposed on the system. This set of variable values is termed the *optimal solution* of the ILP.

In actual fact, most ILPs quite large in scope, often involving thousands of variables and constraints. For example, the ILP used to solve the WRMSP consists of approximately 800 variables and 1200 constraints for a typically sized WRMSP instance.

Those familiar with Integer Programming may notice that problems of this size are generally intractable. This is overcome in the case of the WRMSP by utilising heuristics which exploit structural properties of the WRMSP, specifically the fact that the ILP used within the WRMSP is easily solvable when scheduling for a single vehicle. Using this observation, we iteratively add vehicles one at a time into the model, stopping when we have added the number of vehicles available in the fleet. Although it is generally not possible to establish

strict bounds on the accuracy of this approximation, empirical results suggest that this heuristic produces very effective solutions.

Example Implementation

A preliminary web-based interface to the WRMSP has been developed as part of the model's development. This interface allows operators to select the profile of the jurisdiction for which they want to run the model, provide weather data (automatic gathering of this weather information is planned as future work), and quickly obtain both a schedule of deployments, and a set of graphs estimating road condition over time for each section of road in the network. Using a web based system for this purpose follows recommended practice[3] and also allows for multiple jurisdictions to use the same service (adding new jurisdictional profiles into the server is as easy as creating a configuration file for the profile, based on an existing template). Selected screenshots from this interface are shown in Figure 2 and Figure 3 below.

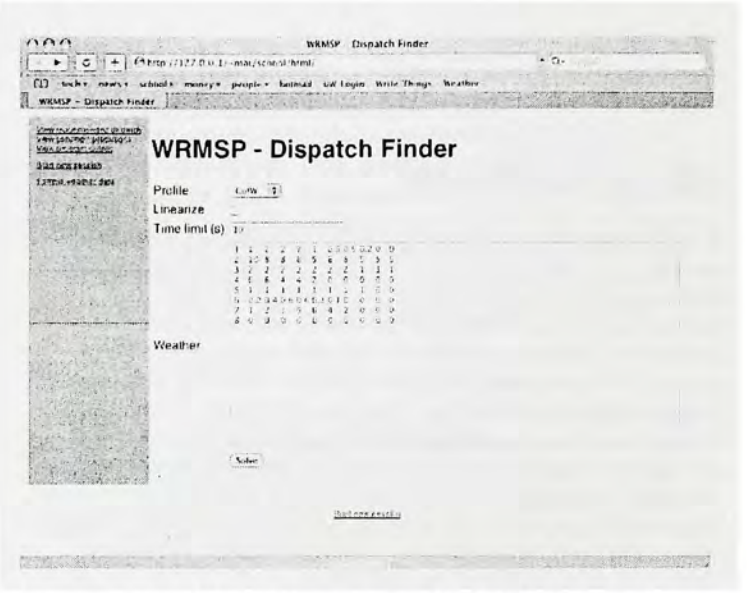


Figure 2 - Typical data entry screen from WRMSP web interface

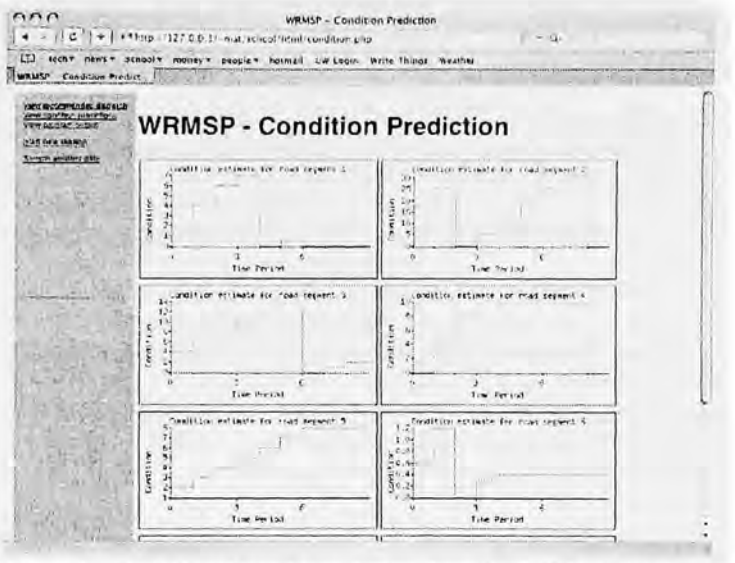


Figure 3 - Typical road condition display from WRMSP web interface

The WRMSP itself is written in a mix of AMPL and C, while the web based interface shown here is implemented in PHP. Thus, the WRMSP is easily extensible for future research and custom deployments.

Future Work

Being a work-in-progress prototype, the WRMSP still has many rough edges. Future work will focus on several aspects of the program, including;

- Making it easier to use the WRMSP platform in a production environment, and also as a basis for future research. This involves improving the weather forecast gathering portion of the WRMSP, im-

proving documentation of the model, and ensuring that both the practical and theoretical aspects of the model are both correct and extensible.

- Adding support for various types of winter maintenance operations. In addition to plowing, anti-icing, salting, and sanding support are all planned for inclusion in the WRMSP.

- Calibrating the WRMSP against real-world scenarios to validate its accuracy.

- Gathering input from practitioners and other researchers as to the features and outputs they would most like to see from the WRMSP.

As previously mentioned, this paper gives only an overview of the WRMSP. Those interested in learning about the WRMSP in more detail should consult the author's MASc thesis[4].

Conclusion

In summary, the WRMSP model allows for the efficient estimation of the overall system-wide effectiveness of a given treatment plan for an arbitrary weather event. This ability proves useful both for developing a plan of action in real-time during a storm event, and also for estimating the effects of altering various system parameters such as fleet size or treatment policies. In addition to providing recommended schedules for vehicle dispatch, road condition estimates are also provided, which can prove useful for emergency response and public information purposes. Although the WRMSP is still in its infancy, our hope is that it will mature into a useful tool for both researchers and practitioners.

References

[1]Andrey, J., Mills, B., Jones, B., Hass, R., Hamlin, W. Adaptation to Climate Change in the Canadian Transportation Sector. Natural Resources Canada, Adaptation Liaison Office, Ottawa. 1999
[2] http://www.rap.ucar.edu/projects/rdwx_mdss/
[3] Castle Rock Consultants. RWIS Data Integration Guidelines, p. 43 Aurora Program Report, St. Paul. 2001
[4] A Framework For Modeling Winter Road Maintenance Operations, MASc Thesis. Trudel, Mathieu. Department of Civil Engineering, University of Waterloo. 2005

The Role of Transportation in Trade Competitiveness: Understanding the Policy Context through Quantitative Measures

Nilam Bedi, Ministry of Municipal Affairs and Housing and Atif Kubursi, Mcmaster University

Introduction

It is certain now that the international economic environment Canada will face in the new millennium will be dramatically different from the world in the 1980's, and even that in the 1990s. International competition is intensifying, space and time are being cumulatively compressed and the basis of economic success is changing. Massive political and economic restructuring are reinforcing technological breakthroughs. Successive rounds of trade liberalisation under the old General Agreement on Tariffs and Trade (GATT) and the newly created World Trade Organisation (WTO), along with increasing cross-border investment and financial deregulation have led to major changes in goods production, international trade, investment and the proliferation of multinational corporations.

Productivity, innovation and efficiency are the triangular fundamentals of competitiveness. Ontario will need to address lagging productivity problems created, as some suggest, by the shield that a depreciating dollar has afforded Canadian producers (McCallum (1999). Some of the issues and concerns about lagging competitiveness in Ontario include the following:

- While Ontario leads most Canadian jurisdictions in GDP growth, it lags in GDP per capita and productivity growth among competing international jurisdictions:

- The Ontario Task Force on Competitiveness, Productivity and Economic Progress ranked Ontario 13th out of 16 peer states and Quebec in GDP per capita in 2003 (http://www.competeprosper.ca/)
- The shortfall is attributed to lower productivity growth in Ontario

- Neighbouring Great Lake jurisdictions are pursuing aggressive policies and investment programs to boost economic productivity. Many U.S. Great Lakes states have spent record amounts on infrastructure development and rehabilitation in

The policy context above does not operate in a vacuum. Rather it is intimately linked with the economic role and contributions of transportation activity. These are substantial but are thwarted by challenges and missed opportunities. A neglected issue but one of serious implications relates to transportation intensity of exports and imports. This is now complicated by the rise in transaction costs presented by border delays and frictions.

The Role of Transportation

When it comes to trade, the costs of transportation are a key consideration. The productivity of this sector has long been of interest to policy makers and researchers. In the U.S. the dominance of transportation sector by trucking has been attributed to the rapid productivity gains of this mode. Boyer (1999) asserts "lower transportation costs have the potential of increasing economic efficiency through regional specialization in the manner described by international trade theory".

Boyer and Burks (2004) have estimated the source of productivity gains in the trucking sector as follows:

- 25% resulting from changes in traffic composition (for example towards longer hauls where trucks are more efficiently used)
- 15% due to declines in the wages of drivers
- 11%because of declines in the cost of owning a rig 18% to the declining price of diesel fuel
- 19% to an increase in ton-miles per year from the truck and driver. And
- the remaining 10% is attributable to other factors, notably to an increase in the fuel economy of over-the-road trucks.

The Policy Challenge

Given that competitive transportation cost are of paramount importance to the trade-dependent Ontario, the concerns over the ability of the transportation sector to keep its costs down are very real. The concern is also legitimate as researchers have pointed out that the sources of past efficiency gains do not portend a rosy outlook. It has to be remembered that several factors act in concert to determine the cost and productivity of the both the infrastructure and the transportation service sector that utilizes the infrastructure::

- infrastructure investments ensure that there is the capacity in the network to avoid congestion and, therefore, time and travel costs of goods
- regulatory actions determine the factor input costs – e.g. vehicle weights and dimensions will affect the tonne-km cost of goods
- broader economic factors and market conditions (e.g. availability of drivers and fuel prices) affect other input costs

How Important is the US Canadian Border to Trade?

The Canada-U.S. Free Trade Agreement (FTA) of 1989, recast in 1994 to include Mexico, predictably strengthened the Great Lakes region's bi-national integration. Ontario and the Great Lakes states

(U.S.) represent the manufacturing core of their respective economies. Ontario accounts for about 55% of Canada's manufacturing output, while the Great Lakes states account for about 25% of U.S. manufacturing.

Because the majority of the region's (Great Lakes-Ontario) trade is comprised of heavy manufacturing and natural resources, vast north-south rail and road networks have developed to accommodate it.

Considerable efforts were made to create a seamless transportation system in the region. Recent manifestations of these efforts include the merger of Canadian National Rail with Illinois Central Rail and the proposed merger by Canadian National Rail and Burlington Northern Santa-Fe Rail, which would link Canada-U.S.-Mexico. Improvements in highway infrastructure and telecommunication technology (such as global positioning systems), in addition to the increasing needs by manufacturing firms for flexible and timely delivery, have also made trucking the favoured mode of transport across the border.

The bi-national Great Lakes economy stands today as a prototype of what economic regions will increasingly look like in the global economy. The region has a long history of highly integrated markets, linkages between industries and cooperation among its people. Lower barriers to trade and investment have recently intensified trade.

The tragic events on September 11, 2001 created a new situation at border crossings with long waiting lines and heightened security concerns. The situation culminated in the endorsement by the two countries of a 30-point Smart Border Declaration on December 12, 2001. Canada's strategy was to lessen U.S. concerns over border issues and to persuade the U.S. authorities that the Canadian government is a partner in making the border more secure for both countries. However, Canada wanted the U.S. to recognize that security concerns cannot ignore economic concerns and that unilateral decisions on border issues ultimately make the movement of people and goods less efficient for both countries. For its part,

Canada has been working hard at closing the credibility gap with the U.S. on security issues.

Economic concerns do not trump security concerns and security concerns have serious economic impacts

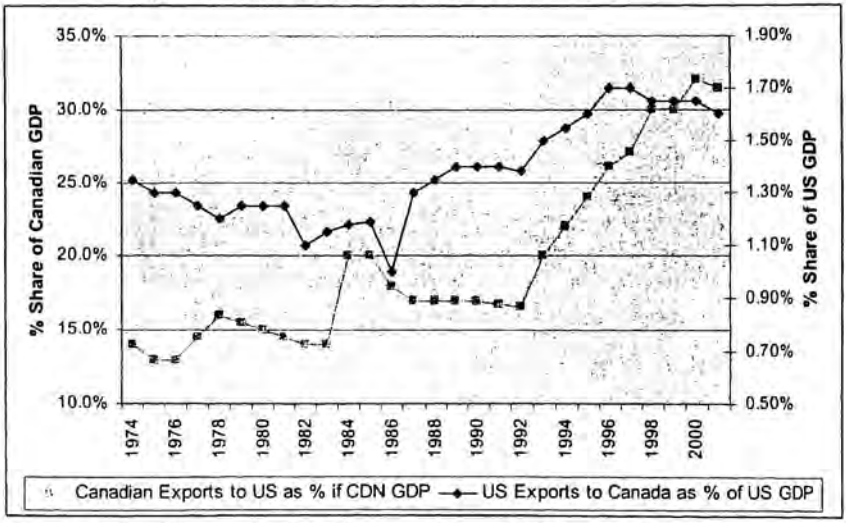
Both countries now recognize the need to formalize a "confidence zone" at the border to enhance both security and trade, while building on a long history of trade and cooperation between the two neighbours.

The following facts underpin the need for this confidence zone:

- In 2004, Canadian exports through Ontario to the U.S. exceeded $132 billion while more than $128 billion in trade flowed from the U.S. into Ontario
- There is increasing integration of the North American economy.
- The importance of international trade in Ontario's economy is increasing. For example, the share of exports in Ontario's GDP soared from 29.4% in 1989 to almost 70% in 2001.
- In 2000, 93% of Ontario's international goods exports were destined for the U.S.
- Ontario shares 14 border crossings with the U.S., which are responsible for the entry or exit of 60% of Canada's total trade.
- In order for Ontario to compete in the North American market, it is essential that its border infrastructure is both sufficient and efficient to meet current and future needs.

- Border investments are also crucial in ensuring the safe passage of people and goods while minimizing economic penalties associated with increased security.
- Canadian exports to the U.S. exceeded 30% of its GDP, whereas U.S. exports to Canada, its largest trading partner, were less than 2% of the U.S. GDP. This asymmetry is a crucial element in negotiating border concerns.

Figure 1 – Growth in US-Canada Trade



Source: Federal Reserve Bank of Chicago. Chicago Fed Letter, July 2002.

Transportation Intensities: Can They Compensate For Border Transaction Costs?

The impact of transportation cost (transport intensity) on the competitiveness of industries in the domestic and international markets is a now a major policy concern in Ontario and Canada at

large. This concern is exacerbated by the increased dependence in Ontario and Canada on exports to the US and by the new complications that emerged after September 11, 2001.
The "transport cost intensity" for an industry is the cost of transport as a percentage of the total cost of its output. To estimate the transportation cost, we need to use the concept of transportation margins. In the Use matrix (the technical coefficients matrix) of an input output table, commodities are valued at producers' prices and the added costs incurred by the purchaser are shown as margins. There are typically seven margins in input output tables--transportation, retail, wholesale, tax, storage, gas, and pipelines margins. The transportation cost of purchasing a commodity in the Use matrix is shown as a purchase from the transportation industry. It is also shown as a transportation margin when it is embedded in the cost of other inputs purchased by the firm to produce its output.

The five sets of input output tables for Ontario in 1979, 1984, 1990, 1996 and 1999, respectively provide invaluable information on the intensity of transportation cost in the value of output of different industries. The input output system of accounts defines the transportation industry in a restrictive sense, as the sum of revenues received by carriers from the users of transportation services. The latter includes commercial (for hire) air, water, rail, and truck, as well as services incidental to transportation. Private freight transportation services, such as private trucking, are considered as part of the internal operations of the establishments and whose costs are included implicitly in the output of the transportation industry.

Transport Cost Intensities and Value Added Contributions

It is clear that in 1979, the transportation sector, other services and the forestry sector had the highest transport cost intensities in Ontario. It is also interesting to note that while wood, furniture and printing do not purchase a great amount from the transportation sector directly, they pay large transportation margins (see Figure 2). Other Services show a transport cost intensity of $0.14 without the transportation margins and about $0.15 when the latter are added. Transportation

activity shows a total transport cost intensity of $0.07 and forestry $0.063. Non-metallic minerals and refined petroleum are next. Perhaps more interestingly, sectors with high transport cost intensities are also sectors with high direct value added coefficients. This fact suggests that sectors with extensive dependence on transportation services, and therefore show high transport cost intensities, are also high-income contributors to the economy.

Sectors with extensive dependence on transportation services are also high- income contributors to the economy

Not much has changed over time in the ranking of sectors by transport cost intensities between 1979 and 1990. Almost the same type of industries that showed high transport cost intensities in 1979, reserve similar rankings in 1984 and 1990. However, the magnitudes of the intensities have changed markedly. Transportation, other services, forestry, mining and refined petroleum show relatively high transport cost intensities, but their values are far below those in the 1979 with the exception of the transport sector itself .

Ranking of sectors with high transport cost intensities has remained constant over a ten-year period

In 1996 and 1999, the rise in tourism expenditures and travel reflects the growing importance of this sector to Ontario. In both 1996 and 1999, this sector shows the second highest transport cost intensity in the economy.

It is quite revealing that most of the Ontario export industries are among those sectors with relatively high transport cost intensities. When the export sectors are aggregated, the transport cost intensity of

international exports is about 2% and that of inter-provincial exports is about 2.5%. The relative shift in Ontario's trade destinations from inter-provincial to international (US) that has been observed after 1989, is perhaps a reflection of the impediments a small fraction of transport costs can place on exports. Intensified international competition and the expansion of markets for Ontario products suggest that efficiencies in transport are critical to maintaining Ontario's international competitiveness.

Figures 2-6 show the relative magnitudes of total transport cost intensities by industry. The results indicate that bulky natural resources and high valued commodities require large transport inputs in their production processes, either directly or through purchased inputs. Conversely, textiles, agriculture, electrical products and clothing have low transportation costs per unit of output, which suggest that these industries are low-income generators for the Ontario economy.

Figure 2 – Transportation Costs per Dollar of Total Cost, 1979

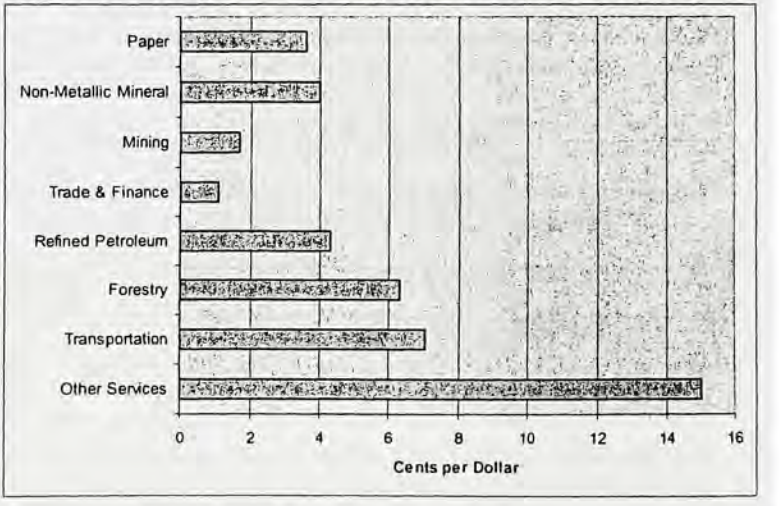


Figure 3 – Transportation Costs per Dollar of Total Cost, 1984

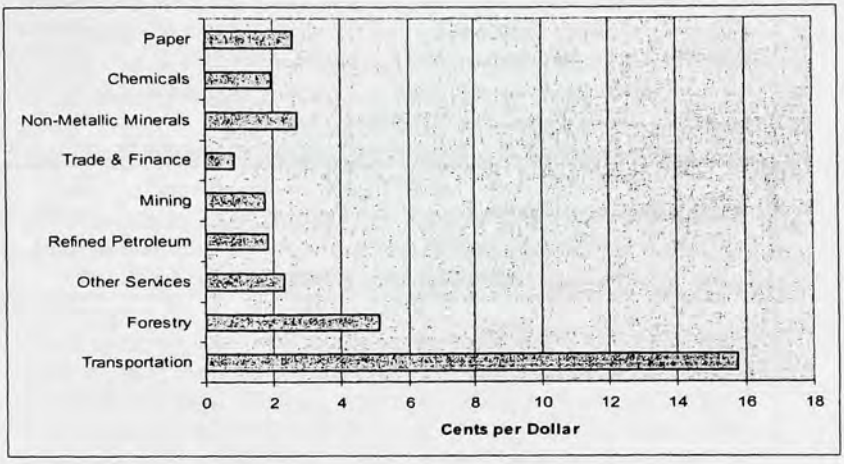


Figure 4 – Transportation Costs per Dollar of Total Cost, 1990

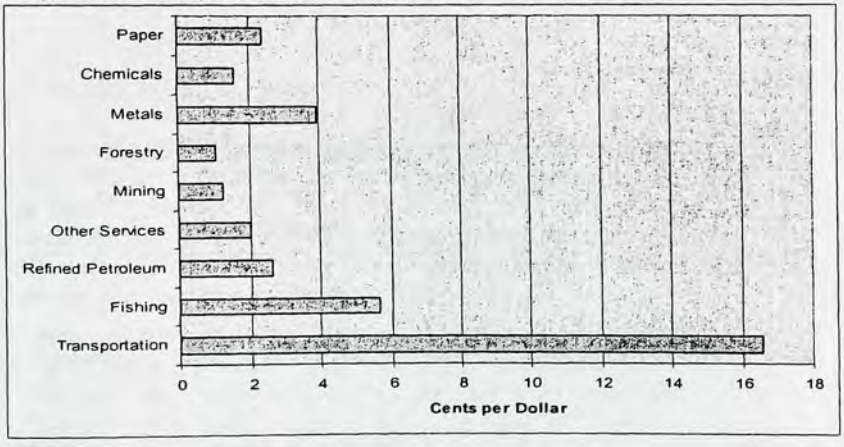


Figure 5 – Transportation Costs per Dollar of Total Cost, 1996

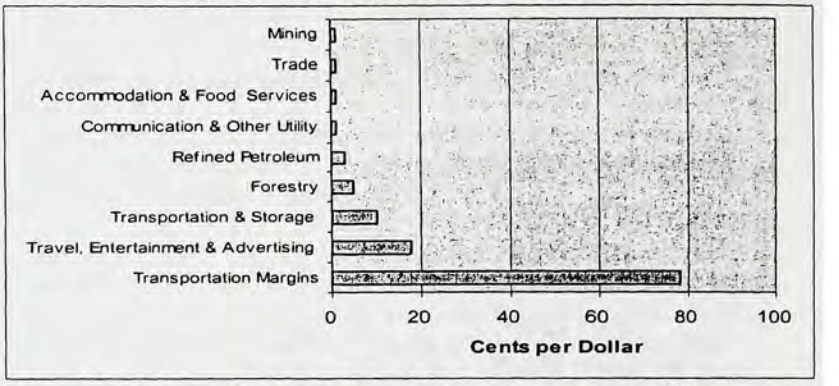
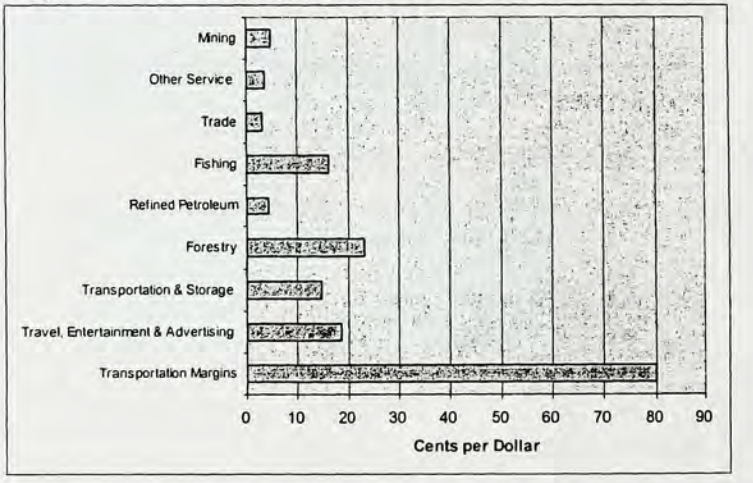


Figure 6 – Transportation Costs per Dollar of Total Cost, 1999



Inter-provincial and foreign exports show high transportation cost intensities that have varied slightly over time between 1979 and 1999. This is true, more so in the case of inter-provincial trade than for foreign exports. The high ratios of transport costs in the export

activity are typical of moving goods to distant markets. Nonetheless, they reduce the competitiveness of Ontario exports, particularly trade to other provinces. The transport cost intensities of foreign exports have increased over the years and those of inter-provincial trade have declined, as trade direction shifted from an east-west axis to a north-south axis. This is the outcome of the increased reach of Ontario's foreign exports to more distant markets, particularly to southern locations in the United States.

Inter-provincial trade is more expensive to transport than international exports, resulting in a shift towards a north-south trade pattern

While transport costs in total exports are rather high, the total Ontario exports of transport equipment to foreign markets remain exceptionally high. They accounted for about 40% of total Ontario foreign exports in 1979, and increased to about 44% in 1990. They account for more than 50% of total exports in the year 2000. If we were to add the exports to other provinces, these shares would rise to almost 60% of total Ontario exports.

Summary and Conclusions

Transportation activity absorbs vast quantities of the province's scarce resources, along with the time of its citizens. Because transportation is generally defined as a means to an end, the contributions of this sector to the creation of wealth, the export base and performance in the province and to the welfare of society at large are typically underestimated or undervalued.

The interplay between transportation and competitiveness is complex and significant. It is real, substantive and measurable. In this study we have made an attempt at defining the links to exports and to other critical sectors in the Ontario economy. We have also quantified these

links. In doing so we had to challenge some of the narrow view of this sector that, in our view, explains the undervaluation of these contributions.

A number of significant results emerge, a few are presented below:

- Transport cost intensities (the fraction of total cost of production represented by transportation) were the highest for the services sector, forestry, mining, refined petroleum, metals and the transportation activity itself until 1992. Travel and tourism sectors joined this group of sectors in the 1990s and assumed the highest transport cost intensities towards the end of the period.

- Sectors with high transport cost intensity are also typically high income contributors to the economy (show high value added coefficients) and are generally export oriented.

- Transportation costs can act as an impediment to exports. Indeed, exports show higher transport cost intensities than domestic production, particularly in the case of exports to other provinces.

References

Boyer, K.D. 1997. " American Trucking, NAFTA, and the Cost of Distance." *ANNALS, APSS,* 553, September.

Boyer, K and Stephen V. Burks, 2004. "Which One Was Driving? Productivity and Cost Trends in the Trucking Industry, 1977-1997", Michigan State University, third draft.

Bureau of Transportation Statistics, U.S. Department of Transportation, 2001. "North American Merchandise Trade by Truck with Commodity Detail, 2001."

http://www.bts.gov/ntda/tbscd/reports/annual01nat/na_com_2001trk2.html,

McCallum, J. 1995. National borders matter: Canada-U.S. regional trade patterns. *American Economic Review*. 85: 615-623.

Institute for Competitiveness & Prosperity, 2004."Realizing Our Prosperity Potential", *Third Annual Report*.

Sourcing of Transportation Inputs:

Comparative Tire Pricing on the Internet and Other Channels

Louis A. Le Blanc
Berry College

1.0 Introduction

The objective of this study is to identify the least cost sources for replacement of automotive and sport utility vehicle (SUV) tires. In the U.S., there is an average of 2.1 vehicles per household.[1] Tires are an expendable commodity, lasting an average of 25 months.[2] Most households face a tire replacement decision annually. For the tire business owner, this seems an attractive proposition, as tire prices average a margin of 24.81 percent.[3]

In the U.S., and likely in Canada, the types of vehicles vary widely. Tires for these vehicles differ extensively as well. For some makes, there can be over 100 choices.. A pure Internet tire seller, with only inventory and a website to manage, could potentially offer significant cost savings when compared to brick-and-mortar competition. However, tires are not easy or economical to ship. The average purchaser cannot install tires, so an Internet purchase will require a trip to an installer. Is it possible to buy a tire over the Internet, drive to an installer, pay labor charges, and still save money over the more traditional option of a complete purchase from a traditional tire dealer?

---

1 www.uscensus.gov
2 MTDealer.com
3 The Babcox Report, Tire dealer profile, September 2000.

2.0 Methodology

The following grid was used to research tire prices. A discussion of the selection of the grid components follows.

Tire Size	Vehicle	Brand	Note	Tire
235/75R1 5	Yukon / Tahoe	Firestone	Standard	Wilderness AT
235/75R1 5	Yukon / Tahoe	Dunlop	Standard	Rover AT
235175R15	Yukon / Tahoe	Michelin	Standard	XC LT4
265/70RI6	Yukon / Tahoe	Goodyear	Off Road Tire	Wrangler HP
265/70R1 6	Yukon / Tahoe	Goodyear	Off Road Tire	Wrangler APT
205/65R1 5	Taurus / Accord	Michelin	Stock Tire	MX4
205/65 RI 5	Taurus / Accord	Varies	Economy tire	Varies
225/55R16	Lexus GS430	Goodyear	Performance Tire	Eagle RS-A
225/55R1 6	Lexus GS430	Goodyear	Performance Tire	Eagle HP
225/55R1 6	Lexus GS430	Michelin	Performance Tire	Pilot XGT H4
225/55RI 6	Mustang GT	Michelin	Winter Tire	Artic Alpine
245/45ZRI 7	Audi A4 Sport	Michelin	Performance Tire	Pilot Sport
245/45ZRI7	Audi A4 Sport	Michelin	Performance Tire	Pilot Primacy
245/45ZR17	Audi A4 Sport	Pirelli	Performance Tire	P-Zero

2.1 The Consumer

Consumers do not enjoy buying tires. Consumers do not want to buy them, but they must. The majority of consumers purchase replacement tires the same day that they discover the need for new tires. Goodyear's internal research reveals that price is the most important consideration when selecting a tire retailer, with service being second. Brand is not very important. When comparing tires, consumers are most concerned with tread-life followed by wet traction.

Tire consumers can be broken down into four groups:

1. Price Constrained - Buys the best brand they can afford.

2. Value Oriented - Shops around for preferred brand at the best price.

3. Quality Buyer - Upscale consumers, research options, looking for the best tire for the application.

4. Commodity Buyer - Loyal, considers value, price and outlet.

Value-oriented and quality buyers purchase major brand tires, such as Goodyear, the majority of the time. Commodity buyers are the least likely to purchase major brands. This market has also proven to be very sensitive to promotions.[4]

For the purposes of this study, the Taurus owner looking for an economy tire represents the price constrained (or tire-indifferent) shopper. The performance tire buyer represents the quality buyer. Several of the choices are a consumer looking for a specific Michelin, Goodyear, or Firestone tire. These are the value-oriented buyers, looking for a specific brand at the best price. The commodity buyer is not represented here.

2.2 Tire Market

Tire purchasing has undergone significant changes over the past 40 years. In 1960, the local gas station was the primary site for a replacement tire purchase. Manufacturer owned centers were second, while independent dealers were the third most likely source for a replacement tire purchase.[5] In 2000, tire sales had shifted away from the service station and into the independent tire dealerships. Independent tire dealers sold 59.5 percent of all replacement tires sold. Mass merchandisers (Kmart and WalMart) were second at 16.5 percent, with tire company stores and warehouse clubs tied for a distant third at eight percent. Service stations had fallen to four percent.[6]

---

4 Bruce Isaacson, *Goodyear, The Aquatread Launch*, Harvard Business School, September 1994.

5 *Modern Tire Dealer*, August 2000.

6 www.mtdealer.com/stats.cfm

There are many varieties of tires available for the consumer to choose from. These can vary from winter tires to summer tires, high performance to touring, speed ratings, and so forth. For examples of the tire choices facing the buyer, refer to any website specializing in tires such as www.tirerack.com.

To narrow down the field, most tires searched were Michelin, Firestone, or Goodyear. These are the top three tire manufacturers combining for over 35 percent market share.[7] A Pirelli was chosen as an outlier to represent the ultra high performance tire. For specific models of tires, there were three approaches. For the Tahoe owner seeking a Firestone, an OEM (Original Equipment Manufacturer) replacement was selected. The same approach was chosen for the Taurus owner seeking a Michelin and the Lexus owner seeking a Goodyear. In all three instances, the owner was seeking to replace the tire on the car with same brand, size, and model of tire.

For light trucks, there are two primary types of tires. One is a standard duty light truck tire, as many of these vehicles never see any off road activity. A second choice is an off-road tire for the person who hunts or simply enjoys projecting an off-road image. Examples of both types were shopped.

A common tire need for the northern states is a winter tire. These highly specialized tires are made for exceptional traction in the snow, and are a common purchase for owners of rear wheel drive vehicles. The top selling example[8] is the Michelin Artic Alpine, and prices for that tire were sought for an appropriate vehicle.

---

7 Modern Tire Dealer, Annual Statistics, 2000.

8 *Car and Driver Magazine*, Winter Tire Comparison, November 2001.

2.3 Automotive Market

There are hundreds of makes and models of vehicles available in the United States. With every year, models receive updates and upgrades, often including suspension changes and wheel and tire upgrades. For 2002, new vehicle sales for the U.S. are presented in Table 1.

Table 1. Annual Vehicle Sales

Ranking	Vehicle	Sales
1	Ford F-Series Pickup	827,319
2	Chevrolet Silverado Pickup	650,523
3	Ford Explorer	385,034
4	Honda Accord	384,649
5	Toyota Camry	357,343
6	Ford Taurus	336,659
7	Dodge Ram Pickup	314,285
8	Honda Civic	310,120
9	Ford Ranger	256,488
10	Ford Focus	252,078

Light trucks and SUV's represented six of the top ten best sellers. A Yukon/Tahoe was chosen to represent all light trucks. A Taurus and Accord were chosen to represent the top four selling vehicles, in large part because they are equipped initially with the same size tires. (See Table 2.)

To represent a luxury/performance vehicle in a non-top 10 purchasing segment, a Lexus GS430 was chosen. The Lexus is the largest selling luxury marquee, and the GS 430 model offers a reasonable amount of performance. A standard tire would not be a primary choice on that car. Lastly, an Audi was added to the mix. For 2002, Audi offered an A4 model with a sport package which includes a 245/45ZR17. This tire is included here as it represents high performance brands, which may include Corvettes, Porches, etc. These are often the quality tire buyer.

Table 2. Tire Sizes, Applications, Brand and Prices

Tire Size	Vehicle	Brand	Application	Tire
235/75R15	Yukon/Tahoe	Firestone	Standard	Wilderness AT
235/75R15	Yukon/Tahoe	Dunlop	Standard	Rover AT
235/75R15	Yukon/Tahoe	Michelin	Standard	XC LT4
265/70R16	Yukon/Tahoe	Goodyear	Off Road	Wrangler HP
265/70R16	Yukon/Tahoe	Goodyear	Off Road	Wrangler APT
205/65R15	Taurus/Accord	Michelin	Stock	MX4
205/65R15	Taurus/Accord	Varies	Economy	Varies
225/55R16	Lexus GS430	Goodyear	Performance	Eagle RS-A
225/55R16	Lexus GS430	Goodyear	Performance	Eagle HP
225/55R16	Lexus GS430	Michelin	Performance	Pilot XGT H4
225/55R16	Mustang GT	Michelin	Winter	Artic Alpine
245/45ZR17	Audi A4Sport	Michelin	Performance	Pilot Sport
245/45ZR17	Audi A4Sport	Michelin	Performance	Pilot Primacy
245/45ZR17	Audi A4Sport	Pirelli	Performance	P-Zero

3.0 Vendor Choices

Vendor choices were chosen to represent the choices facing the tire purchaser in a small community, such as Rome, Georgia. Many tire purchases are completed through independent tire dealers, both small and large. Most of the dealers selected offer some level of service on the Internet, even if only for price comparison.

3.1 Bert Brooks Tire

This Rome (GA) retailer represents the local independent retailer. The outlet carries all of the most popular brands, and commands reasonable loyalty over the local populace. They have no website. Brooks Tire sells Michelin, Pirelli and Goodyear brands.

3.2 WalMart

Wal-Mart represents the mass merchandiser sales outlet. This type has been losing market share over the past five years, primarily to independent tire dealerships. Wal-Mart offers tire choices over the Internet, but this is for price comparison only. All purchases must be made at the local WalMart and Internet prices were consistent with in-store pricing.

3.3 Sears

Sears has seen their tire sale market share erode over the past decades. In 1975, the Sears brand tire represented ten percent of all replacement tire sales. The Sears brand tire now represents only 3.75 percent of all light truck and automotive replacement tires.[9] Sears remains a popular choice for tire sales, and they can order virtually any tire manufactured. Tires can also be ordered through the Internet at Sears.com and delivered to a local Sears store for installation. While Internet prices were not always the same as the local store prices, Sears guarantees that you will receive the lower of the two prices. In most cases this is the store's quoted price.

3.4 Tirerack.com

Tirerack is an Internet based tire store (www.tirerack.com). Starting as a bricks-and-mortar store in Indianapolis in 1979, the dealership took advantage of Internet technology early to become the leading Internet tire merchandiser. They primarily cater to the performance enthusiast, advertising in *Motor Trend*, *Car and Driver*, and other automotive publications. Like many e-business firms, Tirerack.com has an advantage in inventory and availability with four warehouses in the United States. They have partnered with installers in most major metropolitan areas, and can ship directly there to ease the difficulty of finding a garage to install tires purchased over the Internet.

---

[9] Modern Tire Dealer, Annual Statistics, 2002.

3.5 Etire.com

Etire (www.etire.com) is very similar to Tirerack.com, with the exception that there is no home bricks and mortar location. Etire was started as an Internet based tire sales sight, with seven distribution centers located throughout the United States. Like Tirerack.com, they have partnered with local installers to improve convenience for the buyer.

4.0 Observations

The replacement tire market is rapidly changing and ever evolving. The major manufacturers change tires model as quickly as do the automobile manufacturers. The average life of new tire technology is 12 to 18 months.

A second observation is that most dealers can order any of the major brands. Sears carried virtually every major brand, and could get any tire within three business days. WalMart offered the same service, but did not have the same delivery period; and, prices for non-stocked tires were not as advantageous as their advertised brands of tires. The choices available are substantial. For the owner of either a Taurus or Accord seeking a tire, Tirerack.com offered 102 different tires with a wide price range. The sport utility owner faces a similar array of offerings from Etire.com, with 122 tires available to fit the stock rim. (See Table 3.)

Table 3. Tires and Prices by Vendor

Tire Size	Tire	Bert Brooks	Wal Mart	Sears	T-rack.com	Etire.com
235/75R15	Wilderness AT	$85	$87	$118	$98	$101
235/75R15	Rover AT	$90		$107	$103	$83
235/75R15	XC LT4	$112	$123	$124		$126
265/70R16	Wrangler HP	$145		$148	$116	$108
265/70R16	Wrangler APT	$145		$151	$102	
205/65R15	MX4	$110	$95	$171	$149	$116
205/65R15	Varies	$55	$57	$75	$68	$61
225/55R16	Eagle RS-A	$174		$236	$172	$144
225/55R16	Eagle HP		$119	$197		$141
225/55R16	Pilot XGT H4	$185		$169		$139
225/55R16	Artic Alpine			$140	$149	
245/45ZR17	Pilot Sport	$189	$155	$262	$205	$184
245/45ZR17	Pilot Primacy				$205	
245/45ZR17	P-Zero			$414	$203	

For the prospective purchaser wanting to buy the best tire available, both Etire.com and Tirerack.com took advantage of their wealth of data to offer the buyer a tire comparison. Different criteria for tire performance such as tread life, wet traction, road noise, handling, and more were evaluated by past owners. This would especially appeal to the performance tire buyer and the quality buyer.

It is a clear advantage of the online retailers to be able to collect and distribute such data. Not surprisingly, the best tire per the feedback was not always the most expensive tire. Tire manufacturer and independent data was also available for comparison should the prospective buyer not wish to trust the metrics of his fellow consumer. Sears and WalMart, both with websites, did not offer the same comparison service.

Tire installation prices were relatively equal. In many instances, this is a profit center for a tire dealer. Mounting and balancing a tire is a five-minute task, at most, for an experienced technician.

The average charge for this task is $8.29 for a passenger tire, and $9.04 for an SUV tire.[10] Performance tires can be more, with charges averaging $9.64 per tire nationally, and a high of $25 per tire at an Atlanta-area European car service center. At $9.00 per tire, a good technician can balance and install eight tires per hour, including time to remove the wheel. That represents $63 per hour, more than enough to cover overhead, salaries, and expenses.

In the outlets surveyed, installation prices were almost equal, ranging from $8.00 to $10.00 per tire for basic service. In most cases, special applications (high performance, or large rim sizes) were $2.00 extra per tire. These typically require better equipment and more time to install. Virtually all dealers charged $1.50 per tire disposal fee, so that factor was not considered in pricing information and comparisons.

4.1 Statistics

Prices were collected for all available tires at each of the outlet choices. This was not an easy task, as not all tires were available at all outlets. Only the Internet and Sears offered nearly all choices. In the case of Sears, many of the tires were not in stock locally, and a wait for the tire would be needed (similar to the Internet). There are many variations available of a single tire, and head-to-head comparisons were not easily done. (See Table 3.)

---

[10] The Babcox Report, www.babcox.com/editorial/tr/tr9002.htm.

Table 4 summarizes the ten significance tests (paired t-tests with equal variances) between five tire sellers. The table values indicate whether the sample means are significantly different from zero at the five percent level of significance or greater. The table value of "Yes(+)" indicates that the row entry (Bert Brooks at row 1) is significantly less than the column heading (Sears at the top of column 3). For example, WalMart's average prices were significantly less than the comparable average prices for tires from Sears, Tirerack.com and Etire.com, while Sears average prices were significantly higher than those at Tirerack.com and Etire.com.

Table 4. Significant Differences

	Bert Brooks	WalMart	Sears	T-rack.com	Etire.com
Bert Brooks	-	No	Yes (-)	No	No
WalMart		-	Yes (-)	Yes (-)	Yes (-)
Sears			-	Yes (+)	Yes (+)
Tirerack.com				-	Yes (+)
Etire.com					-

Table 5 provides the details of the paried t-test data comparing WalMart with Etire.com. On the six tires compared, WalMart had lower prices on each tire. The difference between the average values for the two retailers was $15.39 and was statistically significantly at the one percent level. (Space constraints in the proceedings do not allow reporting of the details of the ten statistical comparisons.)

4.2 Items Not Quantified

Several factors were not weighed in to the tire purchasing decision, but clearly have some value to consumers. The first was availability of comparison data. This will matter to the quality and high performance buyer. The Internet sites held an edge here, although most retailers were willing to offer advice on the best tire

Table 5. WalMart versus Etire.com

Tire Size	Brand	Tire	WalMart	Etire.com
235/75R15	Firestone	Wilderness AT	$ 87.51	$101.48
235/75R15	Michelin	XC LT4	$123.89	$126.48
205/65R15	Michelin	MX4	$ 95.00	$116.48
205/65R15	Varies	Varies	$ 57.55	$ 61.48
225/55R16	Goodyear	Eagle HP	$119.61	$141.48
245/45ZR17	Michelin	Pilot Sport	$155.99	$184.48
Average			$106.59	$121.98
Minimum			$ 57.55	$ 61.48
Maximum			$155.99	$184.48
Difference			$ 15.39	
Degrees of Freedom			5	
T Statistic			-3.601	
P (T<=t) one-tail			0.008	

A second factor was ease of installation. Clearly, the Internet purchase requires an extra step versus the local tire dealer purchase. The Internet sites have partnered with installers, but none were in the Rome area. Convenience factors will outweigh the comparison factors and cost difference advantage on Internet purchases for most buyers.

5.0 Summary and Conclusions

The tire market alternatives are as varied as the tire buyer. There are many choices in brands, tires (with many different performance characteristics), and purchase methods. The best choice must take into account the consumer's needs and purchasing preferences.

5.1 Price-Constrained

This consumer buys the best brand they can afford, or in many cases is simply looking for the cheapest tire for their car. This consumer was represented by the owner of a Taurus or Accord looking for the cheapest tire available. This buyer will probably look for the best deal. The added hassles of finding an installer and going on line will not appeal to this buyer. They are encouraged to shop locally, or simply go to a local dealer and get the lowest price tire available. In this category, WalMart and Bert Brooks offered the best prices.

5.2 Value-Oriented

This consumer shops around for preferred brand at the best price. For the consumer who simply must have a Michelin, Goodyear, Firestone, or other loyalty brand, they are encouraged to shop the Internet. Individual preferences such as convenience and the experience at the retailer may also weigh into the decision. In this category, WalMart and Bert Brooks also offered the best prices.

5.3 Quality buyers

These are generally upscale consumers who research options, looking for the best tire for the application. Often, these are also high performance tire buyers. The Internet is the clear choice for finding the best tire. This buyer may also find the best price on the Internet. Tirerack.com bills itself as the performance specialists of the tire market. In pricing, selection, and data availability, that is true for the performance tire market. Bert Brooks was very competitive with Etire.com and Tire-rack.com.

5.4 Commodity Buyers

These buyers are loyal, and place a high value on price and outlet. They are not prone to shop around. They will drive to the outlet they have been dealing with for years, and walk away not knowing what they are missing. There is a factor in the simplicity of the purchase which, for them, offsets any foregone savings opportunity. Assuming they are not seeking a performance tire, they are not missing much.

SHORT SEA SHIPPING and the SUPPLY CHAIN

A Review of Cross Lake Ferry Economics and Benefits

40th Annual Conference
Canadian Transportation Research Forum
Hamilton Ontario
May 8-11, 2005

Christopher Wright
President
The Mariport Group Ltd.

EXECUTIVE SUMMARY

The paper reviews the potential role of short sea shipping in US/Canada cross border trade. It demonstrates that marine transportation has a viable role in enhancing supply chain reliability and reducing border congestion.

Reference is made to European experience as well as the problems facing North America in responding to rapidly expanding transportation needs, particularly relative to trucking. The paper reviews US/Canada trade, concentrating on the three primary lower lake gateways. Three proposed ferry operations on Lake Erie and Lake Ontario are reviewed, together with their traffic potential, catchment areas and environmental benefits.

The paper shows that the ferry routes have the ability to significantly reduce border congestion, by offering new gateways between the US and Canada. However, despite attractive cost margins between direct trucking and use of the ferry, institutional barriers, that discriminate against the marine mode, could well prevent them being introduced.

RESUME

Christopher Wright is the President of The Mariport Group Ltd, a specialized marine and port consulting company based in Cambridge, Ontario. He commenced his career in shipping with the Westinform Service in London (UK) in 1964, and after a period with the Chamber of Shipping of the United Kingdom moved to the Bahamas where he was Senior Consultant and a Director of JBC in Nassau. A move to Canada in 1981 created new opportunities in passenger shipping and ferry consultancy, both in North America and internationally.

He is a member of the Chartered Institute of Logistics and Transport; The Society of Naval Architects and Marine Engineers and has been associated with IMTA/Interferry since 1986.

He has been involved in ferry projects from Shetland (UK) to South Korea, and is currently working on five ferry studies in North America, two of which are Short Sea freight ferries.

SHORT SEA SHIPPING AND THE SUPPLY CHAIN

Since Just-in-Time was pioneered by Taiichi Ono for Toyota, in the early 1970's, manufacturing of not just automobiles has been revolutionized. An example of the complexity of the input side of a vehicle assemble operation was given recently for GM Brazil[1]. 600 parts suppliers, 350 in Europe, 150 in North America and the balance in Brazil. As will be appreciated, transportation and its reliability play a critical role in the success of this operation.

In Europe, and in North America, the same approach to outsourcing of parts and components has required a finely tuned distribution system, most commonly based on trucking. However, this is not to imply that marine, and other modes, are not involved. The development of the international container industry, starting in the mid-1960's, enabled manufacturers to place increasing dependence on international outsourcing of products, trading on the inherent

[1] Containerization International, March 2005.

reliability of the container lines[2]. Reliability of transportation, as the glue that holds supply chains together, seems to be little appreciated. Perhaps this is because transportation has been rolled into "logistic", and then outsourced. Even a major study by Cranfield University in the UK for the Dept. for Transportation into resilience of supply chains, gives scant recognition of the critical importance of physical transportation[3]. This report does provide one apposite comment that relates to the distribution problems we face today:

"There is no common language other than cost".

This bottom line driven approach to every link in the supply chain, rather than tempering cost with reliability, is the root cause of transport capacity issues around the world, extending from port and intermodal rail capacity to road and border congestion.

Road congestion is not just a problem in North America, it is creating growing problems in Europe and is driving a thorough reappraisal of goods movements. The recently introduced LKW Maut road toll system in Germany (although 18 months late) is part stick to force the private sector to re-evaluate routing and part revenue generation to fund modal growth. There are plans to introduce a pan-European toll for vehicles over 3.5 tonnes by July 2008 and for all vehicles by July 2011[4].

An example of the impact of road congestion in Europe is that in the UK alone it is estimated to cost industry £20bn per year[5]. Cross border congestion between Canada and the USA is estimated at

---

[2] The severe winter of 1993/94 created considerable schedule disruption in the Europe-Montreal container trade resulting in a number of manufacturers having to fly components into North America to keep production going.

[3] Creating Resilient Supply Chains – A Practical Guide. 2003

[4] Comment by dr. Wolfgang Reinhardt, Director Regulatory Affairs for the European Automotive Association at the UK Freight Transport Assoc. "Freight Summit 2005". London UK, Feb 2005.

[5] Bo Lerenius, CEO, Associated British Ports at the 2nd National Sea & Water Conference. Hull UK, March 2005.

U$4.01bn per year[6]. Another study by the Canada-US-Ontario-Michigan Border Transportation Partnership has not estimated a current US-Canada border congestion cost, but has estimated that costs for the Detroit/Windsor region alone will reach U$7.5-8bn by 2020.

While there is recognition in both Europe and North America that congestion is costly, Europe is actively working on multimodal solutions within both the public and private sector. In North America there is no effective leadership on the issue in either public or private sectors and the marine option is effectively ignored. The Border Partnership study did not explicitly discuss the role of the barge ferry operating at the foot of the Ambassador Bridge, and only mentioned marine options in passing.

Canada-US Trade Flows

Canada's trade with the US and v.v. is considerable and the US Dept. of Commerce reported 2004 value at U$446bn, up from U$392bn in 2003. This was about the same level of activity as US trade with the EU-15. Data from Stats Canada is somewhat lower as the following table shows. Note that in looking at trade data in value terms it is necessary to take into consideration the changes that have occurred in the value of the C$ relative to the U$; the trends are somewhat different in U$ vs Canadian dollars.

US/CANADA CROSS BORDER TRADE
Five Year Trend $bn

	2000	2001	2002	2003	2004
US$ value	397	368	359	378	428
C$ value	589	570	563	531	557
Implied exchange rate	.67	.65	.64	.71	.77

[6] The US-Canada Border: Border Economic Costs, their Causes and Alternative Management Strategies. John C. Taylor & Assoc. Nov 4, 2003. Delays and transit times only, excludes an estimated $6.28bn for trade policy related costs.

Modally[7], in 2000, trucks moved about 66% by value and 41% by weight; rail 24% by weight; water 13% by weight. However, weight data is imputed from value. Actual weight data is not readily available from any consistent source for truck or rail cross border movements and what is available is scattered across several sources[8].

For 2001 it was estimated[9] that rail freight carried 54.8m tonnes of cargo cross border, of which 54% by weight and 71% by value moved through the three border crossings of Buffalo/Niagara Falls, Detroit/Windsor, and Port Huron/Sarnia. Tonnage estimates for truck freight in 1999, based on a Canadian National Roadside Study, indicated 145.3m tonnes of cross border cargo[10]. Of this, 56% by weight and 68.6% by value moved through the three primary Ontario border crossings. This demonstrates the extraordinary concentration of Canada/US trade around the lower Great Lakes of Erie and Ontario.

Projections for rail freight volumes indicate a 30% increase in both volume and value by 2020[11]. The forecast for truck freight indicates a 60% growth in bi-directional transit numbers to 2020[12]. A projection (below) of traffic through the three primary Ontario gateways suggests a somewhat lower, but still high growth.

---

[7] US Dept. of Transport, Bureau of Transportation Statistics – North American Trade & Travel Trends.

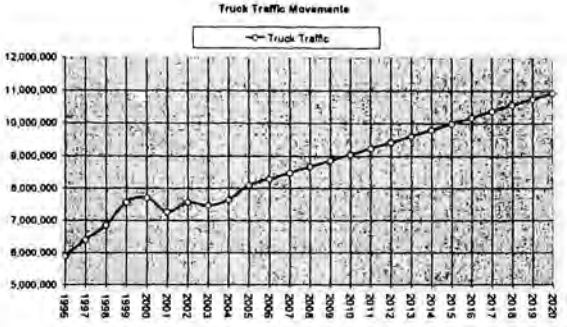
[8] Comments in two studies for the Eastern Border Transportation Coalition.

[9] Cambridge Systematics EBTC study of rail freight crossing the Canada-US border. Nov 2004.

[10] Parson Brinkerkoff Quade & Douglas Inc. Truck freight crossing the Canada-US border. Eastern Border Transportation Coalition. Sept 2002.

[11] A generic forecast in reference 9 from US sources.

[12] Reference 10 does not provide actual weight or value forecasts. Numbers of two-way moves are predicted from regression analysis of long term trends at each crossing. Time series vary.



Growth in road freight volumes is not something that is unique to Ontario and the US border crossings. In the USA, commentators have pointed to growing congestion and the inability to invest in new transportation infrastructure. For example, vehicle miles grew 80% between 1990 & 2000, while highway lane miles grew only 2%[13]. In Europe, goods transport volumes through Germany, which lies at the centre of EU traffic, is projected to rise 90% between 1998 & 2015[14].

Short Sea Shipping as a Solution to Congestion

The European Union, in recognition of the true nature of short sea trade has, since 2001, used the term "motorways of the sea" as the type of activity they wish to develop. Essentially this is to try and control growth of truck traffic. The three main objectives for the Motorways of the Sea initiative are:

- Reducing road congestion through modal shift
- Increase cohesion (within the EU)
- Freight flow concentration on sea-based logistical routes.

[13] Journal of Commerce. July 10-25, 2004.

[14] European Short Sea Network position paper 24 Sept 2004

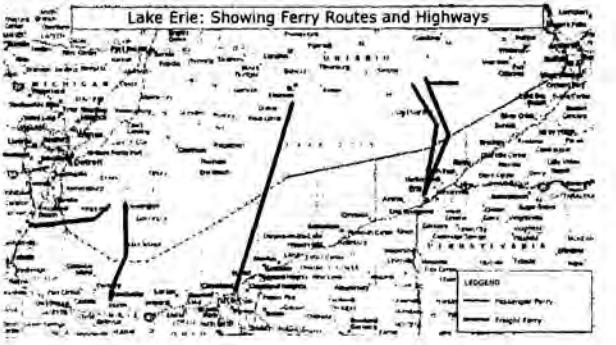
This goal is exactly what is being targeted in North America. As in Europe, there are opportunities for at least four types of long and short haul ferry operations.

- A true intermodal system built around a 53' container, similar in concept to Geest NSL's 13.5m (45') container operation
- Drop trailer freight ferries, probably where runs are over about eight hours
- Full roll-on, roll-off freight ferries on shorter runs
- Ro-pax ferries, although these are possibly less likely, given the significant differences between North American and European passenger traffic.

Proposals for the Great Lakes, in the core Ontario area, cover each of the three ferry types noted above:

- A ro-pax ferry between Cleveland (OH) and Port Stanley (ON)
- A roll on/roll off ferry between Erie (PA) and Nanticoke (ON)
- A drop trailer ferry between Hamilton (ON) and Oswego (NY)

While other ferry routes are possible it is these that will be discussed and the Hamilton-Oswego ferry reviewed in detail. The Cleveland and Erie ferries would both operate across Lake Erie. The Oswego ferry would operate on Lake Ontario.



CLEVELAND-PORT STANLEY

Distance	86sm
Transit time	3hr 30min
Number of vessels	1
Season:	
Closed	Feb 1-Mar15
Shoulder	1 round trip/day Mar 16-May 31, Oct 1-Jan 31
Peak	2 round trips/day Jun 1-Sep 30
Vessel	TBD
Speed	18kts
Capacity	3,600 lane feet, 1,200 pax
Off interstate driving	19sm
Period of study	May 2003-May 2004.

The analysis, for truck/commercial interest, was geared towards time saving primarily between the immediate Cleveland hinterland and the London/St. Thomas area in Ontario. The estimates of activity were derived from broader statistical data using employment to proportion probable goods traffic in the primary market. Overall market projections for the service were in the range 20-25,000 truck moves per annum.

Author's Comments

The traffic potential for the route may well be understated because it concentrated on the immediate hinterland of the ferry termini and did not, specifically, look at cost savings for long haul truck operations. Also, as this is targeted at a roll-on, roll-off type of operation, there are potential benefits for the truck driver being off duty. As the primary freight mode is tractor trailer traffic, space could well be at a premium during summer months, despite two round trips/day, due to competition with passenger-related vehicles. Also, shutting down the ferry for a part of the season will not encourage users to integrate the operation into their cross border strategies.

The primary truck route on the US side of the Detroit/Windsor crossing is I-75/I-71/I-65, with activity levels of about 20,000 trucks

per week to/from Ontario[15]. The ferry could divert a portion of this traffic at Cincinnati via Cleveland saving 172 miles driving. At current long haul costs, fuel surcharges and border costs, this is about $400/tractor trailer before ferry charges. In addition to this traffic about 5,000 vehicles per week each through Port Huron on I-75/I-94 and on I-71/I-90 through Niagara would also be candidate for ferry use

ERIE-NANTICOKE

Distance	48sm
Transit time	3hrs 30min
Number of vessels	1
Round trips/day	2
Season	year round apart from scheduled downtime
Vessel	TBA
Speed	16kts
Capacity	5,900 lane feet, 0 pax
Off interstate	54sm heading west in Ontario
	38sm heading east in Ontario
Period of study	2003.

The route is proposed as a commercial ro-ro service that will complement a separate high speed passenger vehicle ferry between Erie (PA) and Port Dover (ON). The analysis looked at shipping routes and, through a modelling process, predicted a percentage of traffic that could be captured by the ferry. In 2005 the potential volume of business was estimated at just under 100,000truck moves/year. Survey work associated with the analysis indicated a high level of interest in the service for both ro-ro and drop trailer service.

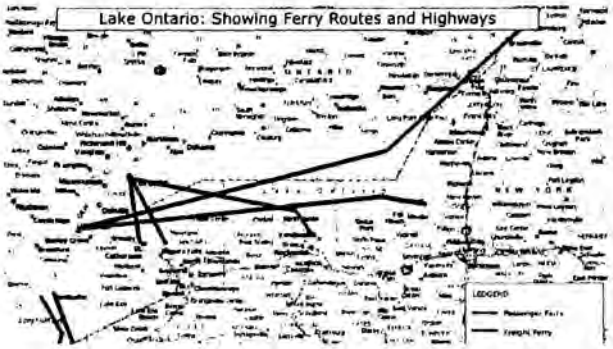
Author's Comments

The traffic potential for the route is probably of the correct order of magnitude as it incorporated both local and long distance movements. However, specific cost savings based on comparative route miles and border costs were not assessed, although it was noted that there would

[15] Based on observation of truck flows in Reference 10

be driver benefits for the time on the ferry. The analysis did not appear to be based on actual candidate truck moves. One of the potential problem areas may be the off highway driving needed on the Ontario side, almost all of which is on tertiary roads.

Volumes on I-90/I-79 through the Niagara border crossing are in the region of 10,000trucks/week to/from Pittsburgh, plus a further 5-10,000 moves per week into and through Cleveland. This latter traffic would be competition for the Cleveland ferry, and Erie Nanticoke could be preferred, given that it is a year round operation , with two sailings per day in each direction. Driven distance savings are not as great as for the Cleveland route, at 110sm. At current short haul truck pricing plus border costs, the savings would be in the order of $400/tractor trailer before ferry charges.



OSWEGO-HAMILTON

Distance	166sm
Transit time	12hrs
Number of vessels	2
Round trips/day	.5
Speed	16kts
Capacity	3,200 lane feet, 0 pax
Off interstate driving	11sm
Period of study	2000 & 2004.

The analysis, given the length of the routing, was targeted entirely at drop trailers and focused on comparative truck routings and costs either direct or via the ferry. This enabled a margin, before ferry charges, to be established based on current truck costs, border crossing costs and surcharges. An analysis of origin/destination pairs that met a minimum margin criteria was then used to help determine the catchment area, and potential traffic level. This showed a potential market of some 200-260,000 moves/year, at 1999/2000 levels, which was 20% of all movements at the Niagara border. Ferry capacity, with two units, was 15% of market or 3 % of all traffic.

TYPICAL TRUCK MARGINS U$[16]
Excludes Border Surcharge & Bridge Fees

	Boston	N.York	Philadelphia	Norfolk	Syracuse
Hamilton	$452	$499	$415	$175	$424
Kitchener	$452	$499	$399	- nc -	$424
N. Bay	$156	$232	$232	- nc -	- nc -

Author's Comments
The traffic potential is probably realistic, but the potential for take up by shippers has yet to be tested. Interest may be higher today than when the study was undertaken as trucking costs have risen considerably and margins are higher. Service reliability was deemed critical, with two vessels offering nightly sailings in each direction. Anecdotal information suggests considerable interest, with added savings to shippers of being able to use a day dray service on either end of the route.

Comparison of Service Factors
An assessment of service factors included with the Cleveland and Erie ferry studies showed a high ranking for Price, Reliability and reduced Customs delays. The same factors are likely to be the key determinants for the Oswego/Hamilton route as well.

---

[16] Based on industry advice as of Oct 2004 of short haul rates of U$2.37/mile, plus $.27/mile fuel surcharge. Long haul rates are about U$1.34/mile with the same fuel surcharge. A recent check has shown that these rates are still valid in 1Q 2005.

These factors were certainly critical in attempts over the last few years to establish a truck ferry between France and Italy. The original proposal of alternate day service between Fos and Savona was changed to daily, but still failed because of high port costs and a route that did not offer a time saving[17]. A service has now been established between Toulon and Citaveccia, which offers a 14-hour sailing time versus 22 hours driving, although again it is only alternate days. However, traffic is growing and a second vessel is expected shortly[18].

Role of the Three Cross Lake Routes

The analysis of potential catchment area for the Oswego-Hamilton route showed the ferry is likely to have a very specific area of application.



Although no comparable analysis is available for the other routes, the brief analyses in this paper, of US routes to the Canadian border and distances saved, strongly suggests that each ferry will have a distinct market, with little crossover.

[17] Lloyd's List. 28 April, 2004.
[18] Lloyd's List. 07 Feb 2005.

What the ferries offer is reduced cost and enhanced reliability relative to border crossings. By providing a comprehensive documentation of en route cargo with advanced notice, clearance by Customs in both Canada and the USA can be expedited. This could be of particular value where a steady stream of goods needs to be moved on a particular route, in that the ferry becomes an integral component of the calculation of time and goods in transit, with border uncertainty removed. This does presuppose a year round reliable marine service.

Environmental Benefits

The short sea ferries proposed for the lower Great Lakes will only be successful if they reduce driving distance and/or time. The margins for the Oswego-Hamilton route were in fact calculated based on trucking costs and distance differentials. Typical savings/trip in terms of driving were estimated at 150 miles, thus at full capacity the two ferries would remove 10.5 million miles of driving in a 350-day operating year. Assuming a one-hour border delay with the truck engine idling, this is equivalent to 13.65m miles of avoided truck driving. The Cleveland and Erie crossings had estimated savings of 172 and 110 miles respectively, thus each crossing could contribute similar savings in terms of reduced emissions, incidents, pavement wear and tyre disposal. The crossings would also mitigate the growing shortage of drivers for cross border activity.

The marine mode is well recognized as the most energy efficient, least polluting and safest transportation mode, as the following table shows. Thus moving cargo to the marine mode should enhance quality of life in urban areas, as well as helping reduce greenhouse gases and atmospheric pollution.

There are other, un-quantified benefits related to the congestion impact of accidents, particularly truck accidents on busy highways. Even a minor accident can close lanes or complete roads for 4-12 hours and in congested urban areas, the time cost of delays and atmospheric emission impact can be significant.

COMPARATIVE ENVIRONMENTAL INDEX OF TRANSPORTATION MODE[19]

	Marine	Rail	Road
Energy efficiency	1	2.2	9.7
Air emissions	1	1.4	7.6
Accidents	1	13.7	74.7
Spills	1	10.0	37.5
Noise	1	1.4	1.3

Institutional Issues

There are a number of factors that materially impact cost and feasibility of short sea ferries on the Great Lakes.

Pilotage

The Great Lakes are a bi-national region; there are no international waters and the US Offshore Petroleum Act of 1990 has created a Pilotage cost nightmare for ships that are neither US nor Canadian flag. A foreign flag operation of the services being considered would incur pilotage costs of a least U$1,500/ship trip.

Putting the ship under Canadian flag is not necessarily a viable option as a non-NAFTA vessel incurs 25% duty plus full upgrades to current MARPOL and SOLAS. The alternative is US flag, which does not incur a duty provision, but does involve upgrade work. However, USCG will consider reasonable alternative and waivers where it can be demonstrated that ship safety will not be compromised.

Customs & Immigration

US and Canadian authorities take very different approaches. On the US side, service is provided without charge, but the operator must meet all the technical requirements as laid out by Customs and Immigration. The main problem is that getting the service in place can take 12-18 months, even with political assistance. In Canada, service will be provided promptly and there is some flexibility with

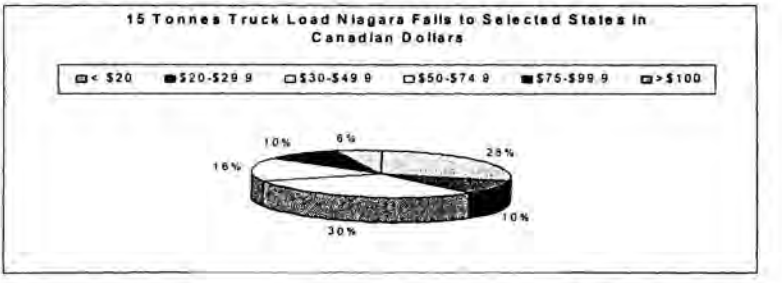
[19] Hamelin-CFORT- Gesco 11.2000 on behalf of SODES. Comparison based on Tonne Kilometres

regard to interim facilities. However, it has to be paid for. Canada has grandfathered existing border crossings and airports as well as some ports of entry for free service. For new operations, or enhanced activities at ports that have resident customs facilities, Canada is in full cost recovery mode. This can add significantly to operating expenses. The Windsor/Detroit ferry, for example, has to pay about $50,000/year for a single customs officer 08:00-16:00, five days a week. This is more than a little limiting.

Habor Maintenance Fee

The HMF is value based at .125% on all imports into US ports (the export portion was struck down by the Supreme Court some years ago). The fee applies equally in domestic as in international trade and is supposed to support dredging programmes in the USA. There are some exclusions, and ferries that carry some truck traffic are excluded where "freight is a minor portion of the business".

Until there is a broad exclusion for freight ferries (i.e. cargo on wheels) it will be difficult to convince shippers that a ferry route which is taxed is better than the traditional all-road route, which is not. An analysis of the Hamilton/Oswego route done using Stats Canada data, see figure below, suggested that 68% of all moves that would be taxed, could be contained at under about U$40/load. There are also a large number of international freight containers through the US to/from Ontario which are exempted from the fee, as well as a significant number of empty trailers which would be zero rated. Thus the overall impact of accommodating the HMF might be manageable in the short term on this route, but it would certainly impact traffic levels and restrict growth.



Transportation Implications of Canada's Aging Population

Darren M. Scott, K. Bruce Newbold, Antonio Páez, Pavlos S. Kanaroglou

Center for Spatial Analysis, School of Geography and Geology, McMaster University, Hamilton, Ontario, L8S 4K1

1. INTRODUCTION

Canada's population is aging rapidly. In fact, by 2026, approximately 21% of Canadians will be at least 65 years old – an increase of almost 9% from 2001 (Statistics Canada, 2005). In absolute terms, the elderly population is expected to grow from 3.935 million to 7.753 million persons over the 25-year period (Statistics Canada, 2005). As noted by Newbold et al. (2005), this unprecedented demographic change is being driven by three factors – notably, the aging of the baby-boom generation, greater life expectancies and historically low fertility rates. At the same time, Canada's older population is healthier, more affluent and more mobile than earlier generations (Chen and Millar, 2000; Manual and Schultz, 2001), leading to potentially greater demand for out-of-home discretionary activities and travel.

While Canadian policy makers have long recognized the importance of an aging society with respect to Canada's health-care system and national pension plan, transportation-related impacts have received little attention. By comparison, other countries such as the United States appear to be more proactive. In fact, the Transportation Research Board identified "an aging population" as one of 14 critical issues or challenges facing users or providers of transport services today and in the near future (Pisarski, 2003; TRB, 2001).

One of the most significant impacts of aging populations on transportation is the increasing "automobility" of the elderly or the

combination of greater and different trip making supported by the car. Recent research from the United States, Australia and some European nations suggests that over time older people (i.e., age 65+) have become increasingly dependent on the automobile to meet their travel needs (Rosenbloom, 2001). Today, a higher proportion of older people in these countries are drivers than just a few years ago. Relative to earlier generations, they tend to take more trips, a greater variety of trips and longer trips, reflecting increased incomes, better health and more active lifestyles (Rosenbloom, 2001).

Concerns have been raised as to the potential impacts of this increasing automobility of the elderly on traffic congestion, emissions of air pollutants and, more generally, the urban environment (see, for example, Carp, 1988; OECD, 2001). Others have also flagged the issue of road safety (Dulisse, 1997; OECD, 2001; TRB, 1994). It is commonly suggested that elderly drivers are relatively more hazardous despite experiencing fewer accidents per driver because they have much higher accident rates when rates are adjusted for mileage (Burkhardt and McGavock, 1999; Chu, 1994; Hildebrand, 2003; Stamatiadis and Deacon, 1998). Other impacts on transportation include the potential demand for increased public transit services to meet the travel needs of the elderly who are no longer able to drive.

To address the transportation-related impacts of an aging population with appropriate public policies, Canadian policy makers must first understand how the travel behavior of Canada's older population differs from that of other age cohorts and has changed over time. More importantly, policy makers must have access to spatially explicit predictive tools to assess how the convergence of an aging population and its increased automobility will impact Canada's urban transportation systems and concomitant environmental systems. GEOIDE[1] project DEC#SCO is a direct response to these challenges.

The purpose of this paper is to highlight, albeit briefly, the four objectives of the project. Detailed information on specific objectives

[1] GEOIDE, or Geomatics for Informed Decisions, is one of 21 Networks of Centers of Excellence (NCE) in Canada. The NCE program fosters powerful relationships between academia, government and industry.

can be found in accompanying papers in this volume. Maoh et al. (2005), for example, present the system architecture for IMPACTS-Hamilton (Integrated Model for Population Aging Consequences on Transportation Sustainability in Hamilton) – the GIS-based decision-support tool developed for the project. Mercado et al. (2005) evaluate the transportation policies of several developed nations in an effort to understand the extent to which these countries are altering their respective policies to cope with aging societies. Finally, Potoglou et al. (2005) detail the results from a series of trip generation models estimated for the Hamilton Census Metropolitan Area (CMA). These models, along with several others, form the travel demand module of IMPACTS-Hamilton.

2. Project Objectives

2.1. Automobility of Canada's Elderly Population

The first objective of our project documents at a national level evidence concerning the automobility of Canada's older population. In total, three investigations have been undertaken along this line of research using data obtained from Statistics Canada's 1986, 1992 and 1998 General Social Surveys on Time Use. Using cohort analysis, Newbold et al. (2005) find that while older Canadians undertake fewer trips and travel for different reasons than those in the labor force, the number of trips by car increases over the study period (i.e., 1986 – 1998). This implies that as the Canadian population ages, it will continue to depend on the private automobile for personal transportation.

Scott et al. (2005) investigate the changing travel behavior of elderly "urban" Canadians between 1992 and 1998. Perhaps the most important finding of this study is the fact that elderly drivers are traveling increasingly during the evening peak period. The changes in trip start times are shown in Figure 1.

Finally, Spinney et al. (2005) investigate inter-provincial and intra-provincial differences in elderly travel behavior in an effort to ascertain whether national transport policies pertaining to the elderly

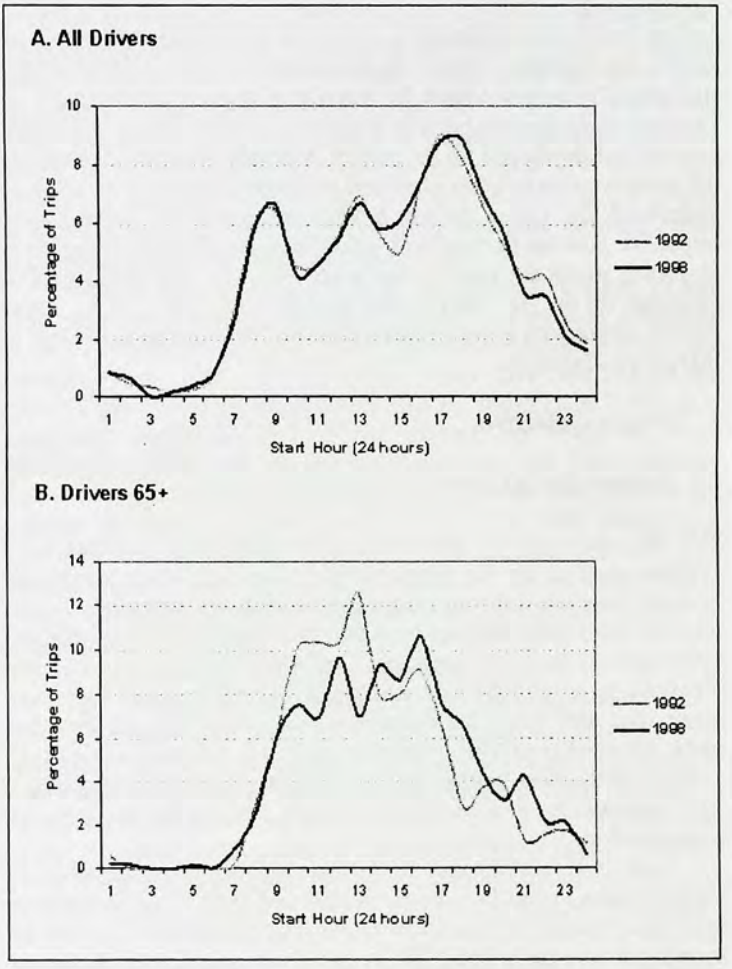


Figure 1. Start time of trips in CMAs for (A) all drivers and (B) drivers age 65+.

are possible. Preliminary evidence suggests that this may indeed be the case.

2.2. Model Estimation for the Hamilton CMA

The second objective of our project concerns the estimation of models underlying basic demographic processes and travel demand in the Hamilton CMA. Specifically, the demographic model is capable of projecting the spatial distribution of population by age and sex. The model accounts for in-migration and out-migration between the eight municipalities comprising the Hamilton CMA, the Greater Toronto Area (GTA), the rest of Ontario and the rest of Canada. The framework makes use of two components: the Rogers multi-regional population projection model and the aggregated spatial logit model. The Rogers model projects the population by age and sex for the CMA's municipalities and the neighboring areas. However, travel demand prediction operates at the level of traffic analysis zones. Thus, the spatially aggregated multinomial logit model is applied to the results of the Rogers model to obtain population estimates for such zones.

Trip generation, trip distribution and modal split models are estimated separately for the elderly and all other people in the Hamilton CMA. Moreover, models are estimated for four periods of the day: morning peak period (6:00 am to 8:59 am), day (9:00 am to 3:59 pm), evening peak period (4:00 pm to 6:59 pm) and night (7:00 pm to 5:59 pm). Potoglou et al. (2005) compare trip generation models estimated via traditional regression and non-traditional geographically-weighted regression (GWR) techniques.

2.3. GIS-based Decision Support Tool

The third objective of our project concerns development of a GIS-based decision support tool for evaluating the impacts of elderly travel behavior on Hamilton's transportation system and, more generally, the urban environment. This tool, known as IMPACTS-Hamilton, is described in depth by Maoh et al. (2005). An important feature of the tool concerns the fact that it employs data that are

readily available via trip diary surveys and the Canadian census. It is therefore possible to recalibrate the tool for any municipality in Canada.

2.4. Impacts and Policy Evaluation

The fourth objective of our project is to quantify over time the likely impacts of an expanding older population coupled with its increased automobility on Hamilton's urban environment. This objective has yet to be completed.

3. Summary

This paper has highlighted the four objectives comprising GEOIDE project DEC#SCO. The research undertaken in conjunction with these objectives provides Canadian policy makers with the necessary information and tools to address the transportation-related impacts of Canada's aging population – an issue that has received little attention until now.

Acknowledgement

This project was supported financially by GEOIDE (Geomatics for Informed Decisions) under grant DEC#SCO (Transportation Implications of Canada's Aging Population).

References

Burkhardt, J.E., McGavock, A.T., 1999. Tomorrow's older drivers: Who? How many? What impact? *Transportation Research Record* 1693, 62-70.

Carp, F.M., 1988. Significance of mobility for the well-being of the elderly. In: Transportation Research Board (Ed.), *Transportation in an Aging Society: Improving Mobility and Safety for Older Persons, Volume 2*. National Research Council, Washington, D.C., pp. 1-19.

Chen, J., Millar, W.J., 2000. Are recent cohorts healthier than their predecessors? *Health Reports* 11, 9-23.

Chu, X.H., 1994. *The Effects of Age on the Driving Habits of the Elderly*. U.S. Department of Transportation, Washington, D.C.

Dulisse, B., 1997. Older drivers and risks to other road users. *Accident Analysis and Prevention* 29, 573-582.

Hildebrand, E.D., 2003. Dimensions in elderly travel behavior: a simplified activity-based model using lifestyle clusters. *Transportation* 30, 285-306.

Maoh, H., Kanaroglou, P.S., Scott, D.M., Newbold, K.B., Páez, A., 2005. A GIS-based decision support tool to study the impact of elderly population on the transportation system in Hamilton Census Metropolitan Area (CMA). *Proceedings of the 2005 Annual Meeting of the Canadian Transportation Research Forum.*

Manual, D.G., Schultz, S.E., 2001. *Adding Years to Life and Life to Years: Life and Health Expectancy in Ontario*. ICES Atlas Report 1. Institute for Clinical Evaluative Sciences.

Mercado, R., Páez, A., Newbold, K.B., Scott, D.M., Kanaroglou, P.S., 2005. Transportation policy in an ageing society: an international comparison. *Proceedings of the 2005 Annual Meeting of the Canadian Transportation Research Forum.*

Newbold, K.B., Scott, D.M., Spinney, J.E.L., Kanaroglou, P., Páez, A., 2005. Travel behavior within Canada's older population: a cohort analysis. *Journal of Transport Geography*, forthcoming.

Organization for Economic Cooperation and Development (OECD), 2001. *Ageing and Transport: Mobility Needs and Safety Issues*. OECD, Paris.

Pisarski, A.E., 2003. Prescriptions for research: reviewing the history of TRB's critical issues in transportation. *TR News* 226, 30-35.

Potoglou, D., Páez, A., Scott, D.M., Kanaroglou, P.S., Newbold, K.B., 2005. Canada's elderly population travel behavior: experience from the Hamilton Census Metropolitan Area. *Proceedings of the 2005 Annual Meeting of the Canadian Transportation Research Forum.*

Rosenbloom, S., 2001. Sustainability and automobility among the elderly: an international assessment. *Transportation* 28, 375-408.

Scott, D.M., Newbold, K.B., Spinney, J.E.L., Mercado, R.G., Páez, A., Kanaroglou, P.S., 2005. Changing mobility of elderly urban Canadians, 1992-1998. Working Paper 007, Center for Spatial Analysis (CSpA), McMaster University, Hamilton, Ontario. Available from <http://sciwebserver.science.mcmaster.ca/cspa/papers /CSpA%20WP%20007.pdf >, March 2005.

Spinney, J.E.L., Scott, D.M., Newbold, K.B., 2005. Elderly travel behavior in Canada: Exploring intra-national and intra-provincial variations. Working Paper, Center for Spatial Analysis (CSpA), McMaster University, Hamilton, Ontario.

Stamatiadis, N., Deacon, J.A., 1998. Are older drivers safer today than in the past? Swedish National Road and Transport Research Institute, VTI Conference 9A Part 4.

Statistics Canada, 2005. Population Projections for 2001, 2006, 2011, 2016, 2021 and 2026, at July 1. Available from <http://www.statcan. ca/english/Pgdb/demo23a.htm>, February 2005.

Transportation Research Board (TRB), 1994. The licensing of older drivers. *Transportation Research Circular* 429.

Transportation Research Board (TRB), 2001. Critical issues in transportation. *Transportation Research News* 217, 3-11.

Transportation Policy in an Ageing Society: An International Comparison

Ruben Mercado (mercadrg@mcmaster.ca), Antonio Páez,
K. Bruce Newbold, Darren M. Scott and Pavlos Kanaroglou

Centre for Spatial Analysis/School of Geography and Geology
McMaster University

1. Introduction

The ageing of the population is one of the modern challenges facing many countries presently and will be increasingly felt over the next two or three decades (UN 2002). It is anticipated that an aging society will have a profound impact on future transport systems so much so that policy movements for its greater recognition in the national agenda have been quite substantial in Europe and North America (e.g. ECMT 2000; OECD 2001; Cobb and Coughlin 2000; Hodge 2004).

In recent years, transportation researchers have recognized the need to see the aging and transportation nexus beyond the usual social equity issues of mobility demand provision to embrace a wide range of health, economic, political, and environmental ramifications (e.g. TRB 1999; Rosenbloom 2001; Scott et al 2004). There is a widely held belief that today's older people and the baby boomers (which will eventually compose the next generation of older persons) have needs, preferences and concerns markedly different from previous generations. While there is yet limited evidence to support this current perception, recent empirical studies have shown that such cohort (or generational) differences are apparent especially in terms of travel needs and behaviour (e.g. Bush 2005; Newbold et al 2004; Scott et al 2004;). This change in features and lifestyles of older persons (which is undoubtedly linked, in general, to their improved health, greater economic resources and better education) will have significant bearing in terms of mobility satisfaction, safety and other societal issues such as urban sprawl, air pollution, environmental

degradation and congestion (e.g. OECD 2001; Scott et al 2004). Many of these concerns are now being articulated more explicitly to shape national transport policies in the context of an aging society and/or to craft a comprehensive national aging policy that cuts across the various development sectors (ECMT 2002; Cobb and Coughlin 1999; Hodge 2004). Canada's current vision and policy framework for transportation has acknowledged a diverse and aging population as one of the forces that will shape economy and society -- a societal change that transportation policy has to adapt in the future alongside the other modern challenges of "continentalism, globalization, environmentalism, urbanization and the high technology explosion" (Transport Canada 2003).

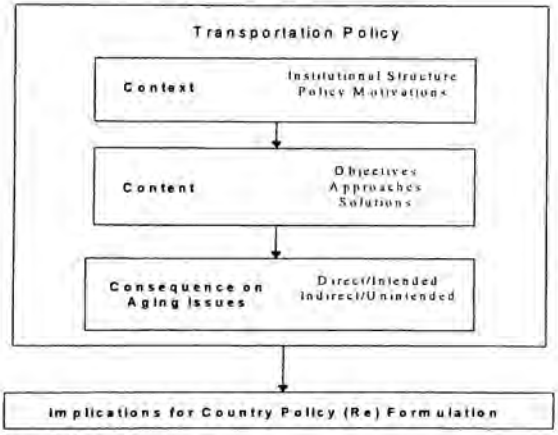
This paper provides a comparative evaluation of standing transportation policies across a number of developed countries. The objective of this comparison is to shed light on the extent to which these countries are gearing up their respective transport policies in the context of an aging society. What are the policy directions of these countries that will impact (directly or indirectly) on transportation needs of older adults now and in the future? What lessons can be drawn for refining or redefining Canadian transport policy that will respond to the challenges of a diverse and aging population that it has so recognized?

The countries for comparison include Australia, Japan, Netherlands, UK, and the US -- all aging and prosperous countries like Canada. While no country is expected to fit exactly the Canadian context, the countries possess certain similarities with Canada in terms of government structure (all study countries), economic system (US), law and culture (UK, Australia), population and economic geography (Australia) as well as social and environmental orientation (UK, Australia, Netherlands, Japan). The inclusion of Japan in the analysis, in particular, is motivated by an interest in the shape and path of transport policy of a country that has already reached the rate of older population that most industrialised countries will not yet achieve in at least ten years hence. The study compares land transport policy objectives, approaches and solutions with respect to key economic, financial, social, environmental and technological dimensions and their direct or indirect impact on older people's mobility. The investigation involves policy evaluation of white papers and related official transport policy documents of the respective

countries. Independent reports and studies from government, academic or research institutes, international organizations, and consultancy groups are also used to supplement the investigation and to identify gaps in terms of research and policy work.

Figure 1 shows the analytic framework of the study, which also reflects the structure of the paper. The second and third sections discuss the policymaking context of the countries being studied expounding on their varying institutional structures and policy motivations. The fourth section compares the articulated transport policy objectives. Next, transport policy approaches and solutions are surveyed and compared. The final section closes the report summarizing the lessons from the countries surveyed, and reflecting on their implications to Canadian transportation policy.

Figure 1
Analytic Framework of Study



2. Institutional Structures

Table 1 provides in a glance the general structure of government and the main transport institutions of the countries under comparison. Apart from the US, all the countries under study are constitutional monarchies with parliamentary form of government. The US, however, share with Canada and Australia a federal system that provides for relative autonomy

of the provinces or states and territories. All countries have basically a transport ministry but in recent years because of the need for policy coordination to integrate concerns of the environment, regional development and land use policy, some countries have either caused a reorganization of their national institutions to include these sectors (Japan, UK) or involved in a more systematic way other departments/ministries in shaping national transport policy (Netherlands, Australia, Canada).

Table 1

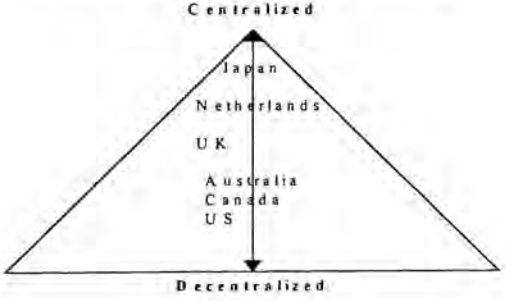
General Governance and Land Transportation Institutions in Comparative Countries

Country	Overall Government Structure/System	National Transport Institution/s	Sub-National Entities
Japan	Constitutional Monarchy Parliamentrary System Unitary State	Ministry of Land, Infrastructure and Transport (MLIT)	Metropolitan district (Tokyo) Urban prefectures (2) "Rural" prefectures (43)
UK	Constitutional Monarchy Parliamentary System Multinational Unitary State (England,Wales, Scotland, Northern Ireland)	Department for Transport (DfT) Office of Deputy Minister (Local Government and Regions)	Regional Governments (3) Greater London Authority Regional Development Agencies (12) Non-metro counties, unitary authorities, councils
Netherlands	Constitutional Monarchy Parliamentary System	Ministry of Transport, Public Works and Water Management (V&W)	Special Regions Provinces (12) municipalities
US	Presidential Federal Government	Department of Transportation (USDOT)	Metro Planning Orgs (MPOs) States (50) Municipalities, counties, townships (39,000) Special local govt. (44,000)
Australia	Constitutional Monarchy Parliamentary System Federation of States	Ministry of Transport and Regional Services Ministry of Local Government, Territories and Roads	States (6) Territories (2) Local Governments (722)
Canada	Constitutional Monarchy Parliamentary System Federation of States	Transport Canada	Provinces (10) Territories (3) Municipalities (3700)

The institutional set-up for transport policy formulation and implementation mirrors the general governance structure of the countries

under study (**Figure 1**). All countries, except the US, are constitutional monarchies and thus state control is expectedly evident. However, centralization occurs mainly in unitary states such as Japan, UK and Netherlands. In contrast, Canada and Australia, being federal states, are more decentralized in terms of institutional arrangements although provincial or state sub-national plans still have to conform to the national policy direction if they are to be granted government funding for costly infrastructure projects. This is also the case in the US as federal government (USDOT) administers federal funds and programs of which states compete for funding. There are recent debates and clamour for greater local control on transportation not only in the centralized but in the more decentralized governments as well. However, given the strategic role of transportation in terms of economic and national security, a more coordinated policy structure has often been the preferred arrangement. In fact, in Australia, where transportation system is fractured in terms of policy, planning and implementation given the autonomy of States and Territories, a national transport plan has been put in place to address regional interconnectivity (especially in freight transport) which is seen as a key requirement for the country's economic competitiveness, being a trading nation. Notwithstanding, the trend in most countries is towards deregulation of commercial transport and decentralization of transport policy, albeit, in a gradual fashion.

Figure 2
Transport Governance System



3. Policy Motivations

Given the variety in economic, social and political circumstances of the countries under study, it is hardly surprising to find distinct visions and thrusts as well as diverging mindsets in their policy directions. On the other hand, it is interesting to note that that these countries are in so many respects in unison in terms of policy end-goals. **Table 2** encapsulates these nuances in policy motivations and, for the purpose of comparison, classifies the same in terms of unique (distinct), common (converging) and differing (diverging) framework messages.

Distinctive. Population aging is a common thread in the overall policy framework in **Japan**. A basic law (Law No. 129, 1995), entitled "The Basic Law on Measures for the Aging Society", has been enacted in order to establish a clear objective, direction and measures to guide the national government and the public in dealing with an aged society. It is for this reason that demographic aging finds centrality in Japan's transport policy. What is also striking is the fact that an aging society is seen from a positive light ("opportunity") rather than viewed, as in many countries, as being problematic. Dealing with demographic changes in Japan sets the platform for harnessing the resources and technologies that are available to create a living environment that will make the country more competitive while addressing the needs of their older populace. The basic precept is that what is good for the elderly will also be for all ages and for the nation as a whole. In the case of the **UK**'s transport policy framework, health and environment take prominence over everything else. There has been an alarming concern that privatization and deregulation that has dominated transport policy for the past 20 years resulted in the decline of public transit damaging people's health and the wider environment. Thus, an "integrated transport policy" is at the heart of the nation's transportation pursuits, allowing "more real transport choice" by making better buses and trains and promoting walking and cycling as equally important travel modes, among others. **Netherland**'s policy framework is also about choices. However, it sees mobility choice as part of the modern society but the user has to take the cudgel of the costs of such freedom. In this respect, government interventions will primarily rely on market forces via pricing policy (e.g. paying for use of car vs. ownership, etc.).

Table 2
Key Messages/Policy Motivations of Country Transport Policy Frameworks

Country / Policy Framework Document	Distinctive	Converging	Diverging
Japan White Paper on Land, Infrastructure and Transport in Japan, 2002	Demographic aging as "opportunity" for transport policy to activate economy and society Towards seamless transportation	Population aging Economic / Global Competitiveness Environment Safety PPPs	Efficient, effective and accessible public transport service and infrastructure
UK White Paper on The Future of Transport (1998) " A New Deal for Transport: Better for Everyone"	Transport policy focus on health and environment "Integrated transport policy"	Environment Safety Expand transport alternatives to car PPPs Population aging	Focus on public transport Reduction of car use
Netherlands National Traffic and Transport Plan 2004-2020 "Nota Mobiliteit" (Mobility Policy Document)	Pricing policy takes central position in policy Decentralisation of traffic and transport policy Shift from supply-oriented thinking to client-oriented thinking	Environment Safety Expand transport alternatives to car Economic/Global Competitiveness PPPs Population aging	Respect for mode choice of the public / No one-sided solution to car vs public transit
US Strategic Plan 2003-2008 "Safer, Simpler, Smarter Transportation Solutions"	World-class transportation security system Transportation responsive to defense needs Changing demographics (aging and immigration) to alter transport requirements and governance	Safety Economic/Global Competitiveness Environment PPPs Population aging	Bring about important changes in all transport modes at the same time
Australia AusLink White paper 2004 "Building Our National Transport Future"	Better land use planning National and interregional Connectivity - focus on strategic network of transport corridors	Economic/Global Competitiveness Environment Safety PPPs Demographic trends	Public transport vs car is a State/Territory issue
Canada Vision / Strategy for Sustainable Transportation "StraightAhead: A Vision for Transportation in Canada(2003)	Challenge of a "scheduled economy" in tandem with environmental imperatives Research and skills development to support innovation	Safety Economic/Global Competitiveness Environment PPPs Population aging	Modal neutrality

A comprehensive pricing policy plan is being drafted and presented for legislation in 2005. The Netherlands, with a tradition of centralised transport policy-making, sees the need to decentralise its operations by

giving greater authority to sub-national entities to effect a more "intelligent and efficient organisation" to implement the national plan.

Of all the countries, the **US** has the highest concern for safety in its transportation policy framework. The repercussions of the 9/11 tragic events gave USDOT an urgent task to ensure safety and security in all modes of transport. Its mandate to establish a world-class transportation system will, thus, require the sector to be on the leading edge of technology. The US policy framework also identified future shapers of transport policy. One of these is the changing demographics (aging and immigration), which is reckoned to cause changes in the future transport requirements and governance - a shift towards increased state and local control of transportation. **Australia**'s transport policy framework focus on national and interregional connectivity triggered by the current "parochial and ad hoc system" that is seen as an impediment to national competitiveness. Land use planning is a key to the strategy and for which a "National Charter of Integrated Land Use and Transport Planning" has been forged to solidify the commitment of all policy actors. In **Canada**, one of the most difficult challenges that it faces is reconciling the needs of a "scheduled economy" (requiring highly reliable, tightly controlled, just-in-time transportation services) and the environmental imperatives. This tension will need clear and innovative policy guidance from the national government and collaboration with various stakeholders. Skills development, R & D and exploitation of new technologies are given a special importance in the overall transport vision viewing transportation as a knowledge-based sector.

Convergence. Economic growth and international competitiveness are paramount concerns of all countries reflecting the importance of transportation especially in a global economy. Safety is also a common concern in all countries, which refer to both national security (from terrorist attacks or calamities) as well as to personal safety in travel modes with the aim to reduce fatalities and accidents. Environmental impacts in terms of congestion and vehicle emissions affecting health and climate change are underlined in country frameworks. All countries have explicitly recognized demographic changes or the aging of the population as a key factor shaping transport system in the long-term especially its

implications to accessibility, equity and transportation infrastructure demands, albeit, in varying degrees. The level of this recognition is highest in Japan where aging and competitiveness are seen as complementary forces. In Canada, accessibility of transport for older people (along with remote regions and disabled) is part of the seven principles of transport policy. The US, UK and Netherlands view the challenge as part of issues on social inclusion and quality of life. Finally, the importance of public private partnerships (PPPs) at all government levels in terms of project implementation, innovation and funding are emphasized in all the country transport frameworks.

Divergence. A major and glaring difference in the policy frameworks of the comparative countries is the attitude placed on car use versus public transit. The policy choice on travel mode is critical as it will influence investment decisions as well as the application of economic instruments that will have consequences on older people's mobility. The **UK** has the most radical and strong policy against car use while providing strong preference to public transit. **Japan** has a comprehensive plan for public transit but seems to be hushed on car use. In the **Netherlands** both pull measures (making alternatives to the car more attractive) and push measures (making car less attractive) are adopted. Since the latter measures are deemed lagging behind the former, there are now serious efforts to adopt and implement pricing policies with more vigour to balance the policy package. The **US and Canada** share a common attitude of modal neutrality, providing a level playing field for all modes. In the case of **Australia**, the federal government left the policy decision to States and Territories. However, it views that walking and cycling should be promoted through land use planning and provision of transport infrastructure to address the national problem of obesity, which affects more than half of the Australian population. This is also reflected in its National Land Use Charter where widening of transport mode choices is seen to increase accessibility and reduce vehicle demand and impacts.

4. Transport Policy Objectives

The policy motivations discussed above find translation in the specific policy objectives for transport. **Table 3** conveniently summarizes

the transport objectives into three categories – economic, environmental and social/political – encompassing the three elements of sustainable transportation system.

Table 3
Comparative Matrix of Standing Transport Policy Objectives
By Country and Policy Themes

Theme / Country	Economic	Environment	Social/Political
Japan	Enhance global competitiveness Ensure smooth transportation everywhere	Achieve environmentally friendly transportation Develop living environment that offers peace of mind to seniors	Ensure safety (personal and national) Deal with graying society
UK	Better planning for development, land use, transport and housing	Better protection for the environment Better places to live Better buses Better trains	Better safety and personal security Fairer, more inclusive society
Netherlands	Strengthen the economy and international competititve position	Improve Quality of life and environment	Accommodate need for mobility Improve safety More responsibility to decentral government
US	Global connectivity	Environmental stewardship	Ensure safety- save lives and reduce accidents Address mobility - congestion and accessibility (inclusiveness)
Australia	Supporting economic growth and social cohesion Managing increases in transport activity	Addressing growing social and environmental costs	Integrated approach to planning and investment (coordinated decisions)
Canada	Efficient and safe movement of people and goods (well-integrated transportation system)	Respect for environment	Reasonable access to national transportation system for remote regions Accessibility and social inclusion

Finding the right balance among these three elements is the fundamental policy challenge in these countries. Each country has differently expressed each of these generic objectives reflecting their

varying circumstances. For instance, Japan, the US and the Netherlands place emphasis on global integration of transport given the role placed on transport in their respective competition policies. While this is also true for the rest of the countries under study, more pressing in-country (urban and inter-urban) concerns of congestion relief and freight bottlenecks surface in most instances in Canada, Australia and UK in view of the anticipated rise in transport activities resulting from expanded trading and business activities. In terms of environmental objectives, US and Canada emphasize commitment to environmental responsibility - sharing a strong resolve to bring balance to economic growth without undermining its responsibility to preserve environmental quality. The other countries frame the environmental objective in terms of problems needing solutions (Australia) or desired outcomes (UK, Japan, Netherlands). Safety and social inclusion are very high on the agenda in the countries under study. Social inclusion is explained in the narrow context of more accessible transportation and/or in broader social and human concerns.

5. Transport Policy Approaches and Solutions

Transport objectives become more diverse when translated to approaches and solutions. **Table 4** takes stock of the various approaches, solutions and instruments of the various countries and classifies them according to policy areas that would directly or indirectly impact on the mobility of older people. These policy areas include modal choices, transportation infrastructure, systems and services, technological solutions, land use and living environment and institutional and policy reforms. In the interest of space, this discussion will be abbreviated to only mention some general observations. It should be mentioned that although most countries have adopted an accommodationist approach to mode choice, transport solutions that promote public transit are still considerable. However, the UK has shown a radical and serious approach and solution to give public transit and other alternative modes of travel a greater edge over the private car not only through the adoption of tough economic instruments but also investments on technological solutions not to the private car (as the rest of the countries do) but to modernize buses and railways. Land use planning and promotion of a better living environment are emphasized in Japan, Netherlands, Australia and UK.

Table 4
Comparative Transport Policy Approaches and Solutions Impacting on Mobility of Older Population

Policy Area	Approaches / Solutions / Instruments Symbols : J=Japan, UK=UK, N=Netherlands, A=Australia, US=USA, C=Canada
Modal Choices	
Car / Driving	Reduce car use through pricing and taxation (UK)
	Shift from fixed car charges to charges dependent on vehicle use (distance and time of day charging) (N)
	Strict enforcement of driving laws (N, US) ; seatbelt laws (US)
	Better driving education (N)
	Special policy focused on drivers "who are getting on in years" (N)
Public Transit (Rail, Bus)	Increase modal share of public transport through market-oriented approach (N)
	Develop a reliable and financially viable intercity passenger rail network (US)
	Put greater emphasis on public transit infrastructure projects (C)
	Towards seamless transportation system (logistics, ticketing, information) (J)
	Single travel ticket (N)
	Barrier free transit (railway cars and low-floor buses) and facilities (stations, bus stops) (J, UK)
	Modern design and better services (UK)
	Government investment in rail, tram and bus infrastructure with quality service (N)
	Accessibility criterion for new city buses (N)
	Full accessibility for bus by 2010 and rail equipment by 2030 (N)
	Scheduled bus service deregulation and liberalization of market entry (reverse onus system (C)
	IC or smart Card in public transport (J, N)
	Bicylce storage spaces in rail stations (N)
	New transport services in areas w/o public transport (J)
Others (Walking / Biking / Special Services Schemes)	Formulate national action plan on cycling and walking (UK)
	Overcome infrastructure barriers (UK)
	Promote walking and cycling through land use planning (A)
	Creative alternatives to traditional individual transportation for elderly with disabilities and disabled (US)
	Public transport subsidies and fare schemes (UK)
	On demand bus (J)
	Shared taxis UK)
	City car clubs (UK)
	Encourage public involvement to develop transport solutions that support community needs (US)
Transport Infrastructure/ Systems and Services	
Modal Integration	Move to a more integrated transport system (C, A, US)
	Encourage and facilitate intermodal transportation planning worldwide (US)
	Given growth of elderly traveling for leisure, adaptation of transport services in international transit and destinations (US)
	Seamless journey (J, UK)
Road Design	Installation of "essential recognisability features" (road, speed limit) (N) indicators for crosswalks (J)
	Residential areas with 30 km/h speed limits (N)
	Improved road safety features such as enhanced signage in view of aging (US)

Policy Area	Approaches / Solutions / Instruments Symbols : J=Japan, UK=UK, N=Netherlands, A=Australia, US=USA, C=Canada
Technological Solutions	
Vehicle Design	Advanced safety vehicle - car (J, N)
	R & D for car model for people with limited mobility (J)
	Encouragement of technological innovations through fiscal incentives (N)
	Consideration of elderly in designing and building vehicles (US)
ITS	Electronic toll collection (ETC) system (J,N)
	ITS to aid pedestrians (J)
	R & D in transport telematics application (ITS) related to public transport
	Sensors placed on vehicles and infrastructure to analyse traffic flows, behavior emissions (N)
	ITS for global connectivity at key transfer points, corridors and border crossings (US)
	Develop an ITS R&D plan (C)
Environment Friendly Vehicles	Development and Promotion of Fuel Cell technology (J, US)
	Development of cleaner, quieter and fuel-economical vehicle technologies (fuel cell and and bio-fuel) (N, A)
	Promote hybrid cars and motorcycles (N)
	Encourage the transition to hydrogen economy through R&D of advanced vehicle technologies (US)
	Stimulate demand for new, more fuel efficient vehicles akin to Energy Star system (C)
	Encourage use of cleaner diesel technology through tax incentives (UK)
Land Use and Living Environment	Good quality residence and housing through tax and financial measures (J)
	Promotion of barrier free buildings and passage to buildings (J)
	Enforcement of standards for automobile exhaust emissions (J)
	Beautiful landscapes and plentiful greenspace (J)
	Improve shopping environment and create better convenience (J)
	Early introduction of public transportation to new housing / development area (UK)
	Revise planning guidance on new housing devt and land use planning for regional/local level; end to mushrooming of out of town shopping centres (UK)
	Transport policy and plans based on national spatial planning policy (N)
	Integrate land use planning and transport infrastructure provision based on National Land Transport Plan (A)
	Explore collaborative efforts to establish supporting transportation management and land use planning frameworks (N)
Institutional Changes / Policy Management	Decentralize traffic, transport and land use policy and planning (N)
	Private sector (transport operators) to control tactical and operational level of public transport services; local govt define services at the strategic level (N)
	Tranportation ministry links with other policy fields including environment, spatial planning, economy, health care and education (N)
	Address and deliver infrastructure spending within a new strategic framework - involving all levels of government and private sector (A)
	Collaborate with industry, all levels of govt and academia on options for road pricing and urban transit infrastructure C)
	Nurture innovation clusters for research and technological application (C)
	Work with community of disabled, senior citizens and transport industry to develop a long-term, multi-modal federal strategy for enhancing accessibility of the national transportation system (C)
	Deregulation in transport industry (J, N)

6. Summary and Implications to Canadian Policy

Beyond Accessible Infrastructure. Like the other countries under comparison, Canada has amply recognized the aging of the population as one of the formidable challenges that will shape transportation in the future. In its transport vision, it has indicated that it will work with the disabled, senior citizens and transportation industry to develop a "long-term, multi-modal federal strategy for enhancing the accessibility of the national transportation system". While this is undoubtedly a step in the right direction, there is a need to go beyond this seeming mindset of providing accessible infrastructure to deal with the mobility needs of older people. This study has shown the breadth of concerns that directly and indirectly impact on older people's mobility and the possibilities of solutions that can be adopted depending on a country's policy orientation. These include policies dealing with: a) travel modalities and the economic instruments that would promote them; b) investments on transport infrastructure and systems that enhance driving or the use of alternatives to the car; c) technological advances that improve the travel mode and the transport environment; d) the strong linkage between transportation and land use planning as it impacts on the living environment of older people; and e) institutional reforms and policy management initiatives that will enhance and help achieve program formulation and implementation.

Research Strategies and Future Directions. There have been significant efforts in other countries especially in the US to allow for empirical and policy-based research to develop a better understanding of the transport implications of an aging society. These studies have contributed to modest policy progress in terms of the significant reflection of aging concerns in national policy, rallying the transport industry to look into the older people's mobility needs and encouraging broader discussion of community design and sustainability for an aging population (Cobb and Coughlin, 1999). However, a number of issues remain especially on the dearth of data that will allow for aged-based transport policy research necessitating a revisit of travel survey design. Interdisciplinary research is also needed. In the US, there are plans to include the private sector and disciplines not usually included in the aging policy community such as business, urban planning, transportation, housing, environment, and law for a "Boomer Interdisciplinary Research" (Hodge 2004). Canada seems to be well-poised in undertaking similar effort given its vision to develop

transportation as an innovative knowledge-based sector and the already well-placed government-launched Canada's Innovation Strategy. Notwithstanding, other complementary research programs can be put together to look into transport-related issues surrounding the more diverse aging baby boomers.

References

Bush, S. 2005 "Forecasting 65+Travel: An Integration of Cohort Analysis and Travel Demand Modeling", Paper presented at 2005 TRB annual meeting.

Cobb, R. W. & Coughlin, J. F. 1999, "Transportation for an Aging Society: Keeping Older Americans on the Move," in *Transportation in an Aging Society: A Decade of Experience*, TRB Conf. Proceedings 27

Department Environment, T. a. R. 1998, *Integrated Transport White Paper: A New Deal for Transport* United Kingdom.

Department of Transportation 2003, *Strategic Plan 2003-2008: Safer Simpler, Smarter Transportation Solutions* United States of America.

Department of Transport and Regional Services 2004, *AusLink WhitePaper: Building our National Transport Future*, Australian Govt.

European Conference of Ministers of Transport (ECMT) 2002, *Transport and Ageing Population.*

Hodge, P. 2004, *Living Younger Longer: Boomer Challenges.* 10-1-2004. Policy Committee for 2005 White House Conference on Aging.

Ministry of Land, I. a. T. 2002, *White Paper on Land, Infrastructure and Transport in Japan*, Government of Japan.

Newbold, K. B., Scott, D., Spinney, J., Kanaroglou, P., & Páez, A. 2004, "Travel behavior within Canada's older population", *Journal of Transport Geography*, vol. article in press.

Organisation for Economic Cooperation and Development (OECD) 2001, *Ageing and Transport: Mobility Needs and Safety Issues* OECD, Paris.

Rosenbloom, S. 2001, "Sustainability and automobility among the elderly: An International Assessment", *Transportation*, vol. 28, pp. 375-408.

Scott, D.M., Newbold, K.B., Spinney, J.E.L., Mercado, R.G., Paez, A. and Kanaroglou, P.S. (2005) *Changing mobility of elderly urban Canadians, 1992-1998.* Working Paper, McMaster University, Hamilton, Canada

Transport Canada 2003, *Straight Ahead: A Vision for Transportation in Canada*, Government of Canada.

Transportation Research Board 1999, *Transportation in an Aging Society: A Decade of Experience* , Conference Proceedings 27.

United Nations 2002, *Report of the United Nations Second World Assembly on Ageing*, United Nations.

Canada's Elderly Population Travel Behaviour: Experience from the Hamilton Census Metropolitan Area

Dimitris Potoglou (potogld@mcmaster.ca), Antonio Páez, Darren M. Scott, Pavlos S. Kanaroglou, K. Bruce Newbold

Centre for Spatial Analysis / School of Geography and Geology
McMaster University
1280 Main Street West, Hamilton, Ontario, L8S 4K1, Canada

1. INTRODUCTION

Canada, like a majority of countries in the industrialized world, is an aging society that will face significant increases in both the number and percentage of older people in the near future. It is estimated, for instance, that by 2021 about 1 in 5 Canadians will be 65 years or older (Statistics Canada, 2003). It is believed that the transportation implications of aging societies are far reaching. At the same time that the aging situation has been recognized as one of the critical challenges for transportation systems (Pisarski, 2003), research to date in several national settings suggests that the elderly in developed countries have become relatively more mobile, with a tendency to make more and more varied trips, while using less public transportation as they have become increasingly automobile dependent (e.g. Rosenbloom, 2001).

Recent research suggests that these international trends in travel behaviour are to some extent shared by Canada (Newbold et al, 2005). In due course they are expected to have significant impacts on Canada's transportation systems (e.g. traffic congestion), urban environments (e.g. air pollution), and the ability of transportation systems and service providers to meet the mobility needs of the elderly (a critical aspect of their quality of life). Despite the clear

importance of this array of issues, our understanding of trip making by the elderly remains for the most part descriptive, and tools to project their future transportation needs are sorely missing. The Urban Transportation Modelling System (UTMS; Meyer and Miller, 2001), the most commonly used tool for assessing urban travel in Canada, is still confined to the evaluation of work-related peak-period travel and ignores the impact of older drivers on traffic congestion and air pollution.

This paper is part of a larger project that aims at developing improved models that take into account these emerging demographic and travel trends. This paper concentrates on the topic of trip generation. A series of regression-based trip generation models are estimated for different age groups (i.e. elderly vs. non-elderly), using spatial analytical techniques to address two research questions: are there age-related differences in trip generation rates, and the factors affecting these rates? And, are these differences dependent on geographical context (e.g. suburbs vs. inner city)?

From an applied perspective, in addition, the models reported here are the first-step in the UTMS, and are therefore a necessary element to project travel demand needs in a region. Eventually, the models will be incorporated into a GIS-based decision-support tool, which will simulate travel demand, traffic congestion and emissions for both work and non-work trips for elderly and adults over a day period. The case study is the Census Metropolitan Area (CMA) of Hamilton, Canada. Data for analysis is drawn from the Transportation Tomorrow Survey, a travel survey conducted every five years in the Greater Toronto Area (GTA).

2. BACKGROUND: TRIP GENERATION ANALYSIS

Conceptually, the issue of trip generation can be thought of as trying to answer the question of how many trips will be produced by a zone, in what usually is the first step in the Urban Transportation Modelling System (for discussions of this system see Ortúzar and Willumsen, 2001; Meyer and Miller, 2001; Miller and Shaw, 2001). This step is followed in a sequential fashion by the analysis of trip distribution (where are those trips going), modal split (how many

trips are made by each transportation mode) and assignment (route selection). Trip generation, being the initial step in this modeling approach, provides the benchmark level of travel demand in a system.

2.1 Regression-based trip generation analysis

Regression analysis is what has been termed a mature technique, one whose properties and characteristics are well understood (Washington, 2000). Models of this type are used in trip generation analysis in two ways. First, regression models can identify variables that highly correlate with the number of trips originating at the unit of analysis. These variables can be of policy significances, when they represent measures over which planners have some amount of control (e.g. development densities). Alternatively, variables of predictive interest are also frequently used (e.g. household structure, age). Although planners may not have control over the latter, strong correlation with number of trips is a useful property for the second objective of trip generation analysis, which is usually to forecast, given changes in the levels of these variables, the number of trips that will be produced by each zone in a region.

Regression analysis can be conducted in two different ways, depending on the unit of analysis. First, zonal-based regression makes use of variables at the zonal level, be it totals (e.g. population in a zone, total number of vehicles, etc.) or rates (people per household per zone, etc.) Modeling the rates is a preferred approach to deal with size effects that tend to plague models that use totals.

A regression equation for trip generation rates is defined as follows, with T_i as the number of trips (the dependent variable), and independent or explanatory variables X_k (e.g. number of household members, number of vehicles in a household, etc.):

$$t_i = \beta_0 + \sum_{k=1}^{K} \beta_k x_{ki} + u_i \qquad (1)$$

The variables transformed in this way would be representative of a typical household in a zone. The coefficients β_k in equation (1) represent the rate of change in trip rates with changes in the corresponding explanatory variable. These coefficients can be

estimated by the method of ordinary least squares under the assumption that the errors terms ε_i are normally distributed, independent and have a constant variance factor. Using the rates achieves the objective of accounting for the effect of size of a zone, in fact reducing the risk of non-homogeneous error terms. There are other advantages to modeling the rates. Their use tends to produce better behaved, more symmetric variable distributions that also present fewer extreme values. And although the coefficient of determination R^2 tends to be much lower when compared to the use of totals, models of the rates are often more informative in terms of the demographics of the problem (multicollinearity is also reduced). It is important to note, however, that the information is at an aggregate level; in other words, the rates are the average number of trips for the average household in a zone.

2.2 Spatial analytical issues

The quality of a regression-based trip generation model is usually assessed on the basis of three criteria: goodness of fit (coefficient of determination and significance of coefficients), theoretical consistency (evaluated based on the signs and magnitudes of the coefficients), and major assumptions (i.e. linearity, normality, independence, constant variance). These criteria are well understood, and are discussed in detail in all major transportation planning and engineering books (e.g. Ortúzar and Willumsen, 2001; Ortúzar and Willumsen, 2001).

Less well known are the implications of modeling what in essence is a spatial process, although recent reviews document an encouraging trend towards a wider recognition of the role of space in transportation modeling (Miller, 1999; Páez and Scott, 2004). A major analytical issue that should be considered when modeling spatial data is spatial heterogeneity. Urban processes (including trip generation) often exhibit patterns of spatial heterogeneity – that is, they do not always operate in exactly the same way over space. This may be a result of large-scale regional effects or administrative subdivisions that delimitate the reach of some process (e.g., zoning).

In statistical terms, spatial heterogeneity can be represented as systematic variation in the mean of the process (spatial parametric

drift), or alternatively as structural variation in the definition of the variance (non-constant variance models). The method of geographically weighted regression corresponds to the first class, as it attempts to capture variation in the coefficients of the model. GWR is particularly attractive in the context of trip generation analysis because it can be conceptualized as a form of "soft" spatial market segmentation. The method can thus be used to explore spatial variations in trip rates, and in particular the relationships between these rates and the factors that influence them.

2.3 Using geographically weighted regression for trip generation

Geographically weighted regression, or GWR for short, is a local form of spatial analysis proposed by Brunsdon et al. (1996) as a simple way of modeling complex spatial variation. The method is based on the geographical concept of distance decay, and the idea of decreasing the weights assigned to individual observations according to their distance from a location termed the focal point. This form of weighting in fact operates as a moving window that can be displaced over a set of observations in order to produce sub-samples of data around specific points in space. The result is locally varying relationships.

GWR is a semi-parametric model that takes the following form:

$$t_i = \beta_0(i) + \sum_{k=1}^{K} \beta_k(i) x_{ki} + u_i \qquad (2)$$

Since observations in trip generation analysis are geo-coded, subscript i in equation (2) above indicates the location of the observation (a zone), and also the location of the focal point for estimation. Note that, unlike the regular regression model used in trip generation [equation (1)], the coefficients in GWR are not spatial constants, but rather are a function of location. It is assumed that the error terms u_i in the above expression are normal, independent, and have constant variance (Brunsdon et al, 1999, p. 502).

The moving window used to locally estimate the model in (2) is defined by a monotonically decreasing function, or kernel, such as:

$$w_{ij} = \exp\left(-\gamma d_{ij}^2\right) \tag{3}$$

where d_{ij} is the distance between the location of the focal point i and observation j. and γ is a parameter called a kernel bandwidth that controls the size of the window and the number of observations in the sub-sample. A commonly used strategy for the selection of this parameter is based on a procedure known as cross-validation. This is essentially a goodness of fit criterion that minimizes the total sum of squared errors between the observed values and the values predicted by the model. Once a value of γ has been obtained, this parameter can be plugged into the following expression (in matrix notation), to obtain a vector of local coefficient estimates. These local coefficient estimates can be used to explore spatial parametric drift:

$$\hat{\boldsymbol{\beta}}(i) = \left(\mathbf{X}^T \mathbf{W}(i) \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{W}(i) \boldsymbol{t} \tag{4}$$

where $\mathbf{W}(i)$ is a diagonal matrix with weights given by equation (3) and $\mathbf{X}$ and $\boldsymbol{t}$ are a data matrix and a vector of trip rates respectively.

GWR is often interpreted as a smoother (it is closely related to the method of locally weighted regression, a scatterplot smoother proposed in statistics by Cleveland, 1979). For this reason, GWR is often capable of obtaining very accurate approximations of the observed variable surface. This feature of the method makes it very attractive for applications in which forecasting is an important objective.

3. CASE STUDY

The following sections provide a brief description of the study area, the travel survey data used, and the results of the models.

3.1 Study Area

The Hamilton CMA is located on the west shore of Lake Ontario approximately 100 km southwest from the city of Toronto. It consists of the New City of Hamilton, the City of Burlington and the town of Grimsby. In 2001, Ancaster, Dundas, and Flamborough, the

Township of Glanbrook and the Cities of Hamilton and Stoney Creek were amalgamated and became the new City of Hamilton.

The population of the area is approximately 650,000 (Census of Canada, 2001) [1] and the new City of Hamilton enumerates the majority of the population (500,000). However, Waterdown (eastern lower part of Flamborough) and Ancaster had experienced the highest population growth in the CMA, at rates of 28.9 and 17.4 percent respectively, between 1996 and 2001. This is the 4th largest urban centre in Ontario and the 9th in Canada.

With regards to the demographics of the region, the 2001 Canadian Census revealed a significant increase in the region of the elderly population aged 75 years and over. A characteristic of this cohort is a dispersed pattern of their places of residences (Figure 1).

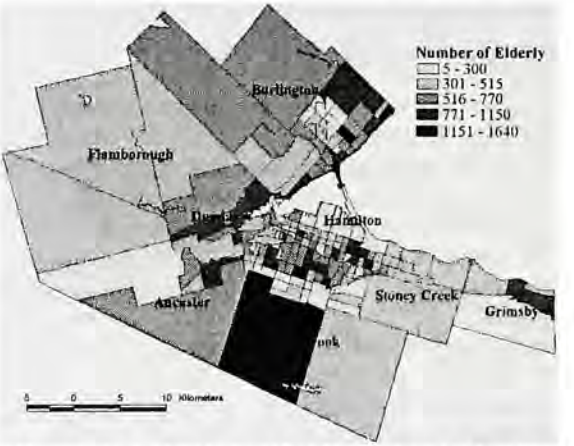


Figure 1. Distribution of Elderly (65+) Population in Hamilton CMA (*Data Source: Canadian Census, 2001*)

3.2 Data - The Transportation Tomorrow Survey (TTS)

The analysis is based on the Transportation Tomorrow Survey. The TTS is a comprehensive travel survey conducted in the Greater

[1] http://www12.statcan.ca/english/census01/home/index.cfm

Toronto Area (GTA) and its surrounding regions once every five years[2]. For the purpose of our analyses we make use of a subset of the 1996 TTS data that corresponds to Hamilton CMA. The dataset includes detailed information of approximately 10,500 households, 27,000 individuals and 51,300 trips recorded on the basis of traffic a analysis zone (TAZ).

Table 1 shows the percentage distribution of trips in the TTS sample by purpose (work, non-work), time of the day-period and age-group, excluding trips of individuals less than 15 years of age (6.2 percent of all trips). A majority of the trips is made during the off-peak day period. Many other trips correspond to commuting trips (from- and to- work) in the morning and afternoon peak period. Almost all elderly trips take place during the day non-peak period accounting in total for 11.28 percent of the all trips in the sample.

Table 1. Percentage of trips by age-group, purpose and time of the day

	Age Group				
	15 < Age < 64		**Age > 64**		**Total**
Age Group in the Sample (%)	65.35		13.90		79.25
Time of the Day	**Non-work**	**Work**	**Non-work**	**Work**	
Morning Peak (6:00 - 8:59)	5.94	11.21	0.38	0.18	17.71
Day Non-Peak (9:00 - 15:59)	23.69	3.28	7.73	0.09	34.79
Afternoon Peak (16:00 - 19:00)	20.17	0.60	1.93	0.01	22.71
Night (19:00 - 6:00 next day)	16.17	1.15	1.24	0.01	18.57
Total	65.97	16.24	11.28	0.29	93.8

Source TTS, 1996

[2] http://www.jpint.utoronto.ca/

3.3 Trip generation models

Trip generation analysis was undertaken using the traditional method of zonal regression and a local form of spatial analysis, the GWR. Using these approaches, adult, elderly and the combination of elderly and adult trips are modeled for the morning peak period.

Zonal - Regression Trip Generation Models

The zonal regression models consider the average number of household trips per TAZ as the dependent variable whereas average household socio-demographic information serves as independent variables. The latter were first expanded using the corresponding sample weights to account for the total population in the study area. Consequently, average household characteristics (e.g., number of vehicles per household) were aggregated at the zonal level.

Table 2 shows parameter estimates for six models each corresponding to a specific age-group and purpose during the morning peak period. In all cases parameter estimates have the expected sign and order of magnitude, however, for comparison purposes between the different models, some specifications include insignificant parameters.

The limited number of elderly work trips (0.18 percent of the sample) prohibits the estimation an acceptable zonal regression and GWR models. All other models demonstrate acceptable levels of goodness-of-fit, which ranges between 28 and 41.3 percent.

The average number of full-time individuals (FULL_EMP) and adults (ADULTS) in a household per zone appear to be the most significant variables, contributing the most to adult and combined adult-elderly (All Ages) work and non-work trips. Also, households with children between 5 and 12 years of age are less likely to undertake work and non-work trips during the morning peak period.

Elderly non-work trips are more likely to occur as the average number of children between 5 and 12 years of age per household increase. This is the same type of relationship found for the case of adult and elderly individuals per household per zone.

Household based regression is not explored in this paper because geo-codes are not available at the level of the household. This prevents any sort of spatial analysis of trip generation – however, if

analysis finds that there is significant variation, the following step could be to try to obtain geo-codes for a spatial analysis at the level of the household.

GWR - Trip Generation Models

The models estimated in the preceding section can be termed global, because they do not account for spatial variability of the parameters. In this section, a series of GWR models were estimated, using the same set of variables and classification of trips. The results of these models appear summarized in Table 3. The figures in the table suggest that there is some substantial variation in terms of the spatial relationships. Although the mean of the spatial coefficients is close to the global coefficient in a number of cases, there are also a number of sign reversals. The range of variation is also considerably large.

The spatial distribution of two coefficients appears mapped in Figures 2 and 3. These are those corresponding to the (average) number of elderly persons per household (β_2) in the non-work trips model). There is a clear north-south trend in the variation of coefficient β_2 in the case of elderly trips, although the range of variation is not large (about 14% between minimum-maximum values), whereas in the case of adult non-work trips, this same coefficient presents substantially more variability (a range of variation of about 140%). The variability is also more complex, suggesting that this variable correlates with a higher trip generation rate in the central part of the city, and in some zones to the west and north of the city. It is worth noting that this variable was not significant in the global model (see Table 2). Examination of the other variables provides some insight into the factors that affect trip generation rates.

Table 2. Parameter Estimates for the Morning Peak Period Trips by Trip Purpose and Age Group

	ADULT		ELDERLY		ALL AGES	
	WORK	NON-WORK	WORK	NON-WORK	WORK	NON-WORK
INTERCEPT	0.224	0.166	-0.011	0.001	0.212	0.167
	1.86	0.89	-1.17	0.09	1.76	0.87
VEH/HLD/ZONE (vehicles per household)	.	-0.391	.	-0.011	.	-0.402
		-4.25		-1.51		-4.25
ELDERLY (number of people 65+ per household)	-0.142	0.127	0.020	0.069	-0.122	0.195
	-1.34	0.71	2.40	4.68	-1.14	1.07
ADULTS (adults (15 to 64) per household)	0.249	0.344	0.008	-0.003	0.257	0.340
	4.37	2.79	1.71	-0.29	4.47	2.68
CHILD_5_12 (children (5 to 12) per household)	-0.284	-0.017	-0.001	0.029	-0.285	0.012
	-3.57	-0.13	-0.12	2.82	-3.55	0.0079
FULL_EMP (full-time employed persons per household)	0.555	0.6762	-0.002	-0.004	0.553	0.672
	9.88	8.12	-0.42	-0.59	9.78	7.84
FT_EMPH full-time employed persons working at home per household	.	0.472	.	-0.047	.	0.425
		1.95		-2.33		1.71
WH_PT part-time employed persons working at home per household	1.447	3.714	-0.028	0.496	1.418	4.210
	2.17	3.64	-0.53	5.88	2.12	4.01
R^2	0.434	0.318	0.032	0.387	0.428	0.304
ADJ-R^2	0.412	0.293	0.008	0.365	0.413	0.280
zonal averages						

Table 3. Geographically weighted regression models: Summary of results.

		Adult Trips		Elderly Trips	
Variable		**Work**	**Non-Work**	**Work**	**Non-Work**
	Max	0.873	2.223	0.061	0.008
INTERCEPT	Mean	-0.025	0.095	-0.012	0.005
	Min	-1.251	-1.072	-0.484	0.000
	Max		0.169		-0.007
VEH_HLD_ZONE	Mean		-0.404		-0.011
	Min		-0.708		-0.015
	Max	1.234	0.602	0.375	0.067
ELDERLY	Mean	0.035	0.209	0.020	0.064
	Min	-0.302	-1.504	-0.065	0.059
	Max	0.803	0.736	0.176	-0.001
ADULTS	Mean	0.379	0.323	0.008	-0.003
	Min	-0.298	-0.546	-0.011	-0.006
	Max	1.141	0.917	0.210	0.034
CHILD_5_12	Mean	-0.416	0.195	-0.001	0.027
	Min	-0.684	-0.688	-0.053	0.021
	Max	0.821	1.010	0.078	-0.004
FULL_EMP	Mean	0.634	0.669	-0.006	-0.005
	Min	-0.654	-1.125	-0.284	-0.008
	Max		4.426		-0.039
FT_EMPH	Mean		1.393		-0.045
	Min		-0.683		-0.051
	Max	3.207	7.458	1.341	0.465
WH_PT	Mean	1.451	4.071	-0.058	0.416
	Min	-1.811	-3.167	-0.560	0.346



Figure 2. Adult Non Work Trips: Variable Elderly

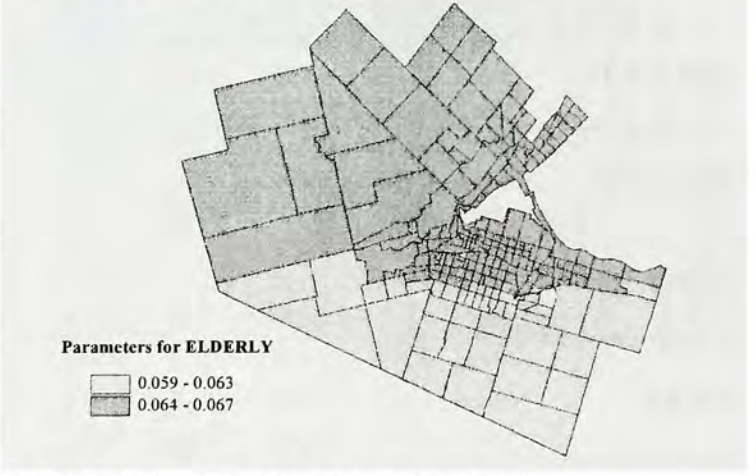


Figure 3. Elderly Non Work Trips: Variable Elderly

4. SUMMARY AND CONCLUDING REMARKS

Canada's population is aging, and this demographic trend is expected to have a significant impact on urban transportation systems. The objective of this paper has been to analyze trip generation trends by classifying the population into two major age groups, namely elderly (65+) and non-elderly. The study made use of regression analysis, a traditional approach in the analysis of trip generation. These models suggest a relative ease in terms of modeling work trips undertaken by the non-elderly (the coefficient of determination of this model was relatively high) and non-work trips by the elderly (with an almost equally high R^2). In addition, a spatial method known as geographically weighted regression was applied in order to explore the question of whether there are contextual variations in the nature of relationships. There appears to be some substantial variation. For example, number of elderly per household correlates with trip rates more strongly in the north of the study area than in the south. The patterns are even more complex in the case of adult work trips.

The modeling exercise in this paper falls squarely within the UTMS paradigm, in terms of conducting analysis at the aggregate level. In recent years, on the other hand, a wealth of new methods has been developed parting from the proposition that such systems are best studied from the perspective of the individual, and her need to engage in work- and social-related activities (McNally, 2000). Despite their conceptual appeal, however, many of these advances have yet to become widely used in practical planning applications. For reasons related to data needs (e.g. availability and confidentiality issues), and the level of technical expertise required for implementation, the state-of-the-practice continues to rely on the UTMS. One objective of this paper was to explore the use of spatially-explicit methods as a way to refine conventional transportation modeling practice. The experience suggests that these models can give more detailed insights into trip generation. Future research should compare the performance of these models with respect to other approaches, such as frequency modeling using nested logit or ordered probit models.

REFERENCES

Brunsdon C, Fotheringham AS, Charlton M, 1999, "Some notes on parametric significance tests for geographically weighted regression" *Journal of Regional Science* **39** (3) 497 - 524.

Brunsdon C, Fotheringham AS, Charlton ME, 1996, "Geographically weighted regression: A method for exploring spatial nonstationarity" *Geographical Analysis* **28** (4) 281 - 298.

Cleveland WS, 1979, "Robust Locally Weighted Regression and Smoothing of Scatterplots" *Journal of the American Statistical Association* **74** 823 - 836.

McNally M, 2000 "The Activity-based Approach" in *Handbook of Transport Modelling* Eds DA Hensher and KJ Button (Pergamon, Oxford) pp. 53 - 68.

Meyer MD, Miller EJ, 2001, *Urban Transportation Planning, Second Edition* (McGraw-Hill, New York)

Miller HJ, 1999, "Potential contributions of spatial analysis to geographic information systems for transportation (GIS-T)" *Geographical Analysis* **31** (4) 373 - 399.

Miller HJ, Shaw SL, 2001, *Geographic Information Systems for Transportation: Principles and Applications* (Oxford University Press, Oxford)

Newbold KB, Scott DM, Spinney JEL, Kanaroglou P, Páez A, 2005, "Travel behavior within Canada's older population: a cohort analysis" *Journal of Transport Geograhy* **(in press)**

Ortúzar JD, Willumsen LG, 2001, *Modelling Transport, Third Edition* (Wiley, New York)

Páez A, Scott DM, 2004, "Spatial Statistics for Urban Analysis: A Review of Techniques with Examples" *GeoJournal* **61** (4) 53 - 67.

Pisarski AE, 2003, "Prescriptions for research: reviewing the history of TRB's critical issues in Transportation" *TR News* **226** 30 - 35.

Rosenbloom S, 2001, "Sustainability and automobility among the elderly: An international assessment" *Transportation* **28** (4) 375 - 408.

Washington S, 2000, "Iteratively specified tree-based regression: Theory and trip generation example" *Journal of Transportation Engineering-ASCE* **126** (6) 482 - 491.

A GIS-Based Decision Support Tool to Study the Impact of Elderly Population on the Transportation System in Hamilton Census Metropolitan Area (CMA)

Hanna MAOH, Pavlos KANAROGLOU, Darren SCOTT, Bruce NEWBOLD and Antonio PAEZ

Center for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University, 1280 Main Street West, Hamilton, ON, Canada, L8S 4K1

1. INTRODUCTION

Over the past five decades, most North American cities became progressively more dependent on the automobile. This was accompanied with massive capital investments in road infrastructure. Urban transport planners and practitioners sought tools to conduct cost-benefit analysis of such infrastructure investment projects. The trip-based Urban Transportation Modeling System (UTMS) was adopted as the state of practice tool by most metropolitan planning organizations (Meyer and Miller, 2001). UTMS has been used to project future traffic and evaluate the need for new road capacity, as well as changes in transit service and land use patterns in cities around the globe. It has done so by focusing on peak-period, work related travel.

UTMS has normally ignored the contribution of older retired drivers on the urban transportation system. However, Canada's population is aging, which means that in coming decades there will be a significant increase in both the number and percentage of older drivers in urban areas. Statistics Canada (2002) estimates that by 2021 one in five Canadians will be at least 65 years old. Recent research for the United States, Australia and some European countries provides evidence suggesting that as people age they become more

dependent on the automobile (Rosenbloom, 2001). Despite these observations, very little has been done to address the issue in the Canadian context.

This paper reports on the effort to develop a GIS-based decision support tool, to study the impact of older population on the transportation system in Hamilton Census Metropolitan Area (CMA) (Figure 1). The significance of our work is twofold: (1) unlike most of the existing four-stage travel demand UTMS; the foundation of our system is a multiregional demographic model that predicts the spatial distribution of population by sex and age (Rogers, 1995), (2) the system is developed as a stand-alone GIS-based decision support tool that is based on the MapObjects spatial software components to enable interactive specification of demographic and travel demand scenarios via map and spreadsheet scenario dialogue boxes.

The remainder of the paper is organized as follows. Section two starts with a background to provide an overview of the basic elements comprising a travel demand UTMS. It also sheds light on the significant role that Geographic Information System (GIS) can offer to enhance the performance of these urban transportation-planning tools. Section three describes the general structure of the devised decision support tool and discusses in detail the different components used in its development. Section four discusses the development environment of the system. The last section provides a conclusion.

2. BACKGROUND

Travel demand modeling continues to be one of the most important aspects of the urban transport planning process. The common practice for travel demand modeling is to use an Urban Transport Modeling System to simulate travel behavior in the city's transport network. UTMS are conceived as predictive models since their purpose is to predict the possible outcomes of alternative policies. Early versions of such models appeared in the mid- 1950s and their use became commonplace by many metropolitan areas around the world by the 1960s (Southworth, 1995). Typically, the UTMS, known also as the four-stage model, consists of four sub-

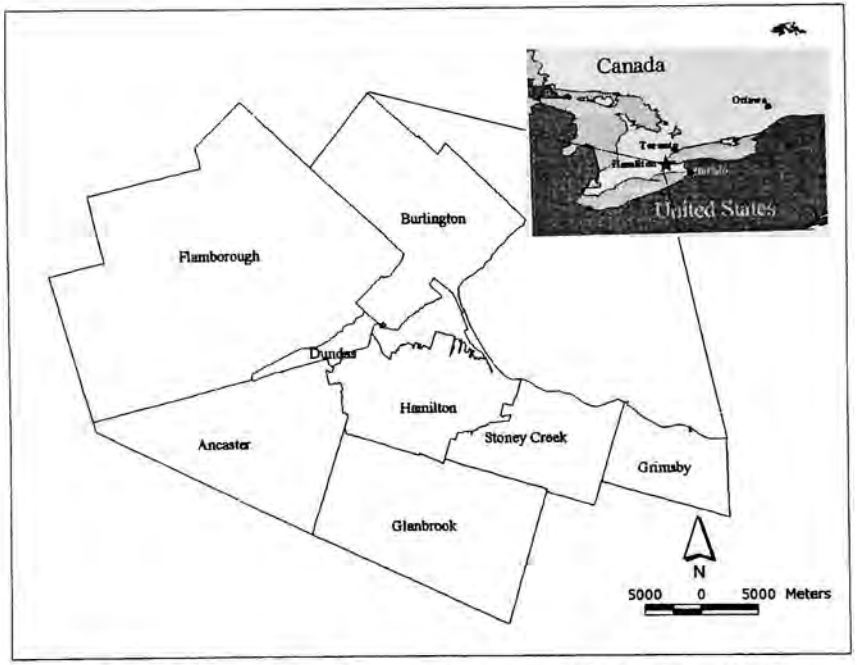


Figure 1: The Hamilton Census Metropolitan Area (CMA) in the regional context

models that are executed sequentially. These are: trip generation, trip distribution, modal split and traffic assignment. The model starts with an exogenous distribution of land uses, population, employment and other economic activities for a particular base year. Therefore, the UTMS makes the assumption that land use planning provides the impetus for transportation infrastructure development. This one-way relationship assumption between the land use and transportation systems is one of the main drawbacks in most of the existing UTMS. There is sufficient evidence in the literature on the two-way relationship between land use and transportation (Wegener, 2004; Kanaroglou and Scott, 2001). Despite this limitation, UTMS remains the most widely used travel demand modeling tool.

Within an UTMS, the metropolitan area is represented with two layers of exogenous information: (1) a set of mutually exclusive zones known as Traffic Analysis Zones (TAZs), and (2) a set of connected links to represent the city's transportation network. The trip generation submodel utilize the exogenous land use information to determine the number of trips originating from each TAZ as well as the numbers of specific purpose of trips destined to each TAZ. Trip generation models are typically based on either regression models or categorical data analysis (Kanaroglou and Scott, 2001). The trip distribution submodel uses the information produced by the trip generation submodel to generate a set of origin-destination trip matrices by trip purpose and time of day. Gravity based spatial interaction models or multinomial logit models are usually used to formulate trip distribution models. Once the trip distribution submodel is executed, the modal split submodel accepts the purpose specific origin-destination matrices and split them by travel mode (motorized, passenger, transit, walking, etc.). The most popular econometric models used to achieve a modal split are the multinomial logit and nest logit models.

Once the modal split submodel is executed, all the motorized origin-destination trip matrices that relate to the same time of day are added up. The resulting matrix serves as input to the traffic assignment sub-model. The objective at this stage is to generate an origin-destination travel-times matrix in a congested situation on the network. Several algorithms have been developed for this purpose,

including the all-or-nothing, user equilibrium and stochastic user equilibrium algorithms (Sheffi, 1985). The whole process starts with an initial origin-destination travel-times matrix that corresponds to a free flow situation. After the first pass through all the sub-models, a new set of inter-zonal travel times is estimated. Then, the trip distribution, modal split and traffic assignment sub-models are run again with this new set of inter-zonal travel time values. The iterations are repeated until the travel costs stabilize. The outcome is the number of total motorized trips on the links of the city's transport network. These trips can then be translated into a system performance indicator to reflect the congestion level on the road network. This could be achieved by comparing the assigned motorized trips on a particular link to that links' design capacity.

While the UTMS is based on the TAZ and network layers that are spatial in nature, early development in UTMS was based on non-spatial computer programs. However, recent advancements during the 1990s in computing and Geographic Information Systems (GIS) have revolutionized the development of decision support tools to study and simulate urban transportation systems. Processing speed and storage capabilities has increased at an exponential rate in the past decade. At the same time, new ideas in software engineering, such as object-oriented languages, have facilitated the maintenance and incremental development of simulation programs. The widespread use of GIS made it easier to organize and display large volumes of spatial data. GIS has also provided an excellent interface to planners to experiment and simulate scenarios without the need to be trained in modeling and programming.

Despite the advances made in GIS during the past decade and a half, a limited number of applications have emerged, which demonstrate the potential of GIS as part of the urban modeling framework. The majority of these applications adopted an encompassing framework as oppose to a modular one. The encompassing framework is based on extending the functionality of the core GIS via scripting or macro languages that are native to the GIS software. This, however, requires the UTMS to be dependent on an existing large GIS-T platform (e.g: TRANSCAD or EMME/2) even if it is only using some of the functionalities of that GIS-T. On

the other hand, a modular framework leads to a stand-alone product since within this framework a set of systems are linked together to fulfill the objective of utilizing specific functions such as mapping to visualize spatial information and graphical user interface to define input to the system and summarize output after the simulations.

An example of the later is the utilization of the MapObjects spatial software components (ESRI, 1996) within the Visual C++ programming language to develop the IMULATE integrated urban land use and transportation model (Buliung et al, 2005). IMULATE has been developed since 1994 to study urban problems that relate to land use and transportation in the Hamilton CMA (Anderson et al., 1996; Scott et al., 1997; Kanaroglou and South, 1999). While each of the two frameworks has its own advantages and disadvantages, a major benefit of adopting an encompassing framework is to avoid carrying around the overhead of a large GIS-T (Buliung et al., 2005). Consequently, the development of our UTMS model will be based on a modular framework similar to the one used to develop IMULATE. Such initiative is deemed innovative and cost-effective, since it will re-use many of the GIS-T functionalities already develop and built in IMULATE. These include mapping functionalities, interactive scenario dialogue boxes, shortest path and traffic assignment routines and graphical output summaries.

3. MODEL ARCHITECTURE

The developed decision support tool aims at quantifying over time the likely impacts of an expanding older population coupled with its increased automobility on the Hamilton's CMA urban environment. At the heart of the devised tool is an UTMS that will allow us to analyze and model the travel behavior of older Canadians residing in the Hamilton CMA. Unlike most of the existing UTMS, the foundation of our system is a demographic model that predicts the spatial distribution of population by age and sex. Population for the 1996 base-year is being evolved in five-year increments to evaluate the contribution of the elderly cohort to travel demand, traffic congestion and transport related pollution levels in the Hamilton CMA.

Figure 2 presents the architecture of the model, which consists of the following three interlinked modules: (1) demographic, (2) transportation and (3) mobile emission. The two layers representing the CMA consist of a TAZ layer that is based on the 1996 census tract divisions and a road network layer that was modified from the IMULATE network to match the 1996 tract divisions. Consequently the devised system has 163 TAZ, 1542 links and 1183 nodes.

The demographic module consists of two sub-modules: (1) Rogers' Multi-regional Demographic Model (RMDM) (Rogers, 1995) and (2) Spatial Aggregate Multinomial Logit Model (SAMNL) (Kanaroglou and Ferguson, 1996; Ferguson and Kanaroglou, 1997). The sequence of operation starts by executing RMDM to simulate the change in the age pyramid of the CMA's municipal population. This change is modeled as a function of the inter-municipal migration and death rates by age and sex as well as the fertility rates of Canadian females. Next, SAMNL model is executed to allocate the predicted municipal population to the census tracts in the CMA. The allocation procedure makes use of a logit-based formula, which predicts the probability of choosing zone i from the set of zones of a given municipality M, by decision maker n who moved to M. The utility of the logit model reflects individual characteristics, such as age and gender, and zone attributes, like the level of accessibility in the zone, the number of new dwellings and the distribution of land uses in the zone. The accessibility measure will vary over time and is based on input from the transportation module for the consecutive simulations as shown in Figure 2.

The transportation module is based on the four-stage model discussed in the previous section. However, to achieve our objective, we model work and non-work trips that are made by all the population, the economically active population (age 15 – 64) and non-economically active population (age 65+) during a typical day. Accordingly, a number of trip generation, distribution and modal split models by purpose for the different cohorts were estimated for the morning (6am – 9am), day (9am – 4pm), afternoon (4pm – 7pm) and night (7pm – 6 am) periods.

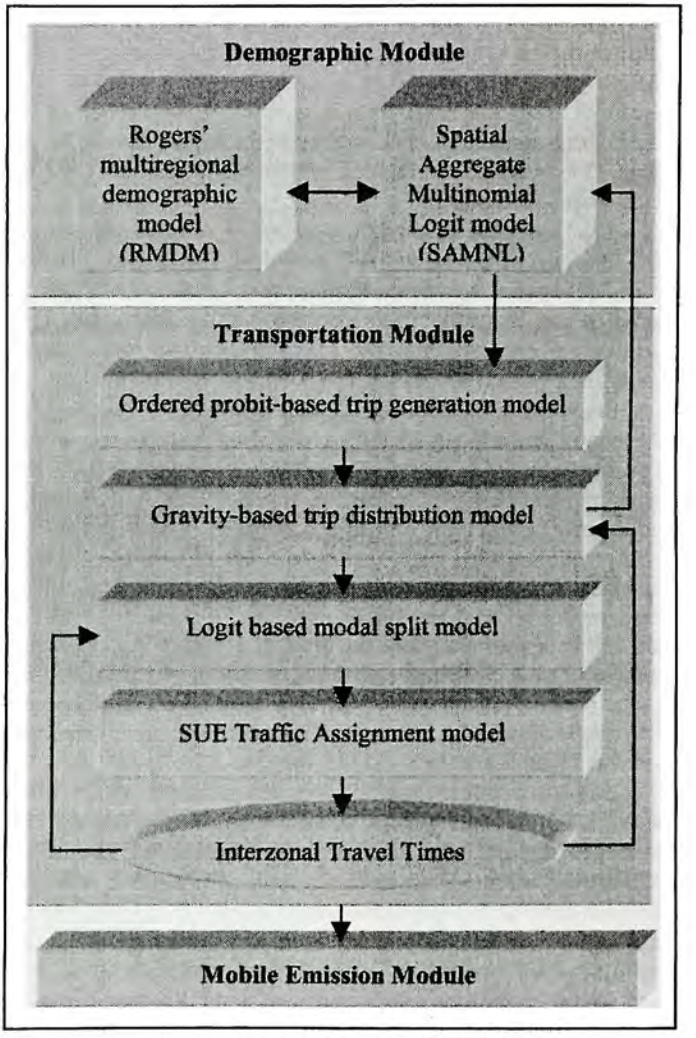


Figure 2: Model General Structure

Generated trips are predicted with ordered probit models, which appeared to be more theoretically consistent than regression models in our case. The generation submodule accepts the information produced by the demographic module and employs it to predict the generated trips. The trip distribution submodule is based on production constraint gravity models, which predict work and non-work origin-destination (OD) trip matrices for the modeled cohorts in the four time periods. Trips that originate in a certain census tract, as provided by the trip generation models, are allocated to other census tracts in proportion to the inverse of a function of impedance cost that separates the origin from its destinations. In our setting, we use interzonal congested travel time to reflect the separation cost between the different census tract pairs.

Initially, the model starts with a free flow interzonal travel times matrix to generate OD trip matrices. The modal split submodule accepts the trip distribution information and splits it by mode to identify motorized trips. Those are assigned to the network via the traffic assignment routine, using a free flow travel times matrix to estimate a new set of travel times in a congested situation. Then, the trip distribution, modal split and traffic assignment sub-modules are run again with this new set of inter-zonal travel time values. The iterations are repeated until the travel costs stabilize. The result is link flows for all the population as well as the economically active drivers. The difference between those two measures is the contribution of elderly drivers to the total flow on the network.

The mobile emission module uses link flows and average speeds from the traffic assignment submodule to provide estimates for CO, HC and NOx pollutants on each link. Average link speed is calculated from the link congested travel time that the traffic assignment produces. Link emissions are estimated after matching the MOBILE5.C emission factors to estimates of link average speed and adjusting by link volumes.

4. DEVELOPMENT ENVIRONMENT

The development of the devised system follows a modular approach that integrates components from different technologies as

shown in Figure 3. The demographic, transportation (excluding the traffic assignment) and mobile emission modules are programmed in the GAUSS language. The choice of this computing technology facilitates the programming of the system since most of the mathematical operations in the systems' models involve matrix algebra. On the other hand, the traffic assignment and shortest path submodules were programmed in C++ as oppose to GAUSS, in order to achieve higher computational efficiency. Currently, the Stochastic User Equilibrium (SUE) traffic assignment routine converges after 11 iterations in approximately 77 seconds when applied to the Hamilton transportation network. This level of performance can be achieved on a workstation with a Xeon 2.0GHz processor.

The developed platform is supported with windows Graphical User Interface (GUI) that is based on the MapObjects spatial software components and Visual C++. MapObjects are software components that enable developers to incorporate mapping or spatial data handling capabilities into applications (ESRI, 1996). Therefore, the user can manipulate, manage and analyze spatial and non-spatial information using familiar GIS and windows controls through the application. Visual C++ provides a set of GUI dialogue boxes and enables the user to devise a certain scenario interactively via maps and spreadsheets. The program creates a set of internal rules to reflect the components of the defined scenario. Once a scenario is devised, the Visual C++ program applies the rules to the systems' database and stores the information in a format acceptable by the GAUSS and C++ routines.

When the user executes a scenario, Visual C++ communicates with the GAUSS code through the GAUSS engine (see Figure 3). The GAUSS Engine is a dynamic library that enables the Visual C++ application to communicate with the GAUSS codes of the different submodules. On the other hand, the traffic assignment and shortest path submodules are executed directly depending on the sequence of operation described in Figure 2. When the system converges, all the output information is updated in the system's database.

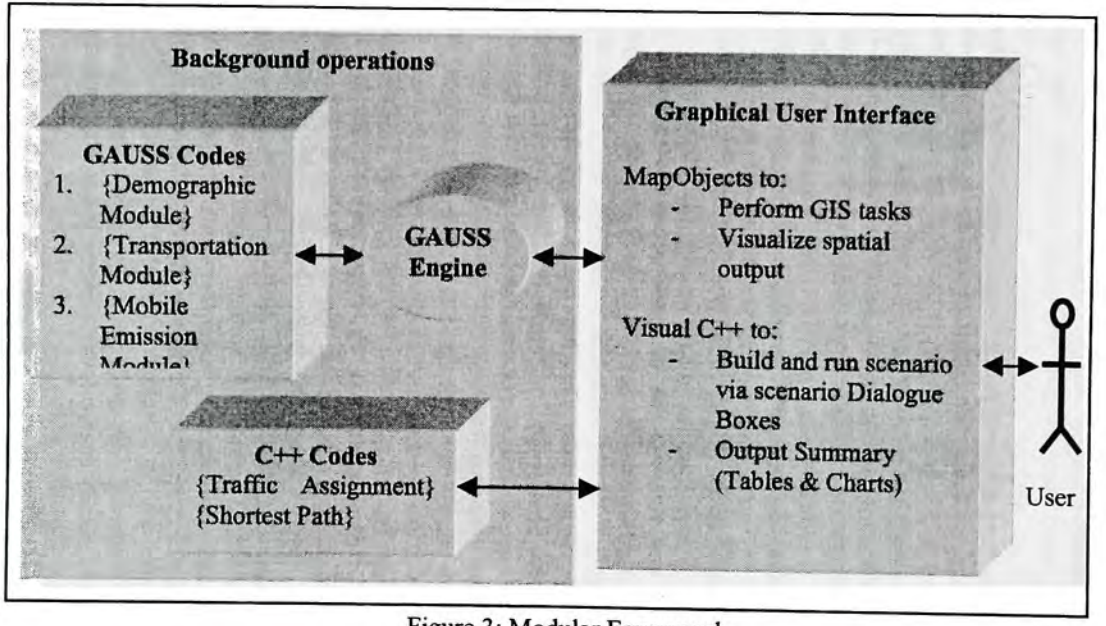


Figure 3: Modular Framework

The output from simulating a scenario is accessible though an intuitive GIS-based or map-driven GUI. The interface consists of a number of standard windows and additional controls that are commonly found in desktop GIS software like ArcView3.3 or ArcGIS 9.0. Users are able to view, query and manipulate spatial and attribute data interactively via an onscreen map. The system provides summary statistics in charts and tables for the most critical indicators by cohort. Following is a list of some of these informing indicators to name few: population pyramids, vehicle kilometers traveled (VKT), vehicles minutes traveled (VMT), congested and free flow link and vehicle average speeds (km/h), city wide emissions (grams), Auto-trip shares, Modal split shares, etc.

5. CONCLUSION

This paper reported on the work we are conducting to develop a GIS-based decision support tool (DST) to study the impact of elderly population on the transportation system in the Hamilton CMA. In recent years, advancements made in geomatics have revolutionized the development of spatial software. For instance, tools like MapObjects allow the development of efficient spatially oriented decision support tools, by offering a wide range of functionalities to manipulate and analyze spatial data without the need to rely on a commercial GIS platform.

An innovative and cost-effective approach has been adopted to develop the DST. We re-used many of the functionalities already built in the IMULATE urban land use and transportation model (Buliung et al., 2005). The result is a simulation model with an intuitive GIS-based graphical user interface that is built using MapObjects and Visual C++.

The developed tool will help us to conduct research that will advance our knowledge on the likely impacts of an expanding older population coupled with its increased automobility on the Hamilton's CMA urban environment. The system also serves as a prototype that can be followed in studying similar demographic and transportation processes in other Canadian cities.

REFERENCE

Anderson, W., Kanaroglou, P. Miller, E. and Buliung, R.. (1996) Simulating automobile emissions in an integrated urban model, Transportation Research Record, 1520: 71-80.

Buliung, R., Kanaroglou, P. and Maoh, H. (2005) "GIS, Objects and Integrated Urban Models" In: The Behavioral Foundations of Integrated Land use and Transportation Models, eds. M. Lee-Gosselin and S. Doherty, Amsterdam, Elsevier.

ESRI (1996). Building Applications with MapObjects, Environmental Systems Research Institute, Redlands CA.

Ferguson, M. and Kanaroglou, P. (1997) An empirical evaluation of the aggregated spatial choice model, Internation al Regional Science Review 20(1&2): 53–65.

Kanaroglou, P. and Ferguson, M. (1996) Discrete spatial choice models for aggregate destinations, Journal of Regional Science 36: 271–290.

Kanaroglou, P. and Scott, D. (2001) Integrated urban transportation and land-use models for policy analysis, in M. Dijst (eds.) Governing Cities on the Move: Functional and Management Perspectives on Transformations of Urban Infrastructure in European Agglomerations., Ashgate, pp. 43-75.

Kanaroglou, P. and South, R. (1999) Can urban form affect transportation energy use and emissions? An analysis of potential growth patterns for the Hamilton Census Metropolitan Area, Energy Study Reviews, 9: 22-40.

Meyer, M. and Miller, E. (2001) Urban Transportation Planning: A Decision-Oriented Approach, Second Edition. New York: McGraw-Hill.

Rogers, A. (1995) Multiregional Demography: Principles, Methods and Extensions. New York: Wiley.

Rosenbloom, S. (2001) Sustainability and automobility among the elderly: An international assessment. Transportation, 28: 375-408.

Scott, D., Kanaroglou, P. and Anderson, W. (1997) Impacts of commuting efficiency on congestion and emissions: a case of the Hamilton CMA, Canada, Transportation Research Part D, 2: 245-257.

Sheffi, Y. (1985) Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods. Prentice-Hall Inc., New Jersey.

Southworth, F. (1995) A technical review of urban land use-transportation models as tools for evaluating vehicle travel reduction strategies, ORNL-6881, The Office of Environmental Analysis and Sustainable Development, U. S. Department of Energy.

Statistics Canada (2001) General Social Survey: An Overview. Catalogue Number: 89F0115XIE. Statistics Canada (2002)

Wegener, M. (2004) Overview of land use transport models, in D. Hensher, K. Button, K. Haynes and P. Stopher (eds.) Handbook of transport geography and spatial systems, Amsterdam: Elsevier, pp. 127-146.

Air Transportation - *Are New Destinations On The Horizon?*

Joseph Monteiro and David Krause*

Introduction

On October 29, 2004, Transport Minister, Jean-C. Lapierre at the *Canadian Airline Investment Conference* said "The traditional industry model - closed domestic markets, and strong national flag carriers, and bilateral agreements - is giving way to a new model. ...the message I heard loud and clear was that the industry was ready for change. ... I believe we have a great opportunity now to create an air policy that will serve Canada well for the next ten years."[1] He also made a statement before the Standing Committee on Transport and presented a reference document on Air Liberalization.[2] Do these recent statements signal the dawning of a new era of air transportation policy for Canada?

In this paper, we review what we believe are the new destinations for air transportation in the new future. We consider respectively domestic air services, transborder air services and international air services. We then present our concluding remarks.

I. Domestic Air Services

During the mid 1980s, regulatory changes gradually evolved in response to the deregulatory move. Liberalization in domestic services was introduced with the Ministerial statement on the 'New Canadian Air Policy' on May 10, 1984. The *National Transportation Act*, 1987 turned de facto deregulation introduced in 1984 to official regulatory reform on January 1, 1988. The Act removed most of domestic economic regulation (i.e., entry, pricing, carrier roles, licence restrictions on frequencies, routes and classes of aircraft) in Southern Canada while retaining it for Northern Canada and for the rest of the world.

The market entry test of 'public convenience and necessity' was replaced by the test of 'fit, willing and able' for licence applications in

---

* *The views expressed here are those of the authors and are not purported to be those of the Commissioner or the Competition Bureau, Industry Canada. The authors acknowledges the help of Sofia Civettini.*

Southern Canada and the advance notice required for exit was 120 days. Further, domestic tariffs had no longer to be filed, however they had to be published. Finally, there was limited control over fares on monopoly routes in the South or routes in the North. In 1996, the restraints on services in the North were removed.[3]

Despite the fact that the industry was nearly fully deregulated, a number of issues remained regarding investment and service rights for foreign carriers. These issues, partly eclipsed by the severe downturn in the industry, were raised by Jean-C. Lapierre in his statement before the Standing Committee on Transport on domestic services: Can we relax ownership restrictions? What about control? Can we allow foreign carriers greater options when providing services between points in Canada? What would be the effects? The findings of a few important reports on these and other issues are considered below.

a) Relaxing regulatory restrictions on Canadian ownership and Control of Canadian Airlines

The National Transportation Act Review Committee (NTARC) in 1993 considered relaxing foreign ownership as one proposal worthy of future consideration in the event of a monopoly developing in the Canadian airline market.[4] Under this proposal, Canada would raise the 25% ceiling on foreign equity participation in a Canadian airline to 49%. It would provide all Canadian airlines, existing and potential, with greater access to capital thereby increasing the possibility of competition.[5]

A few years later in 1999, the House of Commons Standing Committee recommended increasing foreign ownership limits to 49%.[6] The Senate Committee on Transport and Communications in the same year recommended increasing foreign ownership limits to 49% on a selective basis in order to facilitate the access to capital by small carriers competing with the dominant carrier.[7] The Government in its reply to the two reports decided that it would not be beneficial raising foreign ownership limit at this time.[8]

Thereafter in 2001, the Canada Transportation Act Review Panel

(i.e., Panel) recommended relaxing ownership restrictions by raising the limits from 25% to 49%. It stated "Several submissions to the Panel proposed relaxing the current rules restricting the percentage of voting shares in a Canadian airline that foreigners can hold. The Panel agrees that the 25% limit should be raised. This would not guarantee availability of foreign capital, but it would facilitate access to foreign funds."[9]

More recently (i.e., 2003), one study commenting on this issue stated: "A relaxation of the foreign-ownership limits might do more to ensure the survival of domestic competition than any other policy changes. This could increase the sources of capital available to Canadian carriers and, for the well-managed ones, access to more sources of capital should also reduce the costs of capital."[10]

A broader issue is providing control to non-Canadians. While it could be pointed out that equitable treatment with the unregulated sectors of the economy may call for it, the Transport Canada Report *StraightAhead* in its vision for transportation in Canada states "This would be a very substantial step, and one for which there are few international precedents."[11] Opinion polls continue to show that they prefer the domestic airline industry to remain under Canadian ownership. Further, most countries continue to require domestic control of airlines.

b) Canada-only Carriers

The NTARC considered Foreign Carrier Cabotage or Right of Establishment as the second proposal worthy of future consideration in the event of a monopoly. Under this proposal, foreign airlines would be permitted to carry passengers between Canadian points, or to establish stand-alone, intra-Canada operations. The latter could be operated either by the airline itself, a separate division, or a subsidiary. "Because the permanent establishment of foreign carrier operations or routes could adversely affect Canadian operators and investors, this option should be exercised with great care."[12]

Subsequently, the House of Commons Standing Committee indicated that regarding the viability of licencing Canada-only carriers

to operate on Canadian domestic routes, the Committee was of the opinion that it has merit as a means of enhancing competition and recommended that the government examine its viability. Regarding Canada-only Carriers, the Government in its reply indicated that it is committed to achieving a healthy, Canadian-controlled airline industry. Its goal is to foster competition from Canadian carriers.

In 2001, the Panel recommended that Canada negotiate the rights of establishment for foreign-owned domestic carriers **if** negotiations on a North American Common Aviation Area do not succeed.[13]

c) Modified 6th Freedom Rights

The House of Commons Standing Committee in 1999 recommended modification of the 6th freedom rights to permit competition via an intermediate stop (eg. Montreal-O'Hare-Vancouver) by beginning negotiation with the United States but believed that achieving this goal was likely to require significant price reductions.[14]

This was also recommended by the Senate Committee. It stated "In order to increase competition under a dominant owner/carrier regime, the government take steps to negotiate with the United States the necessary authority to allow through ticketing from one point in Canada to another through a US intermediate stop."[15]

Later, the Panel recommended that Canada negotiate with other countries for reciprocal granting of modified sixth freedom rights **if** negotiations on a North American Common Aviation Area do not succeed.

d) Cabotage (foreign carriers)

According to one study "Granting full rights of cabotage to all foreign carriers would be a very big step to take unilaterally and is not one we would advocate at this time. Part of the reason for caution is the tremendous uncertainty regarding what would happen if Canada did go this route. Cabotage, at least outside the current European arrangement, has been so far from the agenda that airlines have not really thought about what they would do if it became possible. That said, we can be confident that the foreign carriers would take a long and careful look at

the Canadian market if we moved to permit cabotage."[16] This study predicts that foreign carriers would begin entry by the tag-end cabotage variety if profitable and most service from foreign carriers would be to and from major Canadian centres by larger carriers. This study supports certain types of cabotage such as modified 6th freedom but believes that providing cabotage to Europeans and not to Americans does not make a great deal of sense. Other types of cabotage such as franchise cabotage and route specific cabotage are not without their complications.

The former Minister of Transport Canada, David Collenette, indicated that he does not support the provision of cabotage rights to foreign airlines. He stated "Cabotage ... is not the answer for competition in this country."[17] In 2003, commenting on increased participation by foreign carriers in domestic markets *StraightAhead* states "In the government's view, however, now is not the appropriate time to undertake this step."[18]

The views or recommendations on some of the above issues in various reports were premised to a large extent on the concern that the restructuring would result in a monopoly carrier. That concern has largely disappeared. The matter has now to be examined against the following background: the overall market share of Air Canada is 54.9% as of October 2004,[19] low-cost carriers have made tremendous inroads into the domestic market over the last few years; the need to encourage domestic competition does not appear to be as pressing today as competition appears to be developing on its own without government intervention, and the industry is just beginning to recover from its crisis. In addition, the new entrant and small carriers continue to need access to foreign capital and Canadian carriers want to expand their markets into the U.S.

As a first step, based on the findings of the majority of the above reports, the government should consider relaxing regulatory restriction on foreign ownership by raising the limits to 49% and should proceed with negotiating with the United States for modification of 6th freedom rights. Providing for Canada-only carriers or cabotage rights are interesting proposals that the government could consider reciprocally if

other factors are conducive.

II. Transborder Air Services

On February 24, 1995, a new historic aviation accord was heralded. It was a major breakthrough in bilateral relations and a first for the two nations in liberalizing air travel. It signalled the dawning of a new era in bilateral negotiations to the world.

The two important features of the Agreement were: first, removal of restrictions on air travel; and second, pricing freedom. Under the first feature: Canadian and US airlines were able to provide unrestricted cross-border services (i.e., without conditions on capacity, frequency and size of carrier) from any point in Canada to any point in the US (Annex I).[20] In addition, all restrictions regarding aircraft size and package size on all US all-cargo service licences were removed and carriers were prohibited from providing all-cargo service with aircraft exceeding 35, 000 pounds to more than one destination in the other country for courier companies. Under the second feature: the Canada-US fare approval process was liberalized (Article 5). Air carriers would be allowed to set their own fares on transborder services and it would no longer have to be filed with the regulatory agencies in the two countries. Fares could be disallowed subject to a dual process only if agreed to by the countries' authorities for reasons of predatory prices, unreasonably discriminatory prices, unreasonably high prices due to abuse of dominance and artificially low prices because of government subsidy.[21]

The Agreement did not provide for total freedom and a decade after signing this Agreement, questions regarding further liberalization have risen again. The Minister of Transport in his statement before the Standing Committee on Transport raised a number questions on trans-border services. Should we consider: allowing trans-border courier services to two or more points; allowing trans-border services to third countries and allowing price competition for these services; allowing U.S. carriers to fly between points in Canada; integrating air markets under a single set of rules; and following uniform or different approaches to passenger and all cargo services. Before considering

liberalizing the Agreement, we should briefly consider whether the Open Skies Agreement has been beneficial.

An examination of the statistical data reveals that the number of new services and the number of carriers serving more than one airport increased between 1995 and 2000 and have resulted in staggering increases in the volume of passengers travelling on transborder routes particularly in the first two years after the signing of the Open Skies Agreement. In addition, the volume and value of transborder air cargo increased sharply between this period. Concentration rose from January 1995 to January 1998 in all transborder markets, though in certain markets it declined. The initial impact of the Open Skies Agreement was to lower prices, however, after mid 1996 prices started rising.

Other studies, in general, also support our findings and make a number of interesting observations.[22] Even Transport Canada recently (i.e., 2003) acknowledged that "There is no debate about the benefits of the 1995 air agreement with the United States for all stakeholders."[23] This holds, notwithstanding, the events since September 2001.[24] The findings of some reports on the issues raised by transborder service are now discussed.

a) Co-terminalization of cargo/courier service

A publication by an executive in the air courier business indicates that "there is an opportunity to build on the 1995 agreement and move closer to a true "open skies" environment by opening the market for cargo. To do so requires several additional steps. The first is co-terminalization; and second is cabotage. ... The agreement allows for designated national cargo carriers to fly only to single city-pairs between U.S. and Canada. This restriction adds to the cost of cargo transport in both countries."[25] "Prohibiting co-terminalization places regulatory constraints on North American air carriers and inhibits the flexibility required to provide reliable and competitive service. The major losers are the users of cargo networks. To maximize global trade opportunities for North American companies, it is imperative that all non-market barriers to trade are removed. Co-terminalization is currently one of the most significant non-market barriers to trade within

the air cargo marketplace." [26]

b) Seventh Freedom Rights for all cargo, scheduled and charter services

Providing unilateral 7th freedom rights (Canada to other countries) to standalone companies for all cargo, scheduled and charter services unilaterally would be beneficial to the air cargo industry. For example, the transhipment program at Winnipeg International Airport allows foreign carriers to drop off cargo destined for a third country, which another carrier can then pick up for the final leg of the journey. The transhipment designation is only available to Winnipeg, Hamilton, Gander, Windsor and Mirabel and is not available for larger airports.

On the provision of eight freedom rights (i.e., carrying cargo from NY to LA on a flight originating from Toronto) or modified 6th freedom rights (consolidating cargo originating in Canada with cargo originating in the U.S. through a U.S. hub for delivery back to a Canadian destination) the previous cited publication states: "The development and maintenance of separate air cargo networks in Canada and in the United States is necessitated by the current Air Transport Agreement. The potential exists to fully maximize economic efficiency of the existing air networks by giving air cargo operators authority to fly wherever and whenever the companies who use their services demand. It is clearly inefficient to require Canadian shippers to use Canadian carriers and U.S. shippers to use U.S. carriers when the market requires otherwise.

By treating all cargo equally, a new Canada-U.S. air transport agreement could facilitate the expansion of transborder trade. It would reduce the supply chain costs and reduce time to market for many industries. Further, it would expand opportunity for air cargo operators to identify and serve market niches - unconstrained by geographic and national identity. ... The net benefit of liberalized air cargo arrangements to Canadian business, Canadian airport operators and the government is real. Restricting economic growth and trading opportunities for Canadian business makes little sense. Liberalizing air cargo would increase international trade opportunities for Canadian stakeholders."[27]

It is also worthwhile observing that Mr. Collenette's successor,

Tony Valeri, has indicated he would like to negotiate the possible easing of restrictions on cargo flights into Canada. Both Mr. Valeri and Mr. Paul Cellucci, US Ambassador to Canada, indicated that "We think it's probably a good time to start discussing things like the cargo issue, as well as things like the cabotage and other issues involving open skies."[28] In addition, "The Airports Council International is pushing for a separation of cargo from passenger traffic rights and endorses open skies for freighters everywhere. Liberalization worldwide is inevitable. Its not a question if it will happen but how, ..."[29] Adopting a dual approach to passenger and all cargo service has merits. Such an approach could lead to a quicker liberalization of all cargo services relative to passenger services.

c) Fifth Freedom Rights (with traffic rights and full price flexibility)

Seeking additional fifth-freedom rights will not only make a route from Canada more viable for a Canadian carrier, but could also provide benefits to travelers from the US or provide service where there is no service from the US. A better utilization of capacity could also result in lower fares to Canadian travelers. It would therefore be to the benefit of Canada to seek removal of existing restrictions bilaterally though pricing issues will have to be considered under the general application of the competition laws. However, it is unlikely that the U.S. will grant complete freedom.

d) Ninth Freedom Rights (i.e., stand alone cabotage in Canada)

The House of Commons Standing Committee on reciprocal cabotage stated "we think reciprocal cabotage would open up the market to more domestic competition, while at the same time allow Canada's air carriers to expand their markets in the United States."[30] It therefore recommended initiating negotiations with the United States but that no agreement should be concluded without a review of the agreement with specific bodies. The Government in its reply rejected the recommendation and indicated that no talks with the United States were then scheduled.[31]

e) A Common North American Aviation Market.

The Panel's most provocative recommendation involved creating a

North American Common Aviation Area with carriers from Canada, U.S. and Mexico in which carriers from these countries would compete freely.[32] Commenting on this issue of the Panel, a former Minister of Transport, Mr. Collenette said that "he didn't think a North American free trade zone for aviation would fly. The likelihood of that is slim. If we're talking about an agreement where there's going to be cabotage, the Americans have shown no appetite for that,.."[33] The amendments introduced in Bill C-26 also did not provide for it. Louis Gialloreto from McGill University commenting on the recent proposal by Jean-C. Lapierre said "If you have a single aviation zone between Canada and the U.S., does that mean we eliminate security, immigration and customs? I don't think the U.S. guys are going to say yes to that, ... given the post-9/11 world, these are not small items."[34]

Based on recent developments and findings of the majority of the studies reviewed, as a first step, the government should consider removing the restrictions on co-terminalization of courier cargo and seek removal of restrictions on fifth freedom rights and provide for 8th freedom rights, modified 6th freedom rights for cargo and 7th freedom rights to aircargo carriers to all transshipment aircargo areas. Providing for ninth freedom rights for the U.S. (passenger cabotage) and a common North American aviation market are interesting proposals that the government could consider bilaterally if other factors are conducive.

III. International Air Services

In December 1994, Canada's new policy for international air transportation was set up. This policy was designed: to make the best use of Canada's international route rights through a 'use it or lose it' approach; to facilitate access to Canada for foreign carriers; and to provide consumers with more travel options and to improve protection for their travel arrangements. In March 1995, criteria for designating a second carrier for international routes were announced.

A few years later, in early 2000, after consultation with various stakeholders, a new policy on April 4, 2000 for international charter passenger air services was announced. "This policy removes fences surrounding international passenger charter services, such as advance

booking and minimum-stay requirements, and restrictions on one-way travel."[35] This policy is expected to enhance international travel options for Canadians, encourage competition, innovation and growth in the charter air industry. In May 2002, Canada announced its multiple designation air policy permitting all carriers to apply to provide international service regardless of size of one way passenger markets.

The present Transport Minister in his statement before the Standing Committee on Transport raised a number of questions with regard to international services : should we further relax or remove restrictions on services between our countries?; should we extend that relaxation to third countries?; and how do we balance all the Canadian interests affected by such changes? He indicated that international services have the greatest number of restrictions and offer the greatest scope for reform. The data of international services will first be reviewed briefly.

The statistical data on international sector reveal that goods carried by Canadian carriers on international routes increased 43.6% over the period 1993-2002 (i.e., twice the increase in the domestic sector); and the operating goods revenue on international routes increased 74.3% (i.e., nine times the increase in the domestic sector) for the same period. Passengers on international routes increased 51.9% compared to 33.3% for domestic passengers for the period 1993-2000. The average international air fare was $359 in 2000, 6.2% below the record average international fare of $382 in 1994. In other words, international traffic outpaced domestic traffic and average international fares declined.

This is also confirmed by another source which states "Over a recent 12-year period, the increase in traffic was almost double the growth in constant dollar gross domestic product. The principal source of growth is international traffic (including the U.S.), where over the same period average annual growth rates were more than double the growth of domestic traffic. These growth patterns are expected to continue in the future."[36] We now review the findings of some report on the above issues.

a) Removal of restrictions on 3rd and 4th freedom rights

A study in the U.S. on the transatlantic market states "Our analysis

shows that partial aviation liberalization through Open Skies agreements has led to a substantial increase -10 percent- in the number of transatlantic passengers overall."[37] Assuming that the same increase holds for Canada then liberalizing these agreements with respect to routing, frequency, pricing and code sharing should result in benefits.

Transport Canada in its 2003 publication *Straight Ahead*, stated "On the international front, the government is not prepared at this time to give foreign carriers unrestricted access to Canada's international market as a matter of policy, out of the concern for the impact on the Canadian airline industry. ... Additional rights, such as those involving passenger or cargo services to third countries, will also be considered for exchange on a case-by-case basis. In making the determinations, the government will weigh the benefits to Canadian communities of enhanced service and the availability of balanced and usable opportunities for the Canadian airline industry. *The government will continue the gradual liberalization of Canada's bilateral air agreements, using the 1995 Canada-United States Open Skies agreement as a guide.*"[38]

b) Removal of restrictions on 5th freedom rights

A submission to Transport Canada recommended that a new international air policy should continue to seek reciprocal fifth-freedom rights as it can make a international route more competitive to the benefit of travelers or provide service where there is no service.[39] Removal of restrictions on existing 5th freedom rights will have the same benefits as in the case of removing restrictions on route, frequency, pricing and code sharing in a domestic market. However, since the advantages or disadvantages that Canadians vs foreign countries receive may not be the same, a case by case situation will have to be examined before any restrictions are removed.

c) Ninth Freedom Rights

This was reviewed under Section I(d) [p. 4] in which one study indicated that granting full rights of cabotage to all foreign carriers would be a very big step to take unilaterally and is not one we would

advocate at this time. d) A Common Aviation Market.

On a theoretical basis, an Open Aviation Agreement will increase efficiency and economic benefits by: more efficient firms replacing less efficient firms*; exploitation of size-related economies over a larger market; closer integration among firms leading to pricing synergies*; output expansion*; cross-border flows of capital; and cross-border flow of labour.[40]

A study on forming the EU-US Open Aviation Area for the European Commission indicates that the EU-US open aviation agreement would generate increases in trans Atlantic travel by 4 to 11 million passengers a year representing increases of between 9% to 24%. This could lead to decreases in travel costs and savings to consumers of more than Euros 2.5 billion on transatlantic travel.[41] This does not include the significant benefits on intra-EU travel (Euros 2.5 billion) or related industries (estimated to be about Euros 8.2 billion). It is also worthwhile noting that the assumptions on which these estimates are based are conservative. It is believed that negotiations will eventually result in a EU-US Aviation Area (with ownership limits and route specific cabotage rather than total cabotage).

If one applies the same percentages indicated above to Canadian transatlantic traffic (using 2000 data), we would estimate the increase in Canada transatlantic traffic would be between .676 million to 1.8 million passengers. This could lead to decreases in travel costs and savings to consumers in Canada.

The US-EU agreement will undoubtedly have an impact on Air Canada, the competitive position of Canadian vs US carriers, Canadian airports, Canadian travellers and the travel balance of payments. This has generated comments in several newspapers on the desirability of Canada joining the US-EU negotiations to form a transatlantic aviation area. One study states "In such a negotiation, it would be in Canada's interests to embrace the EU approach to the nationality provision, the reciprocal removal limits and the right to cabotage operations. The purpose of this would be to encourage the flow of new capital into the Canadian airline industry and allow for the right of establishment by foreign investors. The impact would be substantial and on the whole

positive. Such an agreement would allow European- and US-owned airlines to operate domestic services in Canada, the most obvious routes being those that feed and flow traffic to and from the principal Canadian international airports (Toronto, Vancouver, and Montreal). ... the resulting expansion of service would considerably benefit travellers, shippers and airports."[42]

If one believes that there is a strong possibility of a sea change in the foundation of the international rules governing air transport, then it is suggested that the government should continue removing the restrictions on 3rd and 4th freedom rights and to further negotiate removing restrictions on 5th rights on a reciprocal basis. If Canada does not wish to join the EU-US negotiations, it should continue to monitor the EU-US negotiations and become part of the transatlantic aviation area if an agreement is reached.

IV. Other Air Transportation Issues

Airport Rent Issues: Commentaries on airport rents have surfaced from time to time in various newspapers. In 2002, ground lease payments accounted for 22% of the total expenses of the largest eight Canadian Airport Authorities. The Panel which examined this matter is of the opinion that there is a need to address concerns about the rental formula that determines airport lease payments to the government and "there is room for dispute about the level of payments required to compensate the government fairly for the assets it is making available. The current formula, however, linking lease payments to net revenues, reduces incentives for cost minimization. A simplified formula based on passenger traffic would be more consistent with a governance regime aimed at encouraging efficient management."[43] Recently, airport authorities have called for a review of the government's policy on the grounds that the rent payments threaten their financial viability. It could also affect the competitive position of airports as one carrier has already indicated that it will withdraw their services from Pearson's International Airport. A rent review process is being undertaken by Transport Canada and a re-balancing of all interests is expected to be addressed. We believe the rent policy should be flexible and agree with the Transport

Minister's statement that rents should be reduced, provided the savings are transmitted to the airlines.

V. Concluding Remarks

In this paper, we have described our views on the likely new destinations for air transportation in the near future.

In brief, with respect to the domestic market, we expect the ownership limits to be raised to 49% and the 6th freedom rights to be modified; with regard to the transborder market, we expect the restrictions on co-terminalization for courier cargo to be removed and the restrictions on 8th and 6th freedom rights for cargo to be removed together with provision of 7th freedom rights to aircargo carriers at all transshipment aircargo areas; with regard to the international market, we expect the government to continue to negotiate removal of restrictions on 3rd and 4th freedom rights on a reciprocal basis and to seek 5th freedom rights on a reciprocal basis. In addition, we expect major changes in liberalization of the air policy *if* the United States and the European Union succeed in forming a transatlantic open skies aviation area. The extent of these changes will depend on whether it is a complete open skies transatlantic aviation area or somewhat less.

There are considerable gains to be achieved from having a transatlantic open skies aviation agreement. This will be beneficial not only to Canadian travellers but also to the airline and related industries. Thus, if one believes that the principal objective of Canadian air policy is to permit Canadian air travellers to benefit and to enable the airline industry to adapt and gain from the inevitable long run changes in international aviation, one can expect new developments in the near future - the only question is when? Events such as the Bogor Declaration of 1994 (i.e., free trade within APEC for developed countries by 2010 and free trade with all developed countries by 2020) may also act as a catalyst to any liberalization initiative. The Minister of Transport's willingness to initiate and consider air liberalization is an important and fundamental change in philosophy. This should enable Canada to act promptly when changes are introduced by our key global aviation partners placing air policy on the right flight path.

Bibliography/Endnotes

1. Speaking Notes for Transport Minister Jean-C. Lapierre at the Canadian Airline Investment Conference, Toronto, Ontario, October 29, 2004 (www.tc.gc.ca).

2. *Air Liberalization: A Review of Canada's Economic Regulatory Regime as if affects the Canadian Air Industry*, November 2004.

3. Exit notice is still required. The Agency continues to be responsible for hearing complaints about unreasonable basic (economy) fares and unreasonable fare increases on monopoly routes anywhere in Canada. On review, it can order roll-backs and even refunds.

4. It examined a range of actions which might be proposed to government in the event of a monopoly developing in the Canadian airline market. Six scenarios were discussed: three involving investment restrictions [i.e., relaxing foreign ownership; providing foreign carrier right of establishment; and divesting regional carriers] and three involving expanding service rights for foreign carriers [i.e., cabotage for overseas carriers; limited cabotage rights for U.S. carriers; and indirect services through U.S. gateways]. *Competition in Transportation*, NTARC Report, p. 207.

5. *Competition in Transportation* NTARC Report, p. 120.

6. It stated "We believe that raising the foreign ownership limit to a higher level would remove a significant barrier to entry and enhance competition in the domestic air market, particularly given the capital-intensive nature of the airline industry and the need by new entrants and exiting carriers for capital. In our view, increasing the foreign ownership limit may ensure that Canada has two competing national airlines, rather than one dominant carrier." *Restructuring Canada's Airline Industry: fostering Competition and Protecting the Public Interest*, Report of the Standing Committee on Transport, Dec. 1999, p. 5.

7. *Airline Industry Restructuring in Canada, Senate Committee on Transport and Communications*, December 1999, pp. 8 -9.

8. *Government Response to the Standing Committee on Transport Report Entitled "Restructuring Canada's Airline Industry: fostering Competition and Protecting the Public Interest"*, February 2000, p. 2.

9. See recommendation 7.3, *Vision and Balance*, CTARP Report, June 2001, p. 125.

10. Lazar, Fred, *Turbulence in the Skies: Options for Making Canadian Airline Travel More Attractive*, C. D. Howe Institute, Commentary, No. 181, April 2003, p. 24.

11. *Straight Ahead*, A Vision for Transportation in Canada, TC, 2003, p. 29

12. See recomm. 16 and 17. *Competition in Transportation*, NTARC Report, p. 121.

13. See recommendation 7.3, pp. 122-123. Other recommendations pertained to access to airport facilities, access to frequent flyer program, removal of Agency's power on passenger and cargo fares on monopoly routes, and 180 days notice of exit by Air Canada. See recommendation 7.2/7.7, *Vision and Balance*, CTARP Report, June 2001, pp. 125-129.

14. Id. pp. 12-13.

15. Id. pp. 11-12.

16. Ross, Thomas, *Policy proposals for enhancing competition in Canadian airline markets*, March 11, 2001, pp. 5-20.

17. Alan Toulin, "Extend NAFTA to the skies panel says," *National Post*, Thursday, July 19, 2001, pp. A1-10.

18. *Straight Ahead*, A Vision for Transportation in Canada, TC, 2003, p. 29

19. See footnote 2, November 2004, p. 5.

20. It did contain a few restrictions which were time related that no longer apply eg. phasing of U.S. service and U.S. cargo service to Montreal, Vancouver and Toronto. It also did not affect operation of charters.

21. In addition, to give greater effect to the above liberalization the agreement provided for the availability of slots and sharing of codes (Annex II and Annex V).

22. For a review of these studies see Monteiro, Joseph, David Krause and André Downs, "The open skies agreement between the United States of America and Canada - the results. Does it suggest the need for a wider pact and a more liberal air pact?" *CTRF*, 2002 Conference, St. John's, Newfoundland, May 11 - May 15th, 2002, pp. 319-337.
23. *Straight Ahead*, A Vision for Transportation in Canada, TC, 2003, p. 30.
24. See footnote 2, p. 5. There was marginal difference between the greater discontinuance of transborder services (77) vs new services (68) between 2001-2003.
25. Open Skies for air cargo, *Financial Post*, Thursday, June 3, 1999, p. 6.
26. Id.
27. Id.
28. *Financial Post*, Monday, January 26, 2004.
29. "Airfreight," *Canadian Transportation Logistics*, Nov./Dec. 2003, pp. 29-30.
30. See footnote 6, p. 11.
31. See footnote 8, pp. 4-5.
32. See footnote 9.
33. Keith McArthur, "Extend free trade to air: report," The Globe and Mail, Thursday, July 19, 2001, p. B3.
34. U.S. unlikely to rush in if open skies expands, *Globe and Mail*, Nov. 9, 2004, p. B7.
35. See "Consumers Get More Choices in New Policy for International Passenger Charter Air Services," Transport Canada, *New Release* No. H024/2000, April 4, 2000.
36. New Destinations in International Air Policy, *IRPP*, Vol. 4, no. 2, Oct. 2003, p.19.
37. *The Economic Impact of an EU-US Open Aviation Area*, The Brattle Group, December 2002, p. VI.
38. See footnote 23, p. 30.
39. Submission to Transport Canada, by the Commissioner of Competition, April 20, 2001, p. 5.
40. See footnote 37, p. 1V. * means factors used in estimating gains.
41. Id.
42. See footnote 36, p. 18.
43. See footnote 13, p. 163.

COMPUTER RESERVATION SYSTEMS IN CANADA - CHANGING REGULATIONS AND THEIR ECONOMIC RATIONALE

Joseph Monteiro and David Krause*

I. Introduction

Computer reservation systems (CRS) have been described as the nerve centre of airline marketing, operations planning, revenue management and accounting. Their market power over the airlines in the mid 1980s led to the belief that CRS vendors could engage in a variety of anti-competitive practices. It was therefore not surprising that this led to regulation of CRS in the US, EU and Canada and even to a Code of Conduct by the International Civil Aviation Authority. Two decades after regulations were passed in Canada, the market power of CRS has steadily eroded raising questions as to the need for these regulations.

This paper describes the computer reservation systems in section II. Section III examines the background and history of CRS rules. In section IV and V, the developments between 1992 and 1995 and after 1995 are reviewed. In section VI, the structure of the CRS industry, its economic characteristics, and the economic rationale for CRS regulations in Canada will be examined. In section VII, the future of the CRS will be examined followed by a few concluding remarks.

II. The Computer Reservation Systems

In this section, various forms of electronic airline ticket distribution will be briefly examined.

Computer Reservation Systems in Airline Ticket Distribution: Over the last fifty years, three basic forms of electronic airline ticket distribution emerged: the Internal Reservation Systems (IRS), the Computer Reservation Systems (CRS) and the Internet.

Internal Reservation Systems - The IRS which dates back to the mid-1950s was the initial form of automated ticket distribution. The IRS consisted of a large resident computer database; a software interface, and a set of display/printing terminals. Each airline had its own IRS system which handled requests for ticket information and ticket purchase through the IRS.

---

* *The views expressed are those of the authors and are not purported to be those of the Commissioner or the Competition Bureau.*

Computer Reservation Systems - CRS were developed in the late 1960s and early 1970s as an additional form of ticket distribution in the United States. A CRS consists of three distinct components: an internal reservation system; a distribution system; and a complex computer software with a high-speed switching system. The first contains the carrier's seat inventory together with associated information on seats sold, seats available, prices and other passenger information. The second consists of a network of telecommunication equipment and computer terminals linking CRS and travel agents. The third receives travel agent inquiries, interfaces with various databases, assembles requested information and registers bookings on all hosted reservation systems. The basic functions of CRS originally consisted of listing services and fares, booking airline seats and issuing tickets. Today, however, CRS: access millions of routes and fares; construct complicated itineraries with up-to-date prices; provide additional features such as advance seat selection, boarding passes, and frequent flyer data; and make market calculations.[1]

Internet - The late 1990s witnessed the development of an alternative form of ticket distribution through the Internet. Both airlines and travel agencies have established websites where consumers can research airline service options and make bookings. There are online travel agencies that are available through the Internet. The Internet is not a perfect substitute for the CRS as the extent of the information it provides is usually limited and the fares available through the Internet may not be available through the CRS.

The IRS, CRS and the Internet are used to varying degrees by distribution channels for the sale of airline tickets. The first (IRS) is used by airlines, the second (CRS) basically by travel agents (brick-and-motar) and reservation agents, and the third (Internet) by on line travel agencies, individuals and airlines.[2]

III. Background and History of CRS Rules.

Merger between Reservec and Pegasus: Attention to the CRS industry in Canada resulted from the merger of Reservec and Pegasus. As a result, on March 4, 1988, the Director of Investigation and Research (i.e. Director) under the *Competition Act* filed an application with the Competition Tribunal challenging the merger[3]of the Reservec CRS and Pegasus CRS.[4] The former was owned by Air Canada and the latter by Canadian Airlines International (i.e., CAIL). The merged systems would be operated by Gemini Group Automated Distribution Systems Inc. (i.e., Gemini), and owned by Air

Canada and PWA Corporation - the parent of CAIL. The merged operation would account for approximately 80 percent of the CRS business in Canada. Less than 20 percent would be accounted for by Sabre (which was owned by American Airlines) and the remaining would be accounted for by two or three other CRS. In light of this, relief was sought in the form of a dissolution of the merger or any other order that the Tribunal saw appropriate. On July 7, 1989, the Director requested the Tribunal for an adjournment of the application seeking dissolution of Gemini and filed an amended application for a Consent Order.[5] The Consent Order allowed the merger under certain conditions into a single system operated by Gemini. One of the requirements of the order was for CAIL to keep its internal reservation system in Gemini at least till 1999. This order also contained the first CRS rules.[6]

The 1989 Rules: The rules were designed to prevent unfair, deceptive, predatory, and anti-competitive practices in air transportation by CRS vendors to subscribers (i.e., travel agents or neutral sources of information). The major provisions in the rules were on: 1) Neutral display of information (rule 5);[7] 2) Non-discriminatory contracts with participating carriers (regarding fees, information on fees or fee arrangements with participating carriers, access and tie-ins) (rule 5); 3) Restraints on contracts with subscribers (rule 6);[8] 4) Access to airline information to be provided by the owning carrier on a non-discriminatory basis (i.e., completeness, timeliness and accuracy) (rule 7); 5) Service enhancements to be provided to participating carriers and subscribers on a non-discriminatory basis (rule 8); and 6) Marketing information (non-carrier specific) to be provided to all carriers on non-discriminatory terms (rule 9). These rules, however, did not provide any person an exemption from the operation of the *Competition Act*.[9]

IV. Developments Between 1992 and 1995

Developments in the CRS Industry: Between 1989 and 1992, very few changes occurred in the CRS industry. Ownership of CRS by the airlines continued, though the market share of the CRS firms had changed. Sabre's market share increased from 20 to 40 percent. The alternative to the CRS distribution, the Internet, had just emerged but its potential had not yet been fully visualized. One noticeable development was the change in the financial situation of the airlines, which would inevitably affect the ownership pattern of the CRS as seen in the discussion hereafter.

Application for Changes to the Merger Consent Order: On November 5, 1992, the Director filed an application under section 106 of the *Act* asking the

Competition Tribunal to modify the Consent Order issued by the Tribunal in July 1989. The Director asked the Tribunal to provide for the early termination of a hosting contract which requires CAIL to keep its internal reservation system in Gemini at least till 1999. The Director alleged that due to the deterioration of the financial viability of CAIL, the only alternative open which would not substantially lessen competition was a proposed transaction with AMR Corporation, as CAIL's bankruptcy would leave Air Canada with a monopoly. Part of the transaction was the transfer of CAIL's internal reservation system to Sabre, the CRS operated by AMR Corporation's wholly owned subsidiary American Airlines, Inc.

The initial decision rendered by the Competition Tribunal in April 1993 indicated that it had no jurisdiction to grant the modification sought.[10] However, it later reheard the matter addressing two issues: a) did the Gemini merger cause the likely substantial lessening of competition? and b) what was the appropriate remedy? In its decision, the Competition Tribunal concluded that the Gemini merger did cause the likely substantial lessening of competition and ordered the dissolution of Gemini by November 1994, in the absence of an alternative agreement by the parties. Air Canada, Covia and Gemini announced that they would appeal the Order, however it was later abandoned. The transaction between the AMR Corporation and CAIL, including the transfer of the latter's CRS, was completed in April 1994.[11] This was followed by a formalization of the 1989 Tribunal Rules into regulations.

The 1995 Regulations: On June 6, 1995, regulations pertaining to CRS were passed pursuant to the 1992 *Aeronautics Act* (s. 4.9). This was not unexpected as the Tribunal Rules were to be temporary since with the dissolution of the merger the Consent Order and rules would be rescinded. The intent of the 1995 regulations continued to be the fairness and neutrality in the operation of the CRS so that consumers would have access to comprehensive, and neutral information on passenger air services. Certain elements of the 1989 Rules were retained with new concepts, in particular entitlements of individual consumers. The major provisions of the 1995 regulations were: 1) An obligated carrier provision which imposes participation by an airline in **all** computer reservation systems if it has an ownership in a CRS or if it has a domestic market share of more than 10 percent (ss. 6-8); 2) A non-discriminatory display information clause to ensure neutrality and comprehensiveness to travel agents and an access clause to individual consumers through their personal computers;[12] 3) A clause for preventing

CRS vendors from engaging in price discrimination among airlines (s. 20); 4) A non-discriminatory provision (with regard to promptness, completeness and accuracy) for marketing information to all participating carriers (s. 27); 5) A provision for personal booking information (s. 28); 6) A provision containing obligations of subscribers to consumers (s. 29); 7) A set of clauses containing several prohibitions (eg., a prohibition of tie-ins to obtain any type of benefit; a prohibition in the promotion of any system by participating carriers, a prohibition of any carrier by any CRS vendor, and a prohibition of joint marketing activity by carriers and vendors) (ss. 32-35); and 8) A provision providing for Ministerial authority to prohibit airline participation in any CRS and a provision to deal with circumstances where Canadian carriers have been unable to obtain non-discriminatory treatment in CRS abroad (ss. 36 and 37).

V. Post 1995 Developments

Developments in Alternative forms of Ticket Distribution and Ownership: The first most important post 1995 development was the use of the Internet as a low-cost distribution outlet. As a result, the Internet enabled some airlines to put competitive pressure on CRSs to reduce costs through traditional channels of distribution. According to submissions to Transport Canada on CRS Rules, more than half of travel agency domestic air sales would be done outside the CRS in 2004. For domestic low cost carriers, approximately 95% of their bookings occur outside the CRS environment. For Air Canada, more than 50% of its bookings (within Canada) are through the Internet (either through travel agencies or directly) and for Westjet, approximately 70% of their bookings were through the Internet. The second most important post 1995 development was an end to the vertical ownership relationship between the airlines and the CRS through divestiture. As a result, today no Canadian or US airline has any ownership interest in any of the four major CRS.

The 2004 Amendments: The above two developments had major implications. First, new, innovative, low cost distribution options began to emerge. Second, the existing regulations prevented the industry form acting in the most efficient manner. In light of these changed market circumstances and the need to promote greater reliance on market forces in the distribution system, the CRS regulations were amended in 2004. Several of the existing regulations were removed [13] other than: 1) the requirements that CRS be operated in a comprehensive, unbiased and neutral manner so as not to unfairly disadvantage individual airlines; 2) the obligation of carriers to consumers; and 3) the

prohibition against anti-competitive practices. These were described in greater detail in the preceding section under 2, 6 and 7.

It should be noted that the provision indicating that there was or is nothing in the rules or regulations that exempts any person from the operation of the *Competition Act* continued to be retained.

VI. Structure of the CRS industry, economic characteristics, and the economic rationale for the regulations

Industry Structure: In Canada, the first CRS were *Reservec* and *Pegasus* owned independently by Air Canada and Canadian Airlines International. In the mid 1980s, the two accounted for approximately 80 percent of the business in Canada. Less than 20 percent was accounted for by Sabre (which was owned by American Airlines) and the remaining was accounted for by two or three other CRS. Today, the two major CRSs are *Galileo Canada and Sabre* (the former was called *Reservec* owned by Air Canada; and the latter was operated by American Airlines which later took over *Pegasus* owned by Canadian International Airlines).[14] As of 2004, the market share of CRS airline bookings in Canada for Sabre, Galileo and Worldspan were 55%, 36% and 8%.[15] [16]

Economic Characteristics of CRSs: The CRS business involves high fixed costs most of which tend to be sunk. As one source noted, the sunk costs necessary to enter the CRS industry appear to be large, implying that the barriers to de novo entry are high.[17] In contrast to these costs, the variable costs of operating the system are very low or near to zero. As a consequence, the industry is characterized as a natural monopoly - one that displays large economies of scale. Given these characteristics, there are no technical reasons why total demand could not be met by one mega CRS and the structure of the CRS industry could not gravitate to a monopoly.

CRS derives their revenue from three sources: fees from hosting internal reservation systems; fees (transaction) for other services supplied (seats sold, cars rented or hotel rooms booked); and fees for leasing terminals and providing training and services. The CRS grew rapidly since the mid 1976 for the following basic reasons: it provided travel agencies with access to extensive information reducing consumer search costs; it provided travel agencies an important tool to compete among themselves; it provided travellers using these travel agencies other services with minimum inconvenience; and it provided airlines a primary channel through which its tickets could be sold and through which it could compete with other airlines.

Pre-deregulation period: In the pre-deregulation period, there were no specific concerns raised by CRS as they were seen as labour saving devices. Airlines invested and owned them and when route entry, frequency and fares were regulated, very little market power advantages could be obtained from CRS. As a consequence, CRS did not attract attention as there was no incentive for anti-competitive practices.

Post-deregulation period: In the post deregulation period, all that changed. As one source noted "In deregulated markets, passengers have very many more options for their journeys, in terms of carrier, fares, routing, etc. and they can no longer rely upon any one individual airline to provide them with a list of the available alternatives. For that they have to go to travel agents, the vast majority of whom are linked into one or other of the powerful CRSs owned or hosted by major carriers."[18] As a consequence, the incentives to gain market share and to engage in anti-competitive practices changed. While in other countries, CRS vendors were alleged to have engaged in a variety of anti-competitive practices, concerns in Canada about the CRS stemmed from fears of a CRS monopoly and the absence of competition. These concerns have stemmed from two basic hypotheses about CRS: first, the market power hypothesis; and second, the bias or display preference hypothesis.

The Market Power Hypothesis - According to the market power hypothesis, if CRS have market power they should be able to charge prices that exceeds competitive levels for a significant period of time or the power to exclude competition. CRS vendors participate in two basic markets: 1) travel agents and 2) airlines. Therefore, some of the common sources of market power in these markets will be briefly examined.

1) Travel agents - CRS have information on most airlines and are therefore effective substitutes for one another. As a consequence, travel agents enjoy effective competition on both price and service from CRS. This is suggested by the fact that some CRS offer a low or a negative price to travel agents to increase their travel agency base. This implies that as far as travel agents go, CRS do not have significant market power.[19]

2) Airlines - The airlines do not consider CRS to be substitutes for one another as the majority of agents use only one CRS and each control access to a discrete group of agents. Therefore in the absence of travel agents making bookings outside a system, each CRS constitutes a separate market. As the US Department of Justice, (i.e., DOJ) stated "each CRS constitutes a separate market for air carriers, and each is a monopolist with market power over carriers

that want to sell airline tickets in areas where the CRS has a significant number of travel agents."[20] In addition, "no nationwide carrier can afford to forego listings on any major CRS and therefore, rivalry among the CRS cannot be expected to serve as a check against the exercise of market power by any one of them."[21] Further, CRS owners face little competitive pressure in setting booking fees to airlines. As Air Canada in its submission to Transport Canada states "Airline distribution costs in a CRS environment continue to rise each year. For Air Canada, segment fees have increased on average 8.4% each year. As ticket prices decline, the proportion of CRS distribution costs increase."[22] Airlines have also made attempts to persuade travel agencies to by pass the CRS to reduce their market power.

The above suggests that CRS have market power over airlines, as stated by the US Department of Transport (i.e., DOT) in the US context "Overall, however, we find that the systems currently still have market power over most airlines, although the continuing changes in airline distribution, particularly the growing importance of the Internet for airline's travel agents, and travellers, should continue to erode the system's market power. Our finding that the systems have market power is consistent with the Justice Department's conclusion."[23]

Display Preference Hypotheses - Display preference refers to the practice of systematically improving the position of particular flights on CRS display screens based on the identity of the supplier. Display preference is not unique to CRS, it is a common phenomena where there is a limited supply of preferred display space. There are two basic views on display preference: the competitive vs the anti-competitive. The latter view appears to have initially gained acceptance as regulations were designed to prevent it.

Display preference hypothesis as a competitive device emphasizes that the value a firm places on display preference basically depends on four factors. First, display preference may reduce search costs for their preferred supplier. Second, the willingness of a firm to purchase display preference may provide a signal of product quality. Third, display preference may increase awareness for new suppliers. Fourth, display preference generates new opportunities for entry in the airline industry.

Display preference hypothesis as an anti-competitive device emphasizes two basic factors. First, display preference is used for anti-competitive exclusion. Second, it is used for deceptive trade practices.

Regarding the first, studies indicate that between 70% to 90% of all tickets

sold through the CRS are based on displays in the primary screen. It is therefore claimed that display preference is used as an anti-competitive device since it inhibits competition by diverting traffic to airlines on the primary screen from airlines excluded from the primary screen. Four economic theories have been advanced to support the anti-competitive view: increasing search cost, raising rival cost, leveraging market power, and tying. According to the increasing search cost theory, display preference is anti-competitive as it increases the search costs to customers (Morrison and Winston). According to raising rival cost theory, display preference is anti-competitive as it increases the cost of marketing rival tickets resulting in exclusion (Krattenmaker and Salop). According to the leveraging market power theory, display preference is anti-competitive as market power can be leveraged from the CRS industry into the airline industry (DOJ). According to the tying theory, display preference is anti-competitive as owner-carrier CRS can use their power to limit competition from other CRS by refusing to pay override commissions or other benefits to travel agencies that did not use their CRS.

Regarding the second (deceptive trade practices), preference can mislead travel agents and ultimately their customers because the ranking and editing of airline services are presented on criteria that are not neutral. The DOJ states that "Unlike advertising, display bias does not assist the consumer in making an informed purchase decision. Moreover, ... the consumer may have limited ability to discern the bias and to obtain objective information."[24] Further, bias also causes travel agents and their customers to receive incomplete and misleading information on the available service options. According to the DOT "Display bias, ... provides no apparent consumer benefits."[25]

VII. The Future of the CRS

Besides the development of the Internet as a low-cost distribution outlet and an end to the vertical ownership relationship between the CRS and the airlines other developments will also affect the future of the CRS industry. As stated in the Regulatory Impact Analysis Statement to the 2004 amendments "All stakeholders recognized the profound changes underway within the air travel distribution industry. Furthermore, stakeholders acknowledged that these trends are expected to continue."[26] These developments will be briefly mentioned.

Revolutionary revisions in the air transport, travel agencies and CRS industry - The air transport industry has witnessed revolutionary changes in the last few years. Air Canada has emerged from its attempts to avoid bankruptcy and has been restructured. New competitors such as Westjet, Canjet and Jetsgo

have been able to capture a part of the Canadian market. The events of 9/11, SARS and the economic downturn have made these carriers extremely cost conscious in a competitive and contestable market. Even Air Canada has been accused of attempting to destroy the CRS business in favour of its own proprietary distribution channel in an attempt to cut costs.

The travel agency industry is also undergoing changes. There has been a reduction in the number of traditional brick-and-motor travel agencies and agency locations. New online travel agencies are emerging and gaining popularity. In the US, in 2002 the airlines obtained 58 percent of their bookings from brick-and-motor travel agencies and 15 percent from on-line travel agencies. The travel industry has also been faced with new developments such as an elimination of airlines' base commissions, the growing use of the Internet, limits on bookings outside the CRS, new consumer protection legislation and an overall decline in airline traffic.

CRS are also facing new challenges. CRS are no longer owned by airlines. Though no new CRS firms have entered the business, other forms of distribution have emerged and are a challenge on the horizon capturing a part of their business. For example, the Internet, individual airline websites and on-line travel agencies are now alternative distribution channels that compete for a share of the market accounting for 25 percent of the business. In addition, competition is forcing CRS to sign long term contracts with airlines at lower fees.

As a consequence of the above developments, new structural relationships are emerging between the air transport industry, travel agencies and CRS. Airlines which were prohibited or discouraged from owing CRS are attempting to own on-line agencies so as to cut down costs in airline ticket distribution. Examples include Orbitz and Hotwire being owned and controlled by several airlines. Further, since CRS were regulated, they have attempted to circumvent the regulations by entering or acquiring the unregulated on-line agencies so as to blunt the competitive thrust of new forms of competition into airline ticket distribution and to introduce new service products. An example includes Sabre owning Travelocity. Looser forms of agreements or alliances between them are also beginning to spring up.

New Marketing Services - New commercial information networks are being developed by CRS. Worldspan in its submission to Transport Canada states "...elimination of all non-discrimination provisions ...would lead to more differentiated services and prices and would be more efficient and preferable to the 'one-size-fits all'. ... Dropping the neutral requirement dos not mean that

CRS displays will become biased; all it means is that CRSs will have the opportunity to tailor their services to the needs of their customers. CRSs are providers of information and if the information does not meet their needs ... customers will use alternative information sources ... and CRS will lose share and revenues ...."[27] This has also been noted in the DOT Final Rule which states that CRS vendors are introducing a new crop of more flexible contracts.[28] These developments which increase the ability of travel agents to use several electronic sources of information to make bookings raise the question as to whether each CRS is a relevant market.

New sources of travel information and new services will continue to develop enabling customers to obtain information that suits their travelling needs through online travel agencies and websites. Examples of the largest on-line travel agencies are Travelocity, Expedia and Orbitz.

The Implications of recent regulatory developments

In the US, the Department of Transport (DOT) amended its regulations in January 2004 and terminated all CRS regulations as of July 31, 2004. These developments could affect the competitive position of Canadian vs. US airlines. This opinion was reiterated in several submissions to Transport Canada. For example, Westjet states "If we keep regulations as they are, Canadian airlines will be at a competitive disadvantage to US airlines on transborder [service] as our costs of distribution will be higher."[29] This becomes particularly relevant since U.S. and Canada have an Open Skies Agreement in place.

Different regulatory regimes for similar businesses could also raise calls for equitable treatment. Not regulating on-line agencies till they became viable competitors may have seemed reasonable initially, if one subscribes to the view of distorting market structures to encourage entry and competition. However, is it equitable any longer now that on-line agencies have established themselves as alternative forms of distribution to the CRS? This could call for removal of all CRS regulations as they place CRS at a competitive disadvantage?

The emerging structural changes in the airline, travel agency and CRS industry may have created loopholes in the regulatory regime due to the on going deregulation process. This raises questions as to: whether such schemes fall under CRS regulations or the competition laws raising regulated conduct defence issues (though nothing in the CRS regulations prevents the operation of the *Competition Act)*; and whether it would be administratively convenient to deal with CRS only under the *Competition Act*.

Should the remaining regulations in Canada remain? - Worldspan in its

submission to Transport Canada states "... the October recommendations do not go far enough. We believe an open marketplace is the best way to serve consumers, and that full and prompt termination of CRS Regulations would advance this objective."[30] Other submissions have also provided a similar view. For example, Galileo International, Inc. believes that total deregulation is a superior alternative to partial deregulation and its submission states "...if the Government truly believes that the changed marketplace has substantially eliminated the risk of anticompetitive practices by dominant carriers, the Regulations should be totally eliminated."[31]

Another factor that should be considered in determining this matter is whether the existing regulations are an effective deterrent and impose costs and whether they can be dealt with under the existing laws. While such a study has not been undertaken in Canada, the DOJ concluded that the existing rules have been ineffective in curbing CRS market power and that the rules impose its own costs and should not be readopted.

In sum, the above developments are spearheading our thinking of how competition in the CRS market works. Both the market power hypothesis, and the display preference hypothesis are being increasingly questioned and with this the need for any CRS regulations. The former not only because of the lack of vertical ownership relationship but also because of the broadening of the relevant market to other electronic sources to make bookings. One economist reviewing the situation in the U.S as early as 1992 stated "Yet the two most important antitrust charges, that of market power and display preference, appear theoretically incorrect and derived from important misunderstandings about how competition works in the CRS industry. These misunderstandings have led to unwarranted policy actions in this area."[32]

VIII. Concluding Remarks

The need for regulation arose because of: CRS market power over airlines; airline market power in passenger air transportation markets; and airline ownership of CRSs. It is believed that these basic tenets no longer hold. The Government in its Impact Statement indicated that it is federal policy "To regulate only when and if necessary, given that competition and other market forces may provide equal if not better market discipline ..."[33]

By eliminating many of the CRS regulations in its May 2004 amendments, the Government has moved in the right direction. However, did it go far enough? Some observers believe that it did not and that it should have opted for

complete deregulation for several reasons: the rationale for the bias display hypothesis is no longer valid; the regulations will stifle the development of innovative CRS products; the existing regulations could affect the competitive position of Canadian and US airlines in transborder markets by raising distribution costs in Canada; and the provisions in the *Competition Act* can deal with anti-competitive practices. If consumers need further protection, a less restrictive alternative to the present regulations could be requiring the travel agent to inform purchasers of tickets that the CRS being used has a display algorithm preference.

Finally, the emergence of new developments in the distribution of airline tickets and travel products have also led to the emergence of new questions. Should the Internet distribution, on-line travel agencies, and joint airline websites be regulated? These issues have not been examined in Canada in depth, however, the DOT in its Final Rule on CRS Regulations stated "We still believe that we should not adopt rules governing airline distribution over the Internet, whether through airline websites or on-line travel agencies."[34] This view was also reiterated by the DOJ in its submission to the DOT on CRS Regulations. Similarly, with regard to on-online agencies the DOT stated "We tentatively concluded that we should not adopt rules regulating their conduct..."[35] and with regard to joint airline websites it stated "we decided not to propose rules regulating the operation of joint airline websites. ... Adopting general rules ... would be premature. The enforcement process would be the best means for addressing any problems with deceptive practices and unfair methods of competition created by such a site."[36]

Bibliography

1. Humphreys, B. K., **New Developments in CRSs**, Paris; Institut du Transport Arien, 1994.
2. House of Commons Transport Committee, **Airline Competition: Computer Reservations Systems**, Third Report Session 1987-88, London: HMSO, 1988.
3. Truitt, L.J., Teye, V.B., and Farris, M.T., The Role of Computer Reservation Systems: International Comparisons, **Tourism Management** 11(1), March, 1991, pp. 21-36.
4. **Report of the Department of Justice to Congress on the Airline Computer Reservation System Industry**, 1985.
5. Kleit, Andrew N., "Computer reservation systems: competition misunderstood," **The Antitrust Bulletin**, Winter 1992, pp. 833-861.
6. **Computer Reservation System (CRS) Regulations**, Final Rule, US DOT.
7. Hanlon, Pat, **Global Airlines**, Butterworths, 2000, pp. 66-74.

Endnotes

1. **National Transportation Agency Annual Report**, 1992, p. 46.
2. The market share of sales by different channels of distribution in the US in 2002 were as follows: off-line sales by airlines (17%); on-line by airlines (10%); off-line sales by travel agencies (58%); and, on-line by travel agencies (15%). **Computer Reservation System (CRS) Regulations**, Final Rule, US DOT, p. 7.
3. It was held that the merger prevents or lessens, or is likely to prevent or lessen competition substantially in the provision of CRS services to airlines, travel agents and consumers in Canada.
4. These systems are used by airlines and travel agents for the distribution and sale of airline seats and related travel services.
5. Before this event, AC and PWA Corporation announced a second merger between Gemini and a U.S. based CRS called Pars. The Director sought an amended application seeking dissolution to it and the parties to this merger later abandoned it.
6. In the US, the Civil Aeronautics Board (now the Department of Transport i.e., DOT) first imposed CRSs regulation in November 1984 which were later revised and re-adopted in 1992 by DOT and amended in January 2004.
7. That is the primary display shall include schedules, fares, rates and availability of all carriers in accordance with various factors. These factors are the ordering of information and display of connecting flights shall not be based on carrier identity; the construction of connecting flights shall use at least nine points if possible for each city-pair, if connecting points are used for such construction; and the same standards of care and timeliness to loading the information shall be used.
8. For example, the maximum length is three years, the use of other systems are prohibited, the tie-in between the use of any system by the subscriber and the sale of transportation services are prohibited, the tie-in between receipt of commission or other incentives and the use of owning carriers system is prohibited, and the pricing based on carrier flights sold by subscribers is prohibited.
9. See Consent Order of the Competition Tribunal between the the Director of Investigation and Research and Air Canada et al, CT- 88/1.
10. The Competition Tribunal indicated that had it the jurisdiction, it would have granted the relief requested. The Director and PWA Corporation appealed with respect to the Competition Tribunal's decision regarding jurisdiction, and Air Canada, Covia and Gemini cross-appealed with respect to the Competition Tribunal's decision regarding the Tribunal's conclusion.
11. Gemini was reconstituted to form two separate companies: Advantis Canada and Galileo, the former owned by IBM Canada Ltd. and the latter by Air Canada.
12. Further, charter air carriers can include their services (excluding other tourism services) on payment of a fee (ss. 9 and 10); furthermore, the ordering of information and displaying of connecting flights shall not be based on carrier identity; the selection of connecting points for each city-pair shall not be based on carrier identity or airport identity and the system vendor shall not discriminate between airports; if connecting points and double connect points are used for constructing connecting flights at least fifteen connecting points and ten double connect points where possible shall be used for each city pair in such construction (ss. 15-16); and the same standards of care and timeliness to loading the information shall be used (s. 18).
13. The most important were: the obligated carrier provision to permit airlines the most efficient means to distribute their product; the requirement that CRS charge all airlines the same fee as

it prevents negotiation and leads to higher fees together with the non-discriminatory provision with regard to services; and the provision preventing CRS vendors and travel agents to freely negotiate the terms and conditions of their business relationship as it resulted in restrictive contracts that were not tailored to meet their needs. Others that were removed include how CRS vendors construct connecting flight, exclusion of internet based airline websites, sale of marketing data and prohibitions that could be made by the Minister of Transport.

14. Reservec and Pegasus merged. The two were operated by Gemini. Pegasus was later acquired by Sabre see section VI and footnote 9 for further details why this occurred.
15. See Submissions to Transport Canada on **Computer Reservation System**, February 9, 2004.
16. In the US, initially two CRSs dominated the market: Sabre and Apollo/Covia each independently owned by American Airlines and United Airlines, respectively. Today, the major CRSs in the US are *Sabre, Worldspan, Galileo* and *Amadeus*. The enormous power conferred on the airlines in the US through their CRSs led to the development of Galileo International and Amadeus in Europe and Axess and Abacus in the Far East. In Europe, the two major CRSs are *Amadeus* and *Galileo International* and the two smaller CRSs are *Sabre* and *Worldspan;* and in Asia-Pacific, the two major CRS are *Axess* and *Abacus* and the smaller CRSs are *Sabre* and *Galileo International.* In the United States, the market shares in 2002 were as follows: Sabre (44.7%), Worldspan (26.5%), Galileo (19.7%) and Amadeus (9.2%).
17. Kleit, Andrew N., "Computer reservation systems: competition misunderstood," **The Antitrust Bulletin**, Winter 1992, p. 839.
18. Hanlon, Pat, **Global Airlines**, Butterworth, p. 81.
19. Some argue that this need not be the case if CRS listings of airlines and other information are not the same and if travel agents typically use and have exclusive access to only one CRS through long term contracts. This suggests that CRS have significant market power. A power tempered by the presence of alternatives.
20. **Reply Comments of the Department of Justice**, Notice of Proposed RuleMaking Computer Reservation System, June 9, 2003, p. 5.
21. Footnote 17, p. 843.
22. **Transport Canada Hearings on Deregulation of Canadian CRS Rules**, Presented by Marc Rosenberg, February 9, 2004. Over the 1996-2003 period, Air Canada's average net cost per CRS segment increased from about $2.75 (US) to $4.50 (US). Westjet in its submission to Transport Canada indicates that the CRS/GDS (Global Distribution Systems) price indicies (for Amadeus, Sabre, Galileo and Worldspan) rose from a base of 100 to more than 130. Furthermore, given, the economic characteristics of this industry new entry is unlikely to eliminate the CRS existing market power. Similarly, First Air in its submission to Transport Canada states "Fees have increased 2%-5% year, after year. GDS have not shown any cost saving efficiencies in the last 10 years. Customers are not willing to pay higher fares." **First Air Comments Regarding Changes to the Canadian Computer Reservations System (CRS) Regulations**, 09 February 2004, Toronto, Ontario.
23. See footnote 2, p. 20.
24. See footnote 16, p. 17. If display preference is beneficial to the travel agency or the customers one would expect such preference to be unstable. However, this is not so given the switching costs that travel agents will have to incur and to the fact that customers are unlikely to know that better price and service options were available.
25. See footnote 2, p. 24.
26. Regulations Amending the Canadian Computer Reservation Systems (CRS) Regulations.

Canada Gazette, SOR/2004-91, Extra Vol. 138, No. 5, April 27, 2004.

27. **CRS Comments of Worldspan**, L.P. Before the Parliament Secretary to the Minister of Transport, Presented by Douglas L. Abramson, Toronto, Canada February 9, 2004.
28. See footnote 2, p. 10.
29. **CRS Regulations**, Westjet's Perspective, February 9, 2004, Scott Butler.
30. See footnote 27.
31. **Formal Comments of Galileo, Inc. On the Proposal to Amend Certain Provisions in the Canadian CRS Regulations**, November 24, 2003.
32. See footnote 17, p. 861.
33. See footnote 26.
34. See footnote 2, p. 57.
35. Id.
36. See footnote 2, p. 59.

Capstone Avionics in Alaska: An Economic Appraisal

Dr. Darren Prokop
Department of Logistics
College of Business & Public Policy
University of Alaska Anchorage

(1) Introduction

The Capstone Avionics Program (abbrev. Capstone) is a joint effort on the part of industry, academia and the Federal Aviation Administration (FAA). The goal is to improve aviation safety and efficiency by putting new and emerging technology into aircraft servicing the Yukon-Kuskokwim (Y-K) Delta in Southwestern Alaska. The technology gives pilots the ability to: navigate using an electronic map; broadcast their position to other aircraft for traffic avoidance; and broadcast to ground stations for surveillance.

Capstone represents a critical part of the FAA's goal of achieving a 20% reduction in the annual accident rate in Alaska by 2008. The typical occurrence of accidents is on take-off and landing while those causing fatalities are typically due to navigation. In fact, proportional to the contiguous United States, take-offs and landings, as a point for accidents, are on par; but it is in the area of mid-air collisions that the Y-K Delta differs with higher proportions.

There is an opportunity cost to a flight delayed due to inclement weather conditions. Of course, if the pilot's skill and information available leads him to conclude that a clear and present risk of an incident exists, his decision to stay grounded is prudent. However, if Capstone can overcome the risk then an otherwise grounded flight could take place. The economic activity facilitated by any and all such flights represents the opportunity cost of not

deploying the technology. Conversely, this further economic activity represents a gain to the market area and is a major component of any business case to be made.

(2) Capstone Technology

The Capstone Program, initially funded with $11 million in 1998, is a joint effort on the part of industry, academia and the FAA.[1] Capstone includes the bundling of numerous technologies including global positioning systems (GPS) and Automatic Dependent Surveillance-Broadcast (ADS-B). The Capstone technologies give pilots the ability to navigate using an electronic map and to broadcast their position to other aircraft for traffic avoidance and ground stations for surveillance.[2] Training of pilots and dispatchers on the use of ADS-B have been provided by the Aviation Technology Division of the University of Alaska Anchorage. The Y-K Delta market area comprises Phase I of the Capstone Program. The Alaska panhandle, or Southeast Alaska, has been designated as Phase II. For the purposes of this paper Phase I in the Y-K Delta will be discussed.

Starting in 1999 the FAA provided the ADS-B avionics package to several commercial aircraft in the market area. By 2002, 140 aircraft (85% of air taxi and commuter planes operating in the Y-K Delta)[3] had the package installed. The package link ups with ground infrastructure designed to provide ground to air communication, flight information and surveillance.[4] The pilots receive information covering current location, aircraft recognition (of other Capstone equipped aircraft in the area), terrain awareness and weather conditions. ADS-B has been approved for "Radar-Like Services" in the Capstone operating area. AWOS (Automated Weather Observation Systems) and GPS approaches have been added to provide an instrument approach capability to communities where they were not formerly available in the market area.[5] Capstone provides many on-board capabilities for the pilot. There is no doubt that visual flight rules (VFR) activity in marginal weather heightens the risks of mid-air collisions and weather-related incidents in the market area. Capstone systems have established a usable instrument

flight rules (IFR) infrastructure with terrain, traffic and weather information in the cockpit to reduce the risk of operation.

While the deployment of Capstone technology throughout Alaska is a voluntary and evolutionary process it is the case that Capstone equipped planes are able to detect planes which are not duly equipped provided that: (1) radar services are available in the flight area; and (2) the other plane is currently in radar contact with the ground. Of course, the plane without Capstone will not be in contact with the Capstone equipped plane. Such lost advantages may act as a catalyst for further equipping aircraft in Alaska.

Capstone represents a critical part of the FAA's goal of achieving a 20% reduction in the annual accident rate in Alaska by 2008. Ninety-four general aviation accidents occurred in 2003 of which 6 resulted in fatalities. The average annual amount of accidents was 101 over 1999-2003 of which 7.2 resulted in fatalities. The typical occurrence of accidents is on take-off and landing while those causing fatalities are typically due to navigation. In fact, proportional to the contiguous United States, take-offs and landings, as a point for accidents, are on par; but it is in the area of mid-air collisions that the Y-K Delta differs with higher proportions (**Capstone Phase I Interim Safety Study; ISER 2002; pp. 14-5**). However, current evidence shows a downward trend in accidents and fatalities in the Y-K Delta.

The vastness of the market area implies frequent flights into and out of the air-only accessible villages. The climate, topography and lack of infrastructure in the Y-K market area combine to create volatile and hazardous operating conditions for flights (e.g., white-out conditions, fog and airframe icing). All of these aspects help to explain why accidents in Alaska occur with 2.5 times the frequency and why pilots suffer a fatality rate four-times that in the contiguous U.S. While Capstone is designed to mitigate much of the contributing factors in accidents, it is not designed to overcome issues of aircraft maintenance and fuel handling.

(3) The Y-K Delta Market Area

Bethel, Alaska is the largest hub of aviation activity in the Yukon-Kuskokwim Delta. Located along the north bank of the Kuskokwim River about 60 miles inland from where the river flows into the Bering Sea, Bethel is not accessible by road and is dependent on aviation as the mode of transportation for business and leisure. Maritime access exists for a brief period during the summer. The city is surrounded by tundra and various streams and ponds. The proximity to the Bering Sea, in combination with 50 miles of surrounding flat land means that winds can be very cold and strong during the winter season (November through April). Indeed, Bethel is a fixture in the quality of life for several smaller communities in the market area--- specifically, access to health care professionals is facilitated by airlift into Bethel. There are about 52 villages in the Y-K Delta with each dependent on the Bethel hub.

Currently, the Y-K Delta has 10 AWOS stations installed by Capstone to support the IFR infrastructure. The Capstone project has also installed 10 ground-based transceivers (GBTs) in the Y-K Delta and another in Anchorage at Site Summit to provide a data link with the ADS-B equipped aircraft (***q.v.***, **Map 2**). And in addition to receiving information, such as weather and traffic, these aircraft also send location messages to air traffic for surveillance so that air traffic controllers are able to track and monitor flights.

Bypass mail is a form of subsidy which allows carriers to cover their costs through moving mail in the market area on behalf of the U.S. Postal Service (USPS). With USPS sharing its postage revenues with the carriers, the air carriers' costs are covered to a point where passengers may be carried at fares lower than otherwise possible. Of course, this benefit is mitigated to the extent that insurance rates are higher in the Y-K Delta due to the relatively high prevalence of accidents. These higher rates also bias against passenger carriage and in favor of shipping as a form of economic benefit rather than mobility of residents (***q.v.***, **Table 3**). Of course, Capstone technology on board may signal less risk to insurers so that rates may fall over time (barring, of course, any increased tendency

on the part of pilots to engage in risky behavior based on the Capstone provided equipment). And carriers may use the Capstone technology (especially the decal on side of the plane) as a signal of higher quality--- and passengers/shippers may willingly pay a premium.

Bethel Municipal Airport (airport code: BET) is located 3 miles south west of the city. The airport has 2 runways. The main runway is asphalt and measures 6,400 x 150 feet and the cross runway is gravel and measures 1,850 x 75 feet. Aircraft operation average about 334 flights per pay with the breakdown being: 54% air taxi (village flights); 35% transient general aviation; 6% local aviation; 4% commercial (scheduled and commuter) flights; and 1% military (**www.airnav.com**).

Map 1: Alaska's U.S. Census Areas



Note that the Y-K Delta, by and large, encompasses two U.S. Census areas: Bethel and Dillingham. Of course, as the largest city in Alaska, Anchorage is the trade gateway to these and other areas of the state. **Table 1** compares all three of these areas relative to the state as a whole. The data is assembled to give some idea of the socio-economic activity in these areas.

What may be noticed in **Table 1** is that Bethel and Dillingham are sparsely populated--- in fact there are more square miles in these regions than there are people to populate them. In terms of education, these regions have populations which are less educated relative to the state as a whole. The significantly shorter travel times to work are to be expected of the village populations which predominate. Income levels per capita are significantly lower and poverty levels are significantly higher, than they are for the state as a whole. The number of non-farm businesses is very low on a per capita basis. For Anchorage, the numbers are 0.03 or 3 businesses for every 100 people in the region. For Bethel and Dillingham the numbers are: 0.01 and 0.02, respectively. With 1 (2) business for every 100 people in Bethel (Dillingham), combined with the very small population per square mile, retail sales require travel and shipping over relatively large distances compared to Anchorage residents. The sparseness of business entities combined with the spatial challenge of doing business will have an adverse effect on the size of the economic impact multipliers as discussed below.

(4) The Business Case for Capstone

There is an opportunity cost to a flight delayed due to inclement weather conditions. Of course, if the pilot's skill and information available leads him to conclude that a clear and present risk of an incident exists, his decision to stay grounded is prudent. However, if the Capstone technology can overcome the risk then an otherwise grounded flight could take place. The economic activity facilitated by any and all such flights represents the opportunity cost of not deploying the technology. Conversely, this further economic activity represents a gain to the market area and is a major component of any business case to be made.

Unique to Alaska is a phenomenon known as "bush pilot syndrome"--- an *esprit de corps* which often leads to pilots taking risks. As such, the probability of fatalities and lost cargo is accentuated. The ability of Capstone technology to assist in navigation and provide up-to-date weather information diminishes the

risk of a given flight. Thus, the bush pilot's cost of risk-taking may be reduced.

Table 1: Demographics and Business Indicators

	U.S. Census Region			
Demographics	**Bethel**	**Dillingham**	**Anchorage**	**Alaska**
Population (thousands; 2003)	16.9	4.9	271	648.8
Population (% change, 2000-03)	5.2%	0.2%	4.1%	3.5%
Persons per sq. mile (2000)	0.4	0.3	153.4	1.1
High School Graduates (% of persons 25+; 2000)	71.0%	76.6%	90.3%	88.3%
Bachelor's Degree (% of persons 25+; 2000)	13.1%	16.4%	28.9%	24.7%
Mean Travel Time to Work (2000)	8.6 min.	10.0 min.	19.5 min.	19.6 min.
Home Ownership rate (2000)	61.1%	60.4%	60.1%	62.5%
Income per Capita (1999)	$12,603	$16,021	$25,287	$22,660
Persons below Poverty (%)	20.6%	21.4%	7.3%	9.4%
Business Indicators				
Private Non-Farm Businesses (2001)	211	90	7,979	18,589
Manufacturer Shipments ($ million; 1987)	Not Available	Not Available	$322	$3,305
Retail Sales per Capita (1997)	$4,896	$7,435	$12,392	$10,268

Source: U.S. Census Quick Facts.

The opportunity cost exists in three forms: (1) the time-cost for people and cargo waiting for flight conditions to improve; (2) the congestion that occurs due to flights queuing for short windows of flight opportunity; and (3) the loss of, or injury/damage to, people and cargo due to an incident in the air or on the ground.

A safety study completed for the Y-K Delta estimated the following benefits of Capstone technology if it were placed on all aircraft operating in the market area (**Capstone Phase I Interim Safety Study; ISER 2002; p. iii**):

- 1 out of 7 accidents (or 14%) would have been prevented; while nearly half of those leading to fatalities would have been prevented.

The study covered the period from January 2000 to December 2001. Clearly there are economic and social benefits to these estimated gains.

Of course, there are costs involved in using Capstone. The system routes aircraft in straight lines between points meaning that they are flying closer together than might otherwise be the case. Pilots need to remove their winter gloves in order to operate the on-board avionics. Training requires time and devotion and, to some degree, a change in the culture of bush pilots. Finally, the technology may engender moral hazard on the part of pilots; that is, the feeling of being safer could lead to more risk-taking than otherwise.

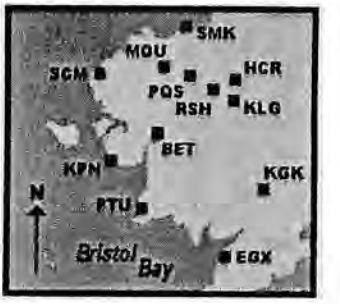
(5) Capstone Participating Carriers

A total of 29 carriers in the market area make use of the 169 aircraft equipped with Capstone technology. Seven of the carriers have head offices in the market area. These are: Craig Air, G & L Air Service, Husky Aviation, Inland Aviation Services, Neitz Aviation, Poe Air, and Yukon Helicopters. This has implications for economic impact leakages as will be discussed below.

There are 12 airports in the market area which have in place a GPS-based non-precision instrument approach capability and an automated weather observing system (AWOS). These are shown in **Map 2**. Other than Bethel, the other airports are located in or near remote villages.

Map 2: Capstone Airports in the Market Area

Bethel (BET)
Egegik (EGX)
Holy Cross (HCR)
Kalskag (KLG)
Kipnuk (KPN)
Koliganek (KGK)
Mountain Village (MOU)
Platinum (PTU)
Pilot Station (PQS)
Russian Mission (RSH)
Scammon Bay (SCM)
St. Michael (SMK)



In order to get an idea of the traffic which takes place within this network of villages **Table 2** shows the amount of round trips in 2002 involving the 29 carriers using the airports listed in **Map 2**.[6] In addition to the total number of inbound and outbound flights, the number of revenue-generating flights by Capstone participating carriers is shown. As can be seen the Capstone participating carriers have a significant amount of traffic at each airport.

The number of inbound flights gives an idea as to how many revenue flights take place which may benefit from the Capstone technology currently available to users of these airports.

Table 3 takes the data from the Capstone carrier flights noted in **Table 2** and calculates the mean inbound flight for each airport in terms of cargo (i.e., freight and mail) and passengers carried. As can be seen, the typical flight involves significantly more mail than freight; and passenger levels are less than 100. The

weighted averages across the 12 airports are: 2,188.1 freight pounds; 8,820.4 mail pounds; and 36.2 passengers.

Table 2: Small Scheduled Air Carrier Frequencies (2002) Listed by Capstone Airport

Airport	Bethel	Egegik	Holy Cross	Kalskag	Kipnuk	Kolliganek
Total No. of Flights	749	29	55	43	49	38
No. of Inbound Flights*	289	16	31	26	28	19
Airport	**Mountain Village**	**Platinum**	**Pilot Station**	**Russian Mission**	**Scammon Bay**	**St. Michael**
Total No. of Flights	69	16	49	55	57	76
No. of Inbound Flights*	38	7	27	29	33	29

* Revenue flights by only carriers with Capstone aircraft in their fleets.
Source: *Small Air Carriers: Schedule T-1; Bureau of Transportation Statistics.* (www.transtats.bts.gov).

Table 3: Mean Flight Characteristics of Capstone Carriers (2002)

Inbound Airport	Bethel	Egegik	Holy Cross	Kalskag	Kipnuk	Kolliganek
Freight (lbs.)	1,609.0	3,343.4	1,647.4	1,520.7	3,712.4	398.0
Mail (lbs.)	6,711.5	8,162.2	5,410.9	14,555.0	12,149.9	4,873.3
Passenger	82.3	52.3	16.7	43.2	40.0	10.9
Inbound Airport	**Mountain Village**	**Platinum**	**Pilot Station**	**Russian Mission**	**Scammon Bay**	**St. Michael**
Freight (lbs.)	1,778.1	2,413	6013.0	1,486.9	758.5	1,576.8
Mail (lbs.)	12,532.8	2,519.4	13,290.4	8,668.2	8,222.8	8,748.1
Passenger	55.5	9.0	63.8	24.1	18.4	18.7

Calculated from: *Small Air Carriers: Schedule T-1; Bureau of Transportation Statistics.* (www.transtats.bts.gov).

A thought experiment may be run in the following manner. In 2000 and 2001 there were 58 recorded accidents in the Y-K Delta (**Capstone Phase I Interim Safety Study; ISER 2002; pp. 58 and 60**). Taking the annual average of 29 and applying it to the total number of 2002 flights in **Table 2** (i.e., 1,285) gives an accident rate of about 2.3%. There were 572 inbound Capstone carrier flights and 34% of the 2000-01 accidents involved outbound causes (i.e., take-off and flight preparation). This leaves 38 accidents (i.e., 66% of 58) as part of inbound activity. Thus, 19 is the annual average of inbound flight accidents and this gives an accident rate of 3.3%. *Assume* that the Capstone avionics on board could prevent all 19 (statistically) expected accidents. By multiplying the weighted averages of the inbound flight characteristics in the previous paragraph by a factor of 19, one obtains the expected extra amount of cargo and passenger business safely received by the 12 airports. This would be: 41,573.9 freight pounds; 167,587.6 mail pounds; and 688 passengers. These numbers show the inbound business activity which Capstone equipment may have successfully facilitated in the Y-K Delta. But this is a very highball estimate. Why? Recall from the previous section that the 2002 safety study found that only 14% of actual accidents would have been prevented had Capstone equipment been on board the planes. Using this statistic, the number of expected annual accidents on Capstone flights falls 14% from 19 to about 16. So with 3 extra flights facilitated by Capstone equipment, the weighted averages of the flight attributes become: 6,564.3 freight pounds; 26,461.2 mail pounds; and 109 passengers. Strictly put, these numbers represent the lowball-expected extra cargo and passenger business induced by Capstone avionics.

Is there a middle ground between the highball and lowball estimates? Yes, because preventing 14% of accidents in the pre-Capstone flight regime likely means a greater percentage of those in the Capstone regime may be dependent on the equipment itself. Why? The equipment helps to prevent two other types of accident not expected to occur the in pre-Capstone regime: (1) those which might otherwise occur because Capstone allows for inclement weather flying; and (2) those which might otherwise occur because a pilot's extra operational risk-taking is being mitigated by the Capstone

information received. Point (2) indicates that moral hazard (i.e., enhanced "bush pilot syndrome"), if it exists, can be mitigated. Such outcomes encourage more economic activity than otherwise. A suggestion for improving the estimate will be made in the on future research.

In short, the extra revenues earned by the carriers (net of insurance claims) and the expenditures induced as a result of the cargo and passengers reaching various points in the Y-K Delta, represent the source of the economic impact of Capstone technology.

(6) The Economic Impact of Capstone Technology

Several ingredients are required in order to provide a sound economic impact analysis. The first is the defined spatial area which has important consequences for economic multipliers and leakages. It is within the spatial area that the economic impacts are estimated. The Y-K Delta is well defined spatially and, for the purposes of this paper, the Bethel hub is made up of well defined spokes.

The total economic impact of increased travel within the market area may be broken down to three sub-impacts.

1) Direct economic impacts are created through the goods and people (as providers of, or demanders for, services) required at various points served by the airports in the market area. The first round of expenditures comes from getting the goods and people to where the market requires them to be;

2) Indirect economic impacts are created through the support services which help to induce the first round of expenditures. Airport maintenance, transient population support (hotels, restaurants, etc.) and the Capstone technology itself are examples of these. They facilitate the critical transportation services which are necessary for village business within the market area;

3) Induced economic impacts are created through the subsequent rounds of expenditures which occur in the market area as a result of the income earned from the direct impacts.

Of course, indirect impacts can be functionally co-dependent thus making it difficult to separate out the various impacts. For example, an airline's purchase of fuel and an airport's maintenance of the runway are common costs necessary for the flight to take place. It is very difficult to allocate the precise revenues induced to the market area from each of these cost centers. The (indirect) economic impact of the Capstone technology lies in the transport value-added of goods and services it induces. A Capstone-assisted flight induces transactions in the market area resulting from the passengers and/or cargo transported. And to the airport go employment opportunities, labor income, revenue and taxes.

Capstone, itself, does provide a direct benefit to those whose revenue is directly traceable to its use. Thus, the air carriers, the training and certification institutions, the airports and aviation regulators are the direct recipients of the benefits.

Since this paper is directed at Capstone technology's total economic impact within the market area and not a particular point in the market area it avoids the problem of allocating airline revenue to specific airports. For example, if an origin or destination is outside of the market area or airline revenues are used, in part, to purchase fuel or equipment outside of the market area then revenues, to an extent, are leaking outside of the market area. Capstone may have induced the spillover but the effect is not of benefit to anyone in the market area.

The indirect benefits of Capstone technology accrue to those whose revenue is traceable to activities induced *after* the completion of the flight. Thus, ground transportation and manufacturing and/or retail services related to the air cargo are secondary recipients. If the flight involves passengers it is the service sector, ground transportation and the hospitality industry which receive the

secondary benefits. The excess of revenues earned by these parties over and above costs are the profits. These profits may be reinvested in their businesses in order to enhance them. This is an example of the indirect economic impact. More jobs are created and further labor income induces yet further rounds of spending in the Y-K market area. This point is expanded upon in the next section.

When a Capstone-assisted flight occurs revenues are generated to various parties as a result. For example, cargo brought into Bethel from Anchorage, and subsequently dispersed among the various points in the market area, mean the cargo's value-added will generate further demands for goods and services. The extra rounds of spending generated indicate a multiplier effect. If revenues accrue to parties outside of the market area the multiplier effect in the area is diminished--- as such this is called a leakage.

Multipliers are convenient because they are a substitute for surveying all induced impact recipients in the market area--- of which there may be hundreds or even thousands. Technically, multipliers are derived from input-output analysis of economic interdependences defined over the market area of study.[7] Simply put, the multiplier is the ratio of the change in total economic impact relative to the change in the induced impact which brought it about.

The multipliers for the Y-K market area, however, are expected to be relatively small for two important reasons.

1) Since the Y-K market area is dependent on Anchorage, located in south central Alaska, as a source of economic health and development, one would expect the leakage effect to be pronounced. In this way the multipliers would be lower than those regions with their own cities and mature industries;

2) Intra-market sources of value-added are not prevalent in the market area. The magnitude of the multiplier is a positive function of the strength of industry in the market area and their linkages.

Thus, no matter how large the expenditures are within the market area the rounds of spending induced, net of leakages, are likely to be few.

(7) Empirical Analysis: Regional Multipliers

To give an idea of the impact of air transportation in Southwest Alaska (inclusive of the market area) **Table 4** shows the multiplier effects[8] on labor income and employment in the region. These are also compared to Anchorage and state-wide; and to alternative modes of transport.[9]

Table 4: IMPLAN Regional Multipliers

		Transport Mode			
Region	Multiplier (2001)	Air Transport	Water Transport	Truck Transport	Ground Passenger
Southwest	Labor Income	1.60	2.40	1.36	1.26
	Employment	1.60	3.31	1.35	1.13
Anchorage	Labor Income	1.80	3.71	1.67	1.49
	Employment	2.28	5.21	1.82	1.22
State-Wide	Labor Income	1.79	3.65	1.62	1.43
	Employment	2.22	5.06	1.80	1.22

Starting with labor income, Southwest Alaska's multipliers are consistently less than those for Anchorage and the state as a whole. However, the differences are least pronounced over the air transport mode. This is because of all the modes shown air travel is least affected by differing topography over the travel route. The multiplier of 1.60 means that for every $1 of labor income generated to air carrier employees in Southwest Alaska a further $1.60 is created in Southwest Alaska.

As to employment, the Southwest Alaska multiplier is significantly lower than those for Anchorage and the state as a whole. For every 1 job created as a result of increased air carrier activity in

Southwest Alaska a further 1.6 jobs are created in the region. Another way to look at this is: the typical air carrier job creates enough economic activity in the region to warrant the hiring of almost two other people in other sectors of its economy.

Table 5: RIMS II Multipliers; Bethel, Dillingham and Anchorage

		Transport Mode			
Bethel	**Multiplier (2001)**	**Air Transport**	**Water Transport**	**Truck Transport**	**Ground Passenger**
Final Demand	**Output**	1.35	1.44	1.28	1.25
	Earnings	0.36	0.28	0.35	0.46
	Employment	9.77	7.37	9.70	30.15
Direct Demand	**Earnings**	1.36	1.94	1.26	1.15
	Employment	1.62	2.73	1.38	1.09

		Transport Mode			
Dillingham	**Multiplier (2001)**	**Air Transport**	**Water Transport**	**Truck Transport**	**Ground Passenger**
Final Demand	**Output**	1.27	1.00	1.00	1.27
	Earnings	0.34	0	0	0.47
	Employment	9.22	0	0	30.90
Direct Demand	**Earnings**	1.29	0	0	1.18
	Employment	1.52	0	0	1.12

		Transport Mode			
Anchorage	**Multiplier (2001)**	**Air Transport**	**Water Transport**	**Truck Transport**	**Ground Passenger**
Final Demand	**Output**	1.99	1.99	1.89	1.92
	Earnings	0.53	0.41	0.48	0.66
	Employment	16.35	12.44	14.44	36.89
Direct Demand	**Earnings**	2.21	3.73	2.05	1.67
	Employment	2.98	5.82	2.38	1.34

Another set of multipliers comes from a model known as the Regional Input-Output Modeling System (known as RIMS II). This model calculates regional economic impacts for 2001 using the U.S.

Census Areas. **Table 5** shows the results for the Y-K market area (i.e., Bethel and Dillingham) as compared to Anchorage.[10]

An examination of the final demand yields estimates of the total economic impact of air transport on dollars of output, dollars of household earnings and employment. For every extra $1 of air transport generated in the Y-K market area a further $1.27 to $1.35 of regional output is produced as a result. Also, each extra $1 is responsible for a further $0.34 to $0.36 in household earnings being generated. The employment figures, by RIMS II construction, show that approximately 9 to 10 jobs will be created in the Y-K market area for every extra $1 million of output generated by the air transport sector.[11] Because of the greater diversification of the Anchorage area it should not be a surprise that the earnings and employment multipliers are larger than those for Bethel and Dillingham.

An examination of the direct demand yields the impact of the air transport sector without consideration for any leakages in earnings from the region. Thus, the earnings multipliers are larger than they are for the final demand. For every $1 in extra earnings payable to air transport employees in the Y-K market area a further $1.29 to $1.36 is earned by *all* households in the area. Compared to Anchorage, however, the leakages are not as pronounced when comparing the final and direct earnings multipliers. Finally, the employment figures show that for every 1 job created in the air transport sector a further 1.52 to 1.62 jobs are created. As expected these figures are less than the 2.98 jobs created in Anchorage.

A survey was sent to the Capstone carriers listed (a copy of which is available upon request). Five surveys were returned (i.e., a 17% response rate). Even smaller was the number of carriers electing to reveal revenue figures. From these two, the aggregated total output derived from cargo and passenger transport exceeded $5.4 million. The average utilization of Capstone avionics within these fleets was 100%. Now if Capstone, by the lowball estimate (see Section 5), prevents 3/572 or 0.5% of these inbound flights from suffering accidents then about $27,000 in extra revenue would be injected into the Y-K Delta as a result those flights alone. In the aggregate, across

all participating carriers, it is likely the figure could be as much as 10 times as high.[12] Thus, with the possibility of $270,000 (or $0.27 million) in extra revenue, a further 3 jobs (i.e., 0.27 x [9.22 to 9.77]) in the Y-K Delta could be attributed to Capstone avionics, according to the employment multiplier. The output multiplier leads to an estimated increase in output of $0.34 to $0.36 million in the Y-K Delta (i.e., 0.27 x [1.27 to 1.35]). These figures give the total lowball estimated economic impact of Capstone avionics within the Y-K Delta.

With the highball estimate of 3.3% in accident prevention, $178,200 in extra revenue is injected into the Y-K Delta based on the survey; or an estimated $1.78 million over the Capstone carrier population. The multipliers, therefore, indicate the following estimates: about 17 extra jobs created and $2.26 to $2.4 million in extra output generated.

Of course, more survey data would add greater precision but revenue information is typically difficult to retrieve from businesses which are not publicly traded. As time passes in the Capstone regime a richer analysis is possible as will be discussed in the next section.

(8) Future Research

Economic impact analyses are static by nature. The multipliers used are valid for the time period in which the underlying relationships were quantified in the input-output model. The analysis must be updated from time to time in order to get a better picture of how an activity is or is not evolving. This is especially important since tracking benefits and costs over time will help determine if net benefits are growing and, if so, at a rate acceptable to stakeholders. Changes in technology, significant price and/or tax changes affecting key goods and services, and changes in the product mix produced in the market area will adversely affect the stability of the input-output model.

Only with a consistent methodology and measurement technique can economic impact analyses be compared across time

and across regions. As the Capstone technology continues in its deployment throughout Alaska, there exists a fine opportunity to monitor its effect socially and economically in Alaska in the long run.

Of course, if the goal of the Capstone Program is to provide the technology throughout all of Alaska the issue of intra-market leakages becomes less of a problem. A state-wide economic impact analysis will provide a better picture of the benefits to be had from Capstone technology.

The lowball and highball estimates applied only to actual accidents thought to be preventable. It was suggested that further flights (and, therefore, further economic activity) would be *induced* by having Capstone avionics on board. Properly tracking the growth of flight activity and Capstone deployment would necessarily raise the question: is Capstone avionics a statistically significant component of *all* flights occurring in the Y-K Delta and elsewhere in Alaska? To answer that question one needs more years of data so that this hypothesis can be tested within an econometric model determining flight activity.[13] The economic impacts discussed above do not, at this time, capture the effects of any Capstone-induced flights.

As noted, Capstone technology has been provided free of charge to those who volunteer to participate. It is likely that in the future aviators will have to pay for the technology on their own. At a cost of $15-20 thousand per aircraft it is uncertain how that would affect the economic impact as the technology becomes available throughout the rest of Alaska. Of course, it is also likely that per unit cost would drop as demand for the technology rises. The price elasticity of demand for the equipment is something to be estimated over time.

Cost benefit analysis could be performed on the question of the net present value of satellite GPS versus that for building the 200 proposed ground stations throughout the state. Cost of acquisition, operation and maintenance of each type of GPS will differ while the benefits will be about the same; that is, GPS coverage occurs

irrespective of the method of deployment. The stream of benefits may differ in terms of the operational life of stations versus satellites.

There may be incentives to deploy Capstone technology in Canada's northern territories for the purpose of facilitating reliable Alaska-Canada cargo trade. As **Prokop (2003)** discusses, there are incentives in place to leverage Anchorage as an international gateway directly into northern Canada. Since Capstone technology, by design, diminishes flight risk in remote areas, trade flows are less risky among the small carriers as well. Trade networks linking Alaskan villages with those in, for example, the Yukon Territory may be developed with Anchorage acting as a collection and distribution hub. Furthermore, as the proposed natural gas pipeline is built and run through Canada, there will be further U.S.-Canada trade links established.

(9) Conclusions

The Capstone Program represents a unique opportunity to improve aviation safety and efficiency in both the Y-K Delta market area investigated in this paper and throughout the rest of Alaska. As a new program it sets up a learning curve for many of the participating carriers. These carriers readily adopted the free technology and are beginning to receive the benefits this avionics package offers. Even if the package is eventually provided at a price it should not abate industry interest. This is based on the assumption that per unit cost would drop as demand for it increases throughout Alaska.

The region used for the Phase I roll-out is highly dependent on air transport and is, therefore, able to receive benefits from avoiding flight accidents and diminishing the risk of flying in formerly prohibitive weather. The effects of enhancing the reliability of air travel will likely have a positive effect on the quality of life in the remote villages situated in the Y-K Delta.

The economic impact methodology provided two estimates of the net benefits received from the deployment of Capstone technology. The various multipliers discussed in this paper described

the incremental net benefits estimated to accrue as a result of increased air travel. Specifically, the incremental net benefits to regional economic output, household earnings, labor income and employment were reported for the Y-K Delta. Furthermore, using an estimate of accidents avoided by Capstone avionics, in combination with financial information derived from an industry survey: 3 to 17 new jobs would be created in the market area and corresponding output would increase in the range of $0.34 million to $2.4 million. It was also noted that these impacts derived from accidents which would otherwise occur were it not for Capstone avionics. The positive economic impact would surely increase when an estimate of the extra flights induced by Capstone avionics can be calculated.

Twelve GPS ground stations are in operation in the Y-K Delta with a total of 200 estimated from deployment throughout Alaska as part of a state-wide coverage. This paper noted that there may be long run implications for using ground GPS versus making an early commitment to a satellite form of GPS. The benefits from Capstone would be conditioned by the initial cost differences and the stream of maintenance costs which are not readily known at this time. As such, a cost benefit analysis of the different ways of deploying Capstone throughout Alaska could be combined with a state-wide economic impact analysis. This would be a useful next step in assessing the overall business implications of this important and innovative technology.

Endnotes

[1] The word "Capstone" became the catchphrase for a program which drew together recommendations from representatives of government, academia and the aviation industry concerning aviation in Alaska. Capstone brings together many technologies under one program.

[2] ADS-B is not a radar system; however, it is approved for "Radar-Like Services". Radar systems send radio waves from the ground to bounce off of airborne targets which are sent back as a reflected signal to be read and interpreted by ground personnel. The ADS-B system relies on a satellite-based global positioning system (GPS) to fix the aircrafts position. ADS-B transmits digital signals through space regarding speed, altitude and flight course of the plane to both GBTs (ground-based transceivers) and to other planes. The information content is superior to radar and is available in real-time to all relevant parties on the ground and in the air. Also unlike radar, ADS-B can be used to monitor aircraft activity on the ground (i.e., on the runways and taxiways) and is less affected by topography (i.e., the mountainous terrain common to Alaska).

[3] Air taxi is a for-hire carrier operating on a non-scheduled, on-demand basis pursuant to 14 CFR 135 (safety regulations) or FAR Part 135 (economic regulations). The aircraft must have a maximum capacity of 30 passengers and a maximum payload of 7,500 pounds. A commuter carrier is a for-hire carrier operating like an air taxi except that it also performs at least 5 scheduled roundtrips per week.

[4] Because of the high cost of satellites for GPS at this time, the FAA decided to use a ground-based approach to support Capstone. Twelve ground stations (i.e., the ground-based transceivers) are active throughout the Y-K Delta (***q.v.*, Map 2**). The FAA estimates that 200 stations would be necessary to provide full coverage throughout Alaska. At this time, it is not known whether the GPS option would become cost effective before these stations are eventually built.

[5] As a substitute for the costly communication satellites (***q.v.***, ***supra*-note 4**) one of the 12 ground based transceivers receives ADS-B signals from the Capstone equipped aircraft and transmits it to the Anchorage-based Air Route Traffic Control Center (ARTCC). The ARTCC, via the transceivers, reciprocates with radar-like services to the aircraft; and it is important to note that this reaches planes where radar is non-existent.

[6] These numbers were derived from: *Small Air Carriers: Schedule T-1; Bureau of Transportation Statistics* (www.transtats.bts.gov). Small carriers are defined as those: holding a certificate issued under Section 401 of the *Federal Aviation Act*; and use aircraft with a seating capacity of 60 and payload of no more than 18,000 pounds.

[7] An input-output model identifies what each business sector in the region must purchase from every other business sector in that region in order to produce $1 in goods or services. These purchases may be traced forward or backward. Forward spending generates employee wages which induce further spending while backward spending on inputs generates income to those producers who, in turn, make purchases to create more inputs. In both cases, there are multiplier effects.

[8] Technically, these are known as SAM multipliers as derived from an IMPLAN computer-based input-output model. I am grateful *Northern Economics, Inc.* of Anchorage for generating the numbers for 2001.

[9] One may notice that the water transport multipliers are significantly higher than those for air, trucking and ground passenger transport. This should not be interpreted as a more productive form of transport. Simply put, this fact arises because the typical shipments in water transport are bulk and low value-added commodities. In other words, because there is more opportunity (or necessity) to perform value-added activity in the recipient region more jobs and rounds of spending are induced; hence the larger multiplier.

[10] **Map 1** shows the Bethel, Dillingham and Anchorage census areas. The RIMS II multipliers are available from the *Regional Economic Analysis Division, Bureau of Economic Analysis, U.S. Department of Commerce*. Empirical tests have shown that estimates of impacts based on RIMS II multipliers have been very close to those generated by more formal and expensive surveys (see **Brucker *et al.* 1990**).

[11] The relatively large employment multipliers for final demand in the ground transport sector are attributable to a peculiarity of the tourism sub-sector: labor intensiveness in the form of a relatively low driver to revenue-passenger ratio.

[12] The factor of 10 is arrived at by extrapolating over the entire population of carriers based on the survey sample's revenue data. From the small sample, it is assumed that the respondents are representative across the population. Of course, the survey was designed to gather enough information about aircraft usage and market network to make the estimates more reliable. It is the case that the Y-K Delta is a major source of revenue for these carriers.

[13] Several more years of data will be necessary because a valid statistical hypothesis test requires "degrees of freedom" which, in part, are obtained by having a stream of annual data to be fed into the model.

Bibliography

AirNav.com. (www.airnav.com).

Brucker, S., S. Hastings and W. Latham. (1990). "The Variation of Estimated Impacts from Five Regional Input-Output Models." *International Regional Science Review*. vol. 13. pp. 119-39.

Capstone Phase I Interim Safety Study. (2002). Institute for Social and Economic Research (ISER).

FAA Alaskan Region Capstone Program. (http://www.alaska.faa.gov/capstone/)

OAG Cargo Guide Worldwide. July 2004. Oak Brook, Il: OAG Worldwide.

Prokop, Darren. (2003). "Analysis of Anchorage as a Logistical Gateway to Canada's North." *Northwest Journal of Business and Economics*. pp. 33-41.

Regional Input-Output Modeling System (RIMS II). Regional Economic Analysis Division, Bureau of Economic Analysis, U.S. Department of Commerce.

Small Air Carriers: Schedule T-1; Bureau of Transportation Statistics Online. (www.transtats.bts.gov).

PUROLATOR COURIER'S "GREENING THE FLEET" INITIATIVE

Douglas J. Kube, QEP, CHSC
Director, Environment, Health, Safety & Security

Purolator Courier Ltd.
5995 Avebury Ave.
Mississauga, Ontario
Canada L5R 3T8

Phone: 905-712-1084
dkube@purolator.com

Copyright © 2004 Purolator Courier Ltd.

1. INTRODUCTION

Through it's "Greening the Fleet" initiative Purolator Courier Ltd. (Purolator) is aggressively and proactively leading the transportation industry's efforts to prevent pollution, meet and exceed current targets set by the Kyoto protocol with a substantial effort to introduce environmentally friendly ("Green") technology into its urban, delivery fleet of curbside vans. As this country's biggest courier company, Purolator is using it's fleet as a "test host" for the deployment of the innovative Green technology sets; and, to prove this Green technology for general commercial use in a fleet environment. Throughout the two phases of this initiative the new Green technology is being tested, demonstrated and evaluated from an environmental, technical and economic perspective to validate the commercial feasibility and sustainability of its application.

Purolator's stepped, integrated approach will allow it to analyze how the various individual components of the entire system interact and will enable Purolator to determine their individual and collective impact on the system as a whole. Furthermore, Purolator believes that by taking incremental steps (like the Japanese car manufacturers are currently doing) and analyzing all system components at each step along the way, it will be able to realize significant efficiencies, and emission reductions, in the implementation of Green technologies.

In Phase 1 of this project emissions generated from Purolator's curbside fleet will be reduced by over 50% with the introduction and adoption of diesel Hybrid Electric Vehicles (HEV). Phase 1 represents the first step towards the ultimate goal towards a zero-emission vehicle.

In Phase 2, the best battery and hydrogen fuel cell technologies are combined, the end result being a zero-emission vehicle. In this phase Purolator is building upon its HEV platform and the knowledge gained from its HEV program to leverage the development and demonstration of a hydrogen Fuel Cell Hybrid Electric Vehicle ("FCHEV"). In this phase Purolator will also explore and

demonstrate the logistics and economics of hydrogen generation and refuelling on-site for fleet operations. This bold and innovative initiative has the potential to generate significant benefits for Purolator, the industry and Canada; they would be as follows:

1) Demonstrate commitment to the Kyoto protocol,
2) Leadership stature for low emission and zero-emission delivery trucks,
3) Sponsor and incubate a sustainable hybrid electric and battery/fuel cell light truck industry,
4) Education and benefit to consumers, and
5) Significantly prevent and reduce the greenhouse gases (GHGs) generated from Purolator's fleet.

2. DESCRIPTION OF THE COMPANY

Purolator is the largest courier company in Canada, controlling a greater portion of the market than all its major competitors combined. As such, it is the dominant player in the Canadian market. Purolator has approximately 12,500 employees across Canada, of which more than 4000 are couriers, making Purolator one of the country's largest employers.

Purolator is distinctly Canadian. It offers express delivery to more communities in Canada than any other courier, shipping to more than 10,000 destinations. This vast Canadian network is a key advantage, helping Purolator maintain its leadership position in the domestic marketplace.

Purolator has an extensive network that delivers more packages to more destination points in Canada than any other courier. We handle more than 6.1 million pieces each week. To move the volumes that Purolator does, more than a million pieces each day, requires an extensive infrastructure of materials processing facilities, transportation vehicles and personnel. Purolator also boasts the largest ground fleet of any Canadian courier. It has over 3,700 courier vans, 390 medium trucks, 860 highway trailers, and 360 tractors. It also has the largest dedicated air express fleet in Canada –

with over 24 dedicated chartered aircraft and 686 pieces of aircraft support equipment. It moves more than 400,000 pounds of air freight each night via its 24 dedicated aircraft.

3. THE "GREENING THE FLEET" INITIATIVE

The main purpose of Purolator's Greening the Fleet initiative is to create and implement a bold vision and implementation schedule for moving the Purolator's fleet towards vehicles and equipment that reduce fuel usage, operating costs, air emissions and negative impacts on the environment and public health.

The main environmental objective of Purolator's Greening the Fleet initiative has been to prevent or minimize the creation of greenhouse gases targeted for reduction under the Kyoto Protocol. In so doing, other air contaminants will also be reduced which supports efforts of various levels of government to improve air quality and reduce the impact on the health of Canadians. In addition, an objective is to establish Purolator as an industry leader in promoting environmental solutions that are affordable, renewable and sustainable.

Purolator's research indicated that it would be possible to significantly cut fuel consumption and reduce GHG, and other, emissions in a cost effective way. When valued on a lifecycle basis, it could be possible to double fuel efficiency for light and medium trucks. Purolator also recognized that being a high profile courier company, it was in a position to exercise leadership and initiative in accelerating the deployment of energy efficient vehicles in Canada.

Purolator's findings coincide with those contained in a 2001 report entitled "Greening the Canadian Courier Fleet: Strategies for Improved Fuel Efficiency" – submitted to Greenpeace and Sierra Legal Defence Fund. In this report Torrie Smith Associates Inc. estimated this savings had a net present value in the range of $16,000 - $28,000 per vehicle for the medium weight vans, even without assuming any real increases in the future cost of fuel. Clearly there are both strong environmental and economic reasons for undertaking this initiative.

This initiative is a multi-phased approach towards the incorporation of environmentally friendly ("Green") technology into Purolator's vehicle fleet operations. Purolator believes that a stepped, integrated approach is key to the successful implementation of Green technology into its fleet operations. Such an approach will allow Purolator to analyze how the various individual components of the entire system (i.e., diesel engine, battery, electric drive, power storage, diesel generator, fuel cell module, hydrogen storage and delivery systems, etc.) interact and will enable Purolator to determine their individual and collective impact on the system as a whole. Purolator believes that by taking incremental steps (like the Japanese car manufacturers are currently doing) and properly analyzing all system components at each step along the way, Purolator will be able to realize significant efficiencies and GHG reductions in the implementation of Green technologies into its operations.

General Description of the Initiative

Through this initiative Purolator is aggressively and proactively striving to meet and exceed current Kyoto targets with a substantial effort to "Green" its urban, delivery fleet of curbside vans. To achieve its objectives Purolator has partnered with Azure Dynamics Corporation, Hydrogenics Corporation, Natural Resources Canada (Canadian Transportation and Fuel Cells Alliance Program) and Industry Canada (Technology Partners Canada's h2 Early Adopters Program).

The baseline for performance utility, emissions, durability and economics begins with the current diesel and gasoline powered curbsides. In Phase 1 of this project, emissions generated from Purolator's urban Hybrid Electric Vehicle ("HEV") curbside vehicle can be reduced by over 50% with the introduction and large-scale adoption of a HEVs. Phase 1 is expected to substantially reduce emissions allowing Purolator to make a significant step towards the ultimate goal of a zero-emission vehicle. In Phase 2, the best battery and hydrogen fuel cell technologies are combined resulting in a zero-emission vehicle. In this phase Purolator also explores and

demonstrates the logistics and economics of hydrogen refueling on-site for fleet operations.

The Hybrid Electric Vehicle's

Hybrid electric vehicles can significantly reduce fuel consumption and exhaust emissions compared to conventional vehicles, without limiting the performance characteristics of the vehicle and without the constraints suffered by other alternative fuel vehicles. Because of reduced engine load, regenerative braking and the reliability of electric motors, a reduction in maintenance requirements is also expected.

On October 12, 2002 Azure Dynamics Corporation ("Azure") delivered, to Purolator one 'Proof of Concept' trial gasoline hybrid electric vehicle - a Purolator van that Azure had converted to hybrid-electric power. The converted vehicle was a 1994, 4.9 litre, Ford/Utilimaster van with a 14 foot body and gross vehicle weight of 10,000 pounds. It now features an Azure hybrid electric power train incorporating a 2.0 litre gasoline engine, a 60 kW generator, a 60 kW traction and a 312 volt energy storage system.

This 'Proof of Concept' test vehicle was placed into parcel delivery service on three routes in Vancouver. These routes offered a mixture of urban and highway driving as well as variable terrain. During this field trial Purolator compared the test hybrid curbside van to a conventionally powered curbside van in terms of factors such as engine displacement, fuel consumption, emission, range (miles) and regenerative braking performance. Purolator's test results indicated that the only characteristic where the hybrid electric vehicle is not equal to or superior to the conventional vehicle is in high speed, a requirement that is superficial to urban delivery routes and which some might view as an additional safety feature of the hybrid electric vehicle. Following the tests in the Vancouver area, in 2003 this trial vehicle was also placed into parcel delivery service in Alberta to test it for cold weather suitability. Again, the HEV's performance was equal to the conventional vehicle.

Purolator built upon the learnings from the gasoline hybrid electric vehicle 'Proof of Concept' experiment. As the next step in this initiative, in late September 2003, Purolator signed a deal with a subsidiary of Vancouver-based Azure Dynamics Corporation ("Azure") to introduce 30 'second generation' diesel Hybrid Electric Vehicles ("HEVs") into its fleet in late 2004/early 2005. As of November 26, 2004 the first of these diesel HEVs was on the road delivering packages for Purolator.

These 30 HEVs will provide a demonstration of full-blown fleet economics and will prove the HEVs' consistency and reliability in a fleet environment. The field trial phase is planned to last six months, and will therefore end in September 2005. At that point, field trial data with respect to fuel economy, operating costs, reliability, etc. will be assembled and reviewed by Azure and Purolator.

The Zero-Emission Fuel Cell Hybrid Electric Vehicle

In this phase Purolator is building upon its HEV platform and the knowledge gained from its HEV program to leverage the development and demonstration of a hydrogen Fuel Cell Hybrid Electric Vehicle ("FCHEV"). For this phase of the initiative Purolator has partnered with Hydrogenics to develop a "World's First" - a complete environmental refueling and vehicular platform.

The FCHEV system will consist of a hydrogen fuel cell power module, ultra capacitors, and high capacity battery storage system optimized to deliver and manage the propulsion system. As this project is based on the HEV platform, it will involve replacing the diesel genset with the fuel cell – which will work in combination with high capacity batteries and ultra capacitors to provide all of the vehicle's power requirements. The integration of a fuel cell to replace the diesel genset in the HEV is the second technological step essential in the validation of the FCHEV platform.

Also, this program will demonstrate on-site refueling of hydrogen and its role in the future hydrogen infrastructure as it pertains to fleet operations. As indicated, combining the FCHEV with on-site

hydrogen production, storage and refueling in a commercial fleet setting, is a "Global First". For this project, Hydrogenics' used its fuel cell Proton Exchange Membrane (PEM) expertise to develop a PEM electrolyzer platform that produces hydrogen via water electrolysis. In this platform conventional city water is conditioned and used as the fuel to produce hydrogen.

This FCHEV is being deployed and tested on a downtown Toronto route and will be maintained and operated out of our "Metro West" site. In addition, Purolator's hydrogen refueler (electrolyzer) has been assembled and is currently using "off-the-grid" power to produce hydrogen.

4. TRANSFERABILITY WITHIN THE INDUSTRY

This project has a wide range of potential applications within the transportation sector, as light-to-medium sized duty trucks are used by the courier, mail delivery, food and beverage industry, and service contracting industry. In 1997 Environment Canada has estimated that commercial trucks were responsible for 27.2% of Canada's transportation related GHG emissions.

The global market size for light and medium duty trucks, Class 2-6, is estimated to $ 52.7 B Canadian. The growth rate varies between 3-15% depending on the segment. The light duty truck urban delivery truck segment is growing at approximately15% per year. By introducing low-emission and zero-emission technology sets into the truck market, we would be focusing on the largest percentage contributor to transportation related emissions.

Purolator's extensive distribution and fleet operations provides the ideal setting to demonstrate and replicate the low-emission and zero-emission technology sets. In the case of the FCHEV, additional multiplier effects are envisioned as the power generation capabilities of the vehicle while parked, can provide stationary, back-up "generator", as well as "peak-shaving" capabilities to the sorting depots. Furthermore, throughout Purolator's extensive distribution network there are numerous functions that may be amendable to this

Green technology, such as: material handling, stationary back-up generators, forklifts, utility work vehicles, aircraft support equipment, etc. The experience and knowledge gained from the integration and operation of the HEVs and FCHEV vehicles will allow this technology to be assessed vis-à-vis current technologies.

5. INNOVATION

This initiative is visionary and "game changing" as the deployment of these innovative technology sets will set the stage for global leadership in the commercial fleet environment. For example, the zero-emission curbside delivery van will involve the integration of hydrogen-compatible technologies, which will combine current "state of the art" fuel cell technology, batteries and ultra-capacitors. The development of the hydrogen infrastructure via the on-site hydrogen production will demonstrate the viability of hydrogen as a safe, economic and non-polluting fuel in a commercial fleet environment.

This project will advance fuel cell science and technology through analysis and implementation of the following:

i) low temperature start-up and operation,

ii) hybridization of fuel cell power module, ultra capacitors and or batteries,

iii) optimized power electronics and management,

iv) high efficiency hydrogen production via electrolysis and 'fast fill' refueling station, and

v) industry and consumer education on the handling and use of hydrogen fuel cell technology.

Furthermore, codes and standards established for the operation of the vehicle, production, storage, safe handling of the hydrogen will be established and serve as a template for other industries as they begin to adopt fuel cell technologies and embrace the hydrogen economy.

Additionally, the deployment of these innovative technology sets will not only provide a sound and viable roadmap for advancing Canada's commitment under the Kyoto Protocol; but also it could potentially

spawn an entire new hybrid light truck industry for Canada. Job creation and training associated with this project will serve to attract and retain skilled workers to support this sector to thereby further Canada's economic growth.

Purolator's extensive distribution and fleet operations provides the ideal setting to deploy of these innovative technology sets providing at the micro level, a reliable and replicable model for other companies to follow. Additional multiplier effects are envisioned as the power generation capabilities of the vehicle while parked, can provide stationary, back up and peak shaving capabilities to Purolator's sorting depots.

6. GREENHOUSE GASES REDUCTION

Purolator operates one of the largest and most recognizable vehicle fleets in Canada. Currently it is estimated that, in the year 2001, the Purolator fleet consumed in excess of 20 million litres of gas and 12 million litres of diesel and emitted approximately 100 thousand metric tones of GHG. This "game changing" initiative by Purolator demonstrates their environmental commitment and as such will influence and significantly impact the truck industry and other major GHG contributors to become engaged in similar initiatives.

This initiative delivers three very significant societal and business benefits for Canada. Firstly Purolator's "Greening the Fleet" Initiative will demonstrate the ability for HEVs to reduce GHGs and other contaminants of concern by 50%; and, in the case of the FCHEVs, eliminate 100% of GHGs and other contaminants of concern currently emitted with conventional gasoline/diesel delivery vehicles.

In terms of the HEVs specifically, if 2,000 commercial HEVs were on the road rather than current standard vehicles, substantial reductions in emissions and fuel use would occur annually: Smog-causing emissions (nitrogen oxides) would be reduced by 340 tons, the equivalent of taking passenger cars off New York City roads for 5 days. Carbon dioxide emissions would be reduced by 16,600 tons,

which is equivalent to planting 400,000 trees. Diesel fuel usage would be reduced by 1.4 million gallons, which requires 200,000 barrels of crude oil to produce.

Based on sales of 200-300 FCHEVs per year, this translates into a GHG reduction of 12,500 tonnes per year by 2008 and a cumulative reduction of approximately 118,500 tonnes of GHGs by 2012.

Secondly, the deployment of these innovative technology sets will provide a future roadmap and economic justification for the sustainability and commercial viability of this technology.

Thirdly, the production of hydrogen via a higher efficiency electrolyzer provides a real-world production process that produces no GHGs versus existing steam methane reformer technology. These two technology demonstrations will validate and provide the roadmap for advancing Canada's commitments under the Kyoto Protocol.

Additionally, the utilization of electricity from renewable sources such as hydroelectric or wind turbine power for the electrolyzer, could provide a complimentary renewable platform for future developments and demonstrations. Combining these two technology sets will provide a world-class demonstration that validates hydrogen compatible technologies. Furthermore, by launching a project of this nature in the truck industry, Canada will be focusing on the largest percentage contributor to transportation related emissions.

7. CONCLUSION

By launching a project of this nature in the transportation industry, we will be focusing efforts on the largest percentage contributor to transportation related emissions. NRCan has estimated that commercial trucks are responsible for 27.2% of Canada's transportation related GHG emissions. This innovative, "game changing" initiative will demonstrate the feasibility for HEVs and FCHEVs in the commercial fleet environment.

Purolator believes that by taking incremental steps (like the Japanese car manufacturers are currently doing) and analyzing all system

components at each step along the way, it will be able to realize significant efficiencies, and emission reductions, in the sustainable implementation of Green technologies.

One of the innovative aspects of this project is that it uses the fleet environment to roll the Green technology into commercial transportation markets. Currently, one of the hurdles to the commercialization of some of the Green technologies encompassed by this initiative is the lack of hydrogen infrastructure. Fleet operations are ideal, given their base operations, for the introduction of hydrogen and for the evaluation and demonstration of on-site hydrogen refueling.

Purolator's HEVs are predicted to eliminate 50%, and FCHEVs are predicted to eliminate 100%, of greenhouse gases currently emitted with conventional gasoline/diesel delivery vehicles. In terms of the HEVs specifically, if 2,000 commercial HEVs were on the road rather than current standard vehicles, substantial reductions in emissions and fuel use would occur annually.

Purolator is confident that this bold and innovative initiative can generate significant benefits for Purolator, the transportation industry and Canadian society in demonstrating commitment to the Kyoto protocol. Purolator's strong visionary role in championing "low-emission" HEVs and "zero-emission" FCHEVs illustrates our commitment to reducing emissions of green house gases, pollution prevention, promotion of the HEV industry, fuel cell industry and hydrogen economy.

A KYOTO PROTOCOL UPDATE
Opportunities and Challenges for Canadian Freight Rail

Dr Malcolm Cairns
Director, Business Research, Canadian Pacific Railway

It has been over two years since the first Climate Change Plan for Canada was released by the federal government, and on February 16, 2005 the Kyoto Protocol came into force following the ratification by Russia in November 2004. On ? the federal government announced changes to its plan...

This paper will provide an informational update on the government initiatives as they might relate to freight transport, and review the opportunities and challenges applicable to Canadian freight railways.

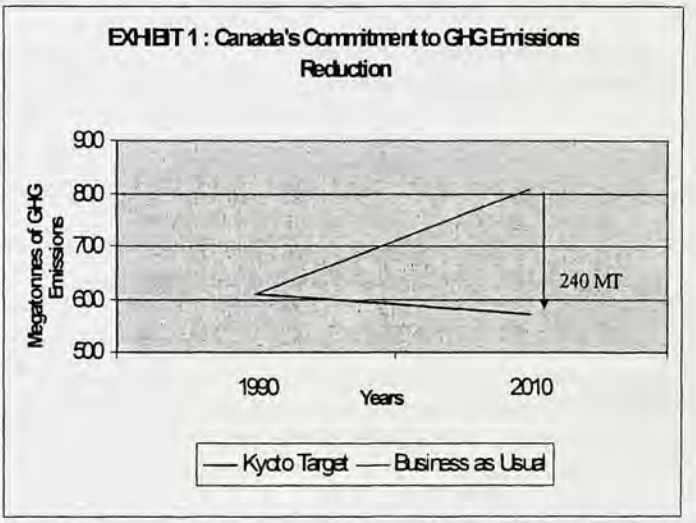
I BACKGROUND

On November 21, 2002 the federal government released its Climate Change Plan for Canada to reduce greenhouse gas (GHG) emissions to 6% below 1990 levels during the period 2008 to 2012. As indicated in Exhibit 1 this would require a reduction of 240 megatonnes (MT) from the business as usual scenario.

Canada's approach includes:

- No region bearing an unreasonable burden;
- Minimizing mitigation costs while maximizing benefits;
- Promoting innovation; and
- Limiting uncertainties and risks;

and is consistent with the need to continue to pursue recognition for



Canada's exports of cleaner energy, and the necessity of recognizing early action by industry. The plan involves three steps: first, there are actions underway to address an 80 MT reduction; second, there are new actions for a further 100 MT reduction; and thirdly, there are a number of potential actions to enable the remaining 60 MT reduction.

Details of the first two steps are presented in Exhibit 2 – note that the estimated reductions are slightly less than the above figures before rounding – and the actions that have bearing on Canadian freight railways are shaded. Exhibit 3 outlines examples of potential actions to close the remaining 60 MT gap.

The various actions may be characterized as involving five specific instruments:

- Innovation and Technology investments;
- Infrastructure investments;
- A Partnership Fund – for co-investment with other parties;Covenants and Emissions trading by industry; and

EXHIBIT 2 : PART 1

ACTIONS UNDERWAY	Business as Usual (megatonnes of GHG)	Step 1 Reductions	NEW ACTIONS	Step 2 Reductions (megatonnes of GHG)	Total Reductions
TRANSPORTATION					
Negotiations for 25% improvement in new vehicle fuel efficiency by 2010		52	Consumer action to improve vehicle efficiency, including off-road vehicles	08	
Increase ethanol production to enable blending in 25% of gasoline		08	Increasing ethanol blending to 35% in gasoline, and targeting 500 million litres of biodiesel in use by 2010	2	
Development and demonstration of refuelling tecnologies and infrastructure for commercialization of fuel cell vehicles		01	Increased use of public transit, alternative approaches to passenger transportation and sustainable urban planning	7	
Reduction in urban transportation emissions		08	More efficient goods transport, including intermodal	23	
Negotiation of voluntary agreements with air, rail, truck and marine sectors to improve fuel efficiency of goods transport		2			
SUB-TOTAL	206	9		12	21
INDUSTRIAL EMITTERS					
Large Industrial Emitters					
Action Plan 2000 programs for Large Emitters		11	Target for emission reduction through covenants with a regulatory or financial backstop	55	
Renewable Energy and Innovative Projects					
Incentive for wind power production		28	10% of new electricity generation from renewable sources	39	
Green power purchase of 20% of governments electricity		02	Identify and develop options to address impediments to new regional hydroelectricy		
Reduction in barriers to interprovincial electricity trade and transmission		54	Demonstrate clean coal technology	45	
CO_2 capture and storage initiative		35	Establish a CO_2 capture and storage pipeline	22	
Sub-Total		12		11	
Smaller Sized Firms and Fugitive Emissions					
Improved energy efficiency in all other sectors		2	Extend voluntary targets to smaller sized firms	07	
			Continuation of Industrial Research Assistance Program		
			Further work to reduce fugitive emissions	4	
Sub-Total		2		5	
SUB-TOTAL	425	25		70	95

EXHIBIT 2 : PART 2

ACTIONS UNDERWAY	Business as Usual (megatonnes of GHG)	Step 1 Reductions (megatonnes of GHG)	NEW ACTIONS	Step 2 Reductions (megatonnes of GHG)	Total Reductions (megatonnes of GHG)
HOUSING AND BUILDINGS					
Making existing buildings more energy efficient		1.2	R2000 energy efficiency for all new housing by 2010	0.7	
Energy efficient evaluations for homeowners		0.7	Energy efficiency retrofit of 20% of housing by 2010	1.5	
Improved standards for equipment and appliances		1.6	New commercial buildings to better Code by 25% by 2010	0.4	
Energy efficiency improvements to federal buildings		0.2	Energy efficiency retrofit of 20% of commercial buildings by 2010	1.2	
SUB-TOTAL	**84**	**4**		**4**	**8**
AGRICULTURE, FORESTRY, LANDFILLS					
Promotion of sinks through Action Plan 2000		5.8	Enable new sink credits to be sold in emissions trading		
Green Municipal Funds for landfill methane capture		2.2	Consider regulation or emissions trading of new landfill gas		
Business as Usual agriculture and forestry sinks		30			
SUB-TOTAL	**94**	**38**			**38**
INTERNATIONAL EMISSIONS REDUCTION					
$15 million investment in World Bank's Prototype Carbon Fund		2	Consult private sector in support of their investments and purchases on the international market		
			Consider purchase of a minimum 10MT of permits by federal government	10	
SUB-TOTAL	**N/A**	**2**		**10**	**12**
TOTAL	**809**	**78**		**96**	**174**

- Targeted measures such as information requirements, incentives, regulations and tax measures;

any of which might, in principle, impact Canadian freight railways.

EXHIBIT 3 THE REMAINDER	Potential Step 3 Reductions (megatonnes of GHG)
Partnership Fund for cooperation, as well as Infrastructure funding	20 - 30
Existing and future R&D investments that reduce emissions	10
Provincial/Territorial actions not involving the federal government	10-20
Community-wide emissions reductions plans by 100 municipalities	10
The one tonne challenge from Rick Mercer beyond Step 2	7
Credits for cleaner energy exports	Up to 70

The economic impact of this plan has been estimated by the federal government under different assumptions:

- An international carbon price in 2010 of either $10 or $50 per tonne of CO_2, representing an expected value and an extreme value;

- Two distinct financing options – tax financed or government financed. In the first option, personal income taxes are increased to maintain government balances unchanged. In the second option, the fiscal impacts of the plan directly affect government balances with the tendency to slow the paying down of the public debt.

The federal government views the $10 carbon price–government financing scenario as the more likely, and the overall impact of this scenario by 2010 is estimated as:

- GDP is reduced by 0.4% - rather than growing by 18% between 2002 and 2010, the economy would grow by only 17.6%;

- Employment would be reduced by 600,000 – rather than increase by 1.32 million jobs, employment would increase instead by only 1.26 million;

- Disposable income per household is unaffected;

- Gasoline prices remain unaffected, while natural gas prices increase in the range 4% to 8%, and electricity prices decrease in the range 2% to 9%.

In Exhibit 4 are presented more detailed estimates of the impacts by 2010 on output by individual sectors – and range from a decrease of 3.1% for refined petroleum to an increase of 0.9% for motor vehicles (due to urban transit measures stimulating vehicle demand). The conclusions drawn by the federal government are as follows:

"The main message from all of these numbers is that the competitiveness issues of the heavy emitting industries appear to have largely been addressed by the design of the domestic emissions trading system. The impacts on output in all cases are relatively small, particularly so in the $10 scenarios. In these scenarios, the effects on the oil and gas industry are modest. Only in the $50 scenarios are the impacts on this industry discernible. Among manufacturing industries, only cement and petroleum refining are significantly affected. For the former, the results reflect the particularly carbon-intensive production process. For petroleum refining, the relatively large impact relates to increased vehicle efficiency and the maintenance of imports at the expense of domestic production."

EXHIBIT 4

SECTOR	Percent share of GDP	Percent change in output
Energy Suppliers and Energy-Intensive Sectors		
Metal Mining	0.7	-0.2
Non-Metal Mining	0.2	-0.6
Pulp and Paper	1.0	+0.2
Primary Iron and Steel	0.6	+0.4
Primary Non-Ferrous Metals	0.7	-0.3
Motor Vehicles	2.5	+0.9
Cement and Clay Products	0.3	-2.4
Refined Petroleum	0.3	-3.1
Industrial Chemicals	0.4	-0.3
Oil and Gas	2.7	-0.4
Electricity	2.3	-2.8
Coal	0.2	-0.9
Consumer Goods and Services	28.2	-0.1
Construction	4.4	-.3
Transportation and Storage	4.6	-0.7
Communications	6.8	+0.2
Agriculture	2	-0.8
Finance, Insurance and Real Estate	15.5	-0.5
Government and Social Services	14.3	+0.1
Sub-Total	**88**	**-0.4**

2 ISSUES IMPACTING CANADIAN FREIGHT RAIL

CPR and CN are not large industrial emitters according to the government, and not subject to covenants for emissions trading – it is understood that the need for continued rail competition with trucking (which clearly does not involve large emitters) precluded designating CPR and CN in any event.

While transportation accounted for 26% of Canadian GHG emissions in 2002 according to Environment Canada, railway emissions accounted for only 3.1% of transportation, and so railways do not

represent a significant target for emissions reductions[1]. Moreover, largely due to the acquisition of over 725 fuel-efficient mainline freight locomotives over the past few years, railway GHG emissions have declined in absolute terms between 1990 and 2002 by 0.5% - despite a 28% increase in traffic as measured either by RTMs or GTMs[2]. While this does not yet represent a 6% reduction from 1990 levels there are still up to 7 years for targets to be met.

Future issues and developments impacting Canadian freight rail are of two categories – those directly involving the rail carriers themselves, and those involving rail shippers and therefore indirectly impacting the railways. These items will be addressed in turn

3 RAIL CARRIER ISSUES

Operational Improvements

There are a number of operational initiatives underway at the railways that will improve fuel efficiency:

- *Co-production Agreements*: Several co-production agreements are being implemented by CPR and CN, including operational changes involving directional running and a reduction in circuity that will imply improvements in fuel efficiency;

- *Train handling*: The proportion of main line locomotives fitted with dynamic brake equipment continues to increase. This allows the increased use of the dynamic brake for control of train speed rather than the use of the air brake

[1] "Canada's Greenhouse Gas Inventory" 1990-2002", August 2004, Environment Canada. This is the official source of data for UN purposes.

[2] "Locomotive Emissions Monitoring Program 2002", Environmental Protection Series, Environment Canada.

system. As the latter does not allow the locomotive engineer to reduce the severity of a brake application already in force, it is frequently necessary to apply power at the same time as the brakes to maintain speed over variable track grades. This causes the fuel consumption to be increased significantly. When the dynamic brake is used to control speed the severity of the application can be varied at will, and the fuel consumption is reduced. As well, the railways have on-going training programs that focus on awareness and on the importance of fuel conservation practices, including locomotive shutdown in yards.

- *Rail lubrication*: Efficient rail gauge face lubrication has been shown in many tests to reduce fuel consumption. The railways have on-going programs to ensure that the system of track mounted rail lubricators is maintained in good operating condition. Railways which have applied on-board locomotive wheel flange lubricators also have programs to keep these in working order;

- *Freight car productivity*: The maximum allowable axle load has been increased on many lines in Canada. This enables the railways to use certain cars with a gross weight on rail of up to 129,800 kg instead of 119,400 kg. The gross to tare ratio of such freight cars is increased so that the quantity of gross tonne-kilometers accumulated to move a given amount of freight is reduced, contributing to the improvement in the ratio of RTK to GTK. The railways have also invested in aluminum cars to replace the heavier steel cars. All these productivity improvements reduce the quantity of fuel required to move a given volume of traffic.

Locomotive energy efficiency

Initiatives are also underway to improve the energy efficiency of locomotives thereby reducing GHG emissions:

- *Continuing fleet renewal*: Canadian railways are progressively renewing their fleets by purchasing EPA compliant locomotives. As well, they are upgrading, their pre-2000 locomotives to Tier 0 upon overhaul. These costly measures will ensure emissions, particularly NOx, will continue to be reduced per unit of work. Locomotives meeting U.S. EPA Tier 2 will be available by 2005. Tier 2 locomotives will have NOx emissions 25.7 percent lower than 2001 locomotives and 59.2 percent lower than pre-2000 locomotives. It is expected that Tier 3 locomotives will focus on lower particulate emissions. These modern locomotives – as well as reducing Criteria Area Contaminants (CACs) such as NOx and particulates - are also more energy efficient than the older locomotive fleet;

- *Low idle applications*: The railways are extending the application of the "Low Idle" feature to more mainline locomotives. This feature allows the diesel engine to idle at a reduced speed with a consequently reduced load from fans. The reduction in fuel consumption can be as much as 10 L/hr, and on the accepted duty cycles can be as much as 3 percent of the annual fuel consumption. The use of the low idle feature is limited in some cases by the ability of the auxiliary power system to generate sufficient power for battery charging. However a continued reduction in overall fuel consumption is expected from this feature.

- *Automatic Start/Stop systems*: Railways are installing devices on both mainline and switching locomotives that will automatically shut down and restart the diesel engine when the locomotive is not in use. The device is regulated by several locomotive system parameters such as water temperature and battery condition. It will restart the engine to idle for a time to prevent freezing and to charge the batteries. The railways now have a policy of shutting down unused engines when ambient temperatures permit;

automatic start/stop systems will allow this practice to be extended all year.

- *Hybrid locomotives*: Railpower Technologies Corporation has developed and tested its 2,000-hp Green Goat hybrid switcher locomotive at CPR, CN and most of the US Class 1 railroads. Tested under various operational and environmental conditions, the results so far have indicated a 40% to 70% reduction in GHGs and an 80% to 90% reduction in NOx and particulates, compared to conventional switchers. By December 2004, Railpower had received firm and conditional orders for 78 of these locomotives to be delivered between 2005 and 2008. On March 14, 2005 CPR announced the purchase of 35 Green Goat hybrids – seven for 2005 and the balance over the following three years;

- *Fuel cell locomotives*: Despite a measure of skepticism[3], fuel cell locomotives are under development by an international consortium, including the US Department of Defense. The five-year project, which commenced in May 2003, will develop and demonstrate a 109 tonne, 1megawatt fuel cell powered locomotive for military and commercial railway applications. Fuel cells are solid-state devices that convert the energy of a fuel into electric power. Based on electrochemistry rather than combustion, they are efficient, quiet, and have zero emissions. Potential commercialization paths and follow-on demonstration projects include subway utility locomotives, switchers, commuter rail, subway trains, light rail, heavy freight, and high-speed rail, according to the consortium.

---

[3] "Participation in the Hydrogen Economy: A Transportation Perspective", CTRF, May 2003. According to the paper, page 763: "Development of fuel cell propulsion for the rail industry is not foreseeable in the near future".

Developments in Railway Fuel

The railways use diesel fuel that complies with the existing engine builders requirements of an average sulphur content of no greater than 5,000 parts per million (ppm). In general, Canadian railways use fuel with a much lower Canadian average sulphur content of approximately 1,500 ppm. While sulphur content does not directly impact GHG emissions, reductions in sulphur content does allow better use of technology to reduce NOx, and NO_2 emissions, that accounted for 11% of the total GHG emissions by railways in 2002. Therefore plans to further reduce the sulphur content of rail diesel fuel in Canada will indirectly reduce the 11% of GHGs accounted for by NO_2.

Despite the government target of 500 million litres of cleaner-burning biodiesel in use by 2010, the current locomotive manufacturers warranties preclude the use of biodiesel in rail locomotives, and so this development is unlikely to impact rail GHG emissions over the medium term.

Modal Shift

Opportunities to shift goods traffic between modes in order to reduce GHG emissions is generally seen in the context of shifting medium and long haul truck traffic from highways to rail intermodal. The railways have advocated for years in favour of "leveling the playing field" between truck and rail – changes to input taxes, capital cost allowances, recovery of highway costs from users – but with little or no success to date.

In addition, in the absence of a recognition by government of the public benefits to the use of railway operations – reduced highway expenditures and traffic congestion, fewer highway accidents, reduced GHG emissions, benefits that represent economic externalities to railways – CPR has chosen to suspend its innovative Expressway intermodal service between Toronto and Detroit in

favour of traffic with higher contributions in this capacity-constrained rail corridor.

Despite this lack of government action, Transport Canada has embarked on two research projects that are encouraging:

- Full Cost Accounting Project: This multi-year project to investigate the full costs of each transportation mode in Canada is intended to tackle, in part, two issues of particular interest; the estimation of the social costs or negative externalities associated with rail and truck; and the allocation of highway costs to the range of highway users and heavy duty trucks in particular. Whether the eventual results from this effort will enable changes in public policy remains to be seen;

- Canadian Railways and the Taxation Regime, and the Environment: This most recent research project is focussed on the relationship between taxation and investment, including the ability of railways to invest in more environmentally sustainable infrastructure and equipment. Again the eventual implication of the findings for changes in public policy remain unclear.

Furthermore, there has been recent interest expressed by shippers in the extent to which the use of rail versus truck reduces GHG emissions.

In one case, an intermodal shipper concerned about its corporate image wanted to know the extent to which the use of rail reduced the emission-equivalent truck traffic. Using rail emission factors from the RAC and truck emission factors from Environment Canada, it was estimated that with the average train consist and fuel consumption between Toronto and Calgary and the average fuel efficiency of heavy-duty diesel trucks over the 3,230 kilometres of highway, that the equivalent trucks would emit 2.6 times the volume of GHGs. Expressed the other way around, the annual 3,500 truckloads that this

shipper moved instead by rail represented the emission-equivalent of 1,350 truckloads moving over the highway – saving the equivalent of 2,150 truckloads that would have otherwise created emissions.

In another case, using the same methodology, it was estimated that moving an annual volume of 400,000 tons of steel by rail from Hamilton to Windsor would save 3,150 tonnes of GHG emissions compared with the movement by truck, and that at a price of $10 per tonne of CO_2, an emission credit of $31,500 would be generated from a modal shift of traffic from truck to rail.

Government Initiatives

Transport Canada has two financial initiatives to reduce GHG emissions from freight transport:

- The Freight Sustainability Demonstration Program (FSDP), a five-year $4.5 million program that is part of its Freight Efficiency and Technology Initiative; and

- The Freight Incentives Program (FIP), a four-year $5 million program that is part of its Freight Efficiency Program.

Exhibit 5 overleaf summarizes the rail projects to date in FSDP, many of which have been identified already. The FIP provides funding to enable eligible applicants to purchase and install efficiency-enhancing technologies, preferably those currently being demonstrated under the FSDP, but there has been no allocation of funds at time of writing. Late in 2004, CPR proposed two new projects for consideration under FSDP to generate fuel conservation and emission reductions: development of a locomotive distribution tool, and demonstration of a process of behavioral modification to improve fuel-efficient train handling methods by train crews.

EXHIBIT 5

Freight Sustainability Demonstration Program (FSDP)

Company	Federal Funds	Project
Kelsan Technologies	$225,000	Top of rail friction control
Athabasca Northern Railway	$48,000	Installation of diesel driven heating systems to reduce idling time
Southern Railway of BC	$25,000	Demonstration of SmartStart to reduce idle time
CN	$198,000	Demonstration of three different anti-idling technologies
CN	$250,000	Demonstration of the effectiveness of using distributed power
Railpower	$250,000	Demonstration of the hybrid switching locomotive
New Brunswick Southern Rail	$60,000	Demonstration of the economic feasibility of new intermodal service
Total	$1,056,000	

Freight Incentives Program

The Freight Incentives Program (FIP) is a four-year, $5 million program designed to encourage the adoption and use of technology and equipment that reduce greenhouse gas (GHG) emissions cost-effectively

Environment Canada is currently negotiating with the railways through the Railway Association of Canada (RAC) to update the Memorandum of Understanding (MOU) between the RAC and Environment Canada regarding locomotive emissions monitoring. It is intended that the new MOU come into effect on January 1, 2006 for a 10-year period, and that the MOU will establish voluntary performance targets for locomotive emissions of GHGs and CACs. Details of these targets are as yet undecided, but CPR, CN, VIA and Go Transit will have to develop and share Action Plans on how they will achieve the targets.

4 SHIPPER ISSUES

The impact of the Climate Change Plan for Canada will also affect rail shippers directly in their businesses and therefore indirectly impact the railways. Exhibit 4 indicates that sectors negatively impacted by the plan, and which ship by rail, include mining, cement manufacture, agriculture, petroleum, and chemicals – the extent of

impact on output ranging between -0.2% to –3.1%. If the extent of these impacts are to be believed, then the overall drop in demand for rail transport is not especially significant. It should be noted however that any increase in costs to the shippers will put pressure on the negotiation of rail freight rates.

The decision by the US government not to ratify the Kyoto Protocol also implies that there will be further competitive pressure from US businesses competing with Canadian businesses. The Canadian government has indicated that a number of US States are taking steps to reduce GHG emissions – including Oregon, Massachusetts, New Hampshire, and California – which should alleviate competitiveness concerns. Canada is on record that "we need not – and we will not – export Canadian jobs in order to meet our climate change goals".

There is one particular issue that continues to be of concern to railways. CN, and CPR in particular, export very significant quantities of metallurgical coal from Alberta and BC to Japan and north Asia for use in steel-making. According to the Kyoto Protocol, the GHG emissions from the burning of this coal are counted as emissions from the Asian countries not Canada. However, Canada continues to pursue recognition for Canada's export of cleaner energy – natural gas – to the US. To date this recognition has not been forthcoming, but in the event that this strategy is successful, then the logical corollary is that Canada should be debited for the export of dirtier energy – coal – to Asia. Such a change could have a significant negative effect on the coal industry and the railways that move the product to market, but the likelihood of such an outcome is presently unknown.

To a lesser extent, the railways move thermal coal from western Canada to Ontario for electricity generation. The intention of the Ontario government to phase out coal-fired generating stations by 2008 to help meet climate change commitments will also negatively impact railways, but the extent of the impact as indicated in Exhibit 4 for the coal sector does not appear very significant.

5 RECENT FEDERAL BUDGETARY INITIATIVES

In the February 23, 2005 federal budget it was stated that "the budget makes significant investments aimed at reducing GHG emissions, and developing science and technology that will help build the framework for a comprehensive and sustainable response to the challenge of climate change". Beyond that "hot air", the budget allocated $5 billion over 5 years (with more than two-thirds of the money to be spent beyond 2007) distributed among specific provisions including:

- $1 billion over 5 years for the Clean Fund to encourage cost-effective projects and actions that reduce GHG emissions;
- $250 million to create a Partnership Fund with provinces/territories; and
- An estimated $295 million through a tax incentive program, under which companies will be able to write-off half of their investments in environmentally-friendly equipment and technology such as fuel cells.

What does all this mean? In the words of one media commentator[4] on the budget:

"All of which is expressed in that unfathomable dialect of euphemism that is the real official language of Ottawa, a sing-song patois of government "commitments" to this and "initiatives" in support of that, as if it were their money. Take this gem from the section on Kyoto, where the government is still struggling to come up with a plan to reduce Canada's GHG emissions, eight years after it committed itself to reduce them. The previous grab-bag of ditsy schemes having failed utterly, at huge expense, the budget as much as announces it has gone back to the drawing board. But listen to how they say it. "The government of Canada will learn from past investments". I had grown accustomed to hearing "investments"

[4] Andrew Coyne, National Post, February 24, 2005.

used as a euphemism for spending, but I had not heard it used before to mean "mistakes".

While not all commentators were quite so cynical, there was concern expressed about the possibility of a carbon tax applied to the users of energy generated from fossil fuels. However, all that the budget indicated was that "opportunities to use the tax system to advance environmental goals will continue to be actively considered". It appears that the potential implications will have to await further announcements of federal policies.

6 LATEST DEVELOPMENTS AND THIER IMPLICATIONS FOR FREIGHT RAIL

At budget time it was stated that further details on the broader Kyoto plan would follow "within weeks". However, with government "dithering", no announcements have been forthcoming before the March 21st deadline for papers to be submitted for inclusion in these conference proceedings. In the meantime, media rumours circulate that the budgetary requirements might exceed $10 billion rather than the $5 billion allocated in the February 23rd budget, and that Canada may need to raise its emission reductions to between 270 and 300 MT rather than the earlier figure of 240 MT.

Further analysis will be presented at the conference if available by that time.

Climate Change, Productivity Shocks and Demand for Freight Ground Transportation in Atlantic Canada

Yuri Yevdokimov, Assistant Professor, Departments of Economics and Civil Engineering, University of New Brunswick, P.O. Box 4400, Fredericton, NB E3B 5A3, yuri@unb.ca

Oleg Byelyayev, MSc Candidate, Department of Civil Engineering, University of New Brunswick, P.O. Box 4400, Fredericton, NB E3B 5A3, oleg.b@unb.ca

Introduction

While most of the existing studies are dedicated to the impacts of climate change on transportation infrastructure, which from an economic standpoint represents the supply side of transportation, very little, if anything, has been done in order to estimate the changes in the demand for freight transportation. In addition, the lack of the regional prospective in climate change studies with regard to the transportation sector is another underdeveloped area. Therefore, this study considers the demand for freight transportation as demand by other sectors of the regional economy, which results in a clear regional perspective. If added to the findings of the supply-side studies, it can provide a better picture of what should be expected from climate change impacts on transportation in the regional context.

The ultimate goal of this study was to analyse potential climate change impacts on demand for freight ground transportation in Atlantic Canada through its links with other sectors of the regional economy. It recognizes that freight transportation is produced to meet the demand for freight transportation by other sectors of the regional

economy, which means that demand for freight transportation is a derived or factor demand. It is derived as a by-product of profit maximization or cost minimization by producers of various commodities in respective sectors of the regional economy. Hence, the main idea of the designed methodology is to model climate change impacts as productivity shocks in relevant sectors of the regional economy, and then to trace consequences of these shocks for the demand for freight ground transportation on the basis of computer simulation.

Structure of the Analysis

In order to achieve the above-stated goal, the following steps were undertaken:

- Collection of the climate related data in Atlantic region. The data reflects major climate indicators across all Atlantic Provinces. In addition, the collected data is rich enough to conduct comprehensive time-series analysis;
- Collection of aggregate economic data that reflects the economic performance of the regional economy as a whole;
- Design of a growth model that links regional economy to the freight transportation. The designed growth model is based on the state-flow relationships between economic variables over time to capture the linkage between the regional economy and freight transportation;
- Modeling of the climate change impacts as productivity shocks in the regional economy to trace their consequences for the demand for freight transportation;
- Evaluation of different scenarios of future developments in regional economy subject to the climate change impacts.

As a final result, computer simulation produced patterns of the future demand for freight transportation due to climate change impacts as well as the loss function under different climate change scenarios.

Model and Assumptions

Freight transportation is produced to satisfy demand for transportation by other sectors of the regional economy that produce various commodities and use freight transportation to deliver them to final users. Since in an economic sense the demand for freight transportation is a derived factor demand, it is obtained as a by-product of profit maximization or cost minimization by producers in relevant sectors of the regional economy. It means that there is a clear link between commodities produced by these sectors and the volume of freight transportation. In other words, there is a link between regional GDP and the volume of freight transportation.

However, in general it is not clear whether GDP enhances volume of transportation in a regional economy or transportation increases GDP because the two are interrelated and the actual relationship between them goes both ways. In statistical terms, there is no strong one-way causality between the regional GDP and the volume of freight transportation. Therefore, in such a case, it is appropriate to treat both variables as endogenous with the help of a vector autoregression (VAR). Typically, VAR is used to describe a causal relationship amongst mutually interdependent economic variables to study their evolution over time (Gordon and Boccanfuso, 2001). It explains each variable by its own past values as well as the past values of all other variables included in the VAR.

Although statistical specification of the VAR is straight forward, there should be a rigorous economic context in it. It can be done through the so-called structural VAR. Structural VAR is a system of simultaneous equations describing an economic system's dynamics on the basis of economic theory.

In this study, the economic system was defined as follows. Since freight transportation is a part of the regional economy, it is reasonable to assume that the regional GDP time path is affected by current and past realizations of the volume of freight transportation

and, at the same time, the time path of the volume of freight transportation is affected by current and past realizations of the GDP. If Y_t is regional GDP in time period t and T_t is the volume of ground freight transportation in the region in time period t, then the following system arises:

$$\begin{aligned} Y_t &= b_{10} + b_{11}T_t + \gamma_{11}Y_{t-1} + \gamma_{12}T_{t-1} + \varepsilon_{Yt} \\ T_t &= b_{20} + b_{21}Y_t + \gamma_{21}Y_{t-1} + \gamma_{22}T_{t-1} + \varepsilon_{Tt} \end{aligned} \quad (1)$$

In this system, ε_{Yt} and ε_{Tt} are pure innovations or independent productivity shocks. In classical structural VAR they are called random uncorrelated structural disturbances (Enders, 1995; Gordon and Boccanfuso, 2001). Hence, ε_{Yt} is a pure productivity shock in the regional economy as a whole while ε_{Tt} is a pure productivity shock in transportation sector.

In turn, VAR in standard form, which can be derived from system (1), is:

$$\begin{aligned} Y_t &= a_{10} + a_{11}Y_{t-1} + a_{12}T_{t-1} + e_{1t} \\ T_t &= a_{20} + a_{21}Y_{t-1} + a_{22}T_{t-1} + e_{2t} \end{aligned} \quad (2)$$

In system (2), error terms e_{1t} and e_{1t} are composites of the productivity shocks ε_{Yt} and ε_{Tt}, defined earlier, because

$$\begin{aligned} e_{1t} &= \frac{1}{1-b_{11}b_{21}}\varepsilon_{Yt} + \frac{b_{11}}{1-b_{11}b_{21}}\varepsilon_{Tt} \\ e_{2t} &= \frac{b_{21}}{1-b_{11}b_{21}}\varepsilon_{Yt} + \frac{1}{1-b_{11}b_{21}}\varepsilon_{Tt} \end{aligned} \quad (3)$$

Since the main goal of this study is to measure climate change impacts on demand for freight transportation through productivity shocks, it is perceived that these shocks have long run impacts on the

variables in the model. In addition, the following two assumptions were formulated:

Assumption 1: Transportation has no contemporaneous effect on GDP. Therefore, it is assumed that the volume of transportation T depends on income Y (as any demand does) in period t but GDP or Y does not depend on the current volume of transportation because such dependence comes from the supply side. However, the next period GDP does depend on the current volume of transportation.

Assumption 2: There are no productivity shocks in the transportation sector itself. By this assumption supply side shocks in transportation are ignored. Mathematically, it implies that $\varepsilon_{Tt} = 0$.

As a result of the above two assumptions:

$$\begin{aligned} e_{1t} &= \varepsilon_{Yt} \\ e_{2t} &= \rho\varepsilon_{Yt} \end{aligned} \tag{4}$$

in which ε_{Yt} reflects aggregate productivity shock in sectors other than transportation, and ρ is the correlation coefficient between the overall productivity in the regional economy and the volume of freight transportation demanded. The two systems of equations (2) and (4) combined show that dynamics of the demand for freight transportation is driven by the overall productivity shock ε_{Yt}.

Following the methodology introduced by Blanchard and Quah (1989) on decomposing productivity shocks into demand-side shocks and supply-side shocks, in this study the aggregate productivity shock was decomposed into (i) aggregate, climate-unrelated productivity shock, (ii) aggregate, climate-related productivity shock, and (iii) other influences.

In addition, the following assumptions were made to capture the climate-related productivity shock:

Assumption 3: Aggregate, climate-unrelated productivity shock is permanent and time invariant.

Assumption 4: Climate-related productivity shock depends on climate variables such as annual average temperature, annual level of precipitations, which is represented by rainfall and snowfall, sea level and weather related extreme events expressed by the vector C.

Assumption 5: Climate-related productivity shock is permanent and autoregressive, which reflects its cumulative dynamic nature.

Assumption 6: Other influences are totally random events. Statistically they can be specified as a white-noise process with zero mean and constant variance.

Given these assumptions, the following decomposition arises:

$$e_{it} = \varepsilon_{it} = \alpha_0 + A_1 C_t + A_2 C_{t-1} + u_t \tag{5}$$

in which α_0 is aggregate, climate-unrelated productivity shock, C is the vector of climate variables, A_1 and A_2 are vectors of coefficients, and u_t is the white-noise process (independent and identically distributed influences by assumption).

Given the dynamics of the economic system represented by (2), dynamics of the climate related productivity shock that comes from (5), and the expected changes in climate variables in the future, it is possible to evaluate the consequences of the climate change impacts on the demand for freight transportation. The climate related productivity shock can be imposed on the system (2) in the following format:

$$e^{c}_{i(t+1)} = A_1 C_{t+1} + A_2 C_t \text{ and } C_{t+1} = c + C_t \tag{6}$$

where C is the vector of climate variables, A_1 and A_2 are the vectors of coefficients defined by (5), c is the vector of annual average

changes in climate variables over the 2000-2100 period as suggested by the forecast, and t is a time period of *2000-2099*. In this specification, the vector c must include only additional changes on top of those occurred prior to 2000.

Data

A critical issue in any climate change study is the choice of a set of indicators against which the economic impacts can be measured. In this study, the set of climate variables is defined as follows:

- Average annual temperature.
- Average level of precipitations. Since the seasonal variation in precipitations can offset each other (e.g. lack of rain during the summertime versus abundance of snowfall over the wintertime), the level of precipitations is estimated based on two variables - total rainfall and total snowfall.
- Sea water level.
- Extreme weather events as represented by the number of major climate related disasters and the associated deaths.

The key source of climate data in Atlantic Canada is the Canadian Daily Climate and Environment Data Files recorded by the Environment Canada (2004). The database is recorded on CD-ROMs and is divided into two segments that cover eastern and western parts of the country. The eastern Canada CD-ROM contains daily temperatures and precipitations for 6,774 locations in Ontario, Quebec, the Atlantic Provinces, Yukon Territory, Nunavut and the Northwest Territories.

In order to obtain representative regional data, eleven meteorological stations out of 6,774 locations within the Atlantic Provinces were selected. The weather stations are located in Moncton, Fredericton, Edmundston, and Saint John in New Brunswick, Charlottetown in Prince Edward Island, Yarmouth, Halifax, Sydney in Nova Scotia, and Sainte Johns', Corner Brook and Gander in Newfoundland. The reason these stations were chosen is that the Environment Canada

climate database contains full records of the climate variables on these stations from 1950 to 2002.

The sea level is represented in terms of annual centimetres above Chart Datum for Atlantic Canada water level data adjusted for variations in atmospheric pressure. This data comes from the Coastal Impacts of Climate Change and Sea-Level Rise Synthesis Report produced by the Climate Change Action Fund Project CCAF A041 (Shaw, 2001).

The data on extreme weather events is based on a comprehensive study conducted by Environmental Adaptation Research Group, Environment Canada and Institute for Environmental Studies, University of Toronto (2002) prepared for the Round Table on Environmental Risk, Natural Hazards and the Insurance Industry. One of the goals of this study was the analysis of the social and economic impacts of hydro-meteorological hazards and disasters in Canada. Partly it is also based on the historical survey of Canadian disasters.

The data on freight ground transportation was obtained from Statistics Canada publications such as Trucking in Canada and Rail in Canada (Statistics Canada, 1980-2000). Trucking in Canada, Catalogue number 53-222-XIB, presents a comprehensive overview of the Canadian trucking industry, both for-hire and private (own account). Principal information includes statistics on revenues and expenses, equipment operated, investment, employment, and commodities transported from point of origin to point of destination.

Rail in Canada, Catalogue number 52-216-XIB, provides information pertaining to the size and structure of the Canadian rail transport industry (Statistics Canada, 1980-2000). A general overview of the rail industry for a six-year period is followed by a more in-depth analysis on the economic performance, financial structure and equipment use for the reference period. This source also provides commodity origin and destination data series supplied by the National Transportation Agency.

The data on the Gross Domestic Product (GDP) of the Atlantic Provinces was collected through Statistics Canada's Socio-economic Database (CANSIM II). CANSIM II is an online resource for the Canadian economic statistics. It provides the data to analyse long-term economic trends and to study economic activity using reliable data (Statistics Canada, 2004).

In general, CANSIM II is a multidimensional database containing 25 million time series organized in over 2,000 tables. For the purpose of this study, the data was retrieved from Table 379-0025 "Gross Domestic Product (GDP) at Basic Prices, by North American Industry Classification System (NAICS) and Province, Annual (Dollars)". Specifically, the values of provincial GDPs from 1980 to 2000 stated in 1997 constant dollars were obtained from this table. The Atlantic Canada's annual GDP was calculated by a simple summation of the provincial GDPs at the basic prices in 1997 constant dollars.

Results

In order to explain the dynamics of the economic system defined by (1), VAR in its standard form (2) was estimated. The results of the OLS estimation of the VAR with the help of the EVIEWS statistical package are shown below:

$$\begin{aligned} Y_t &= 9.23308899 + 0.6753247205Y_{t-1} + 0.1349954886T_{t-1} + 0.1380296295\varepsilon \\ T_t &= 6.91961614 + 0.0344555065Y_{t-1} + 0.5053337031T_{t-1} + 0.6094505003\varepsilon \end{aligned} \quad (7)$$

with R^2 adjusted 0.96 and 0.92 respectively. It implies that the estimated VAR nicely captures the dynamics of the economic system under study over the 1980-2000 period.

Next the problem of the climate change in Atlantic Canada was addressed. In doing so, a comprehensive time series analysis was performed on major climate variables such as mean temperature,

rainfall, snowfall, sea level, and the number of deaths per climate related accident over the period of 1950-2000 using methodology described in Yevdokimov (2001). The following table summarizes the findings and shows the break years (climate regime change) for all climate variables used in this study:

Table 1 The Break Years for the Major Climate Variables

Climate variable	Break year
Mean temperature	1980
Rainfall	1979
Snowfall	1979
Sea level	1980
Deaths per accident	1980

Based on these results, the following conclusion was made: the problem of climate change in Atlantic Canada does exist. Moreover, it can be dated back to 1979-1980. In a dynamic sense, the 1980-2000 period has been subject to a different dynamic process compared to the 1950-1979 period.

In addition, the time series analysis of the climate related variables showed that over the 1980-2000 period mean annual temperature was increasing by 0.03°C per year, total precipitations were fluctuating around 1,174 mm per year and sea level was increasing by 0.32 cm per year (Environment Canada, 2004).

According to the assumptions 3, 4, 5 and 6 introduced in methodological part, and decomposition (5), auto-regressive climate related productivity shock e_{1t}^{c} was derived as follows:

$$e_{1t}^{c} = 1.0833MT_t + 0.1556MT_{t-1} - 0.0021TP_t - 0.0007TP_{t-1} - 0.1738SL_t + 0.1327SL_{t-1} \quad (8)$$

Then the derived climate related productivity shock (8) along with the expected changes in climate variables over 2001-2100, given by the Environment Canada forecast, was imposed on the economic system (7). As a result, the following three scenarios of possible changes in

the demand for freight transportation in Atlantic Canada were evaluated.

The best-case scenario was based on the assumption that every climate variable performs at its best level given the Environment Canada forecast (Government of Canada, 2004). Therefore, in this scenario the following specific assumptions were applied:

- The mean annual temperature increases by extra 0.01°C per year on top of the existing 0.03°C or in total it increases by 4°C by the end of the century;
- The sea level rises by extra 0.18 cm per year on top of the existing 0.32 cm or by a total of 50 cm by the end of the century;
- The total precipitations remain the same;
- No extra major climate related disasters occur over the 100-year period.

Computer simulation with the help of the Mathcad-8 computer package showed that the conditions, specified in the best-case scenario, lead to the cumulative loss in demand for ground freight transportation of 7.213 billion tonne-kilometres over the 2001-2100 period.

The expected-case scenario was based on the assumption that every variable performs at the most likely level, which is actually a moderate change in the climate variable of interest according to the Environment Canada forecast (Government of Canada, 2004). In this scenario, the following specific assumptions were applied:

- The mean annual temperature increases by extra 0.02°C per year on top of the existing 0.03°C or in total it increases by 5°C by the end of the century;
- The sea level rises by extra 0.38 cm per year on top of the existing 0.32 cm or by a total of 70 cm by the end of the century;
- The total precipitations decrease by 1 mm per year or it becomes 1074 mm per year in 2100 compared to 1174 mm now;

- One extra major climate related disaster occurs every three years.

Computer simulation showed the cumulative loss in demand for ground freight transportation in Atlantic Canada of 582.127 billion tonne-kilometres over the 2001-2100 period.

The worst-case scenario was associated with the largest values of each climate variable in the Environment Canada forecast (Government of Canada, 2004). The following specific assumptions were applied:

- The mean annual temperature increases by extra 0.03°C per year on top of the existing 0.03°C or in total it increases by 6°C by the end of the century;
- The sea level rises by extra 0.68 cm per year on top of the existing 0.32 cm or by a total of 100 cm by the end of the century;
- The total precipitations decrease by 2 mm per year or it becomes 974 mm per year in 2100 compared to current 1174 mm;
- One extra major climate related disaster every two years.

The cumulative loss was estimated as 610.92 billion tonne-kilometres over the 2001-2100 period.

Conclusion

Overall results of the computer simulation showed significant decrease in the demand for ground freight transportation in Atlantic Canada due to the climate change impacts under all three scenarios. According to the derived loss function, the loss in demand mainly occurs during the first half of the 21st century with some recovery in the second half.

One possible explanation for the decreasing loss in the second half of the 21st century is that with passage of time, the economic system

better adapts to the new climate conditions. However, only under the best-case scenario the system was able to completely recover.

Since the best-case scenario is unlikely to occur, some policy instruments to reduce the climate change impacts should be put in place. These instruments may include climate change related taxes or standards in economic sectors that contribute the most to the climate change. Transportation is one of the largest contributors, and therefore, it should be the primary target for such regulations.

References

Blanchard, O.J. and Quah, D.. The Dynamic Effects of Aggregate Demand and Supply Disturbances. *American Economic Review Vol. 79 (4), pp. 655–73.* 1989.

Bosello, Francesco; Lazzarin, Marco; Roson, Roberto, and Tol, Richard. *Economy-Wide Estimates of the Implications of Climate Change: Sea Level Rise.* Climate Change Modelling and Policy Working Paper. Social Science Research Network Paper Collection. 2004.

Deke, O., Hooss, K., Kasten, C., Klepper, G., and Springer, K. *Economic Impacts of Climate Changes: Simulations with A Regionalized Climate-Economy Model.* Kiel Institute of World Economics. Kiel, 1065. 2001.

Environment Canada. 2002 CDCD East CD-ROM. Internet. Available on-line at *http://www.climate.weatheroffice.ec.gc.ca /prods_servs/cdcd_iso_e.html.* Accessed August 11, 2004.

Environment Canada, and University of Toronto. *Coping with Natural Hazards in Canada: Scientific, Government and Insurance Industry Perspectives.* A study written for the Round Table on Environmental Risk, Natural Hazards and the Insurance Industry. Ottawa, ON. 2002.

Galeotti, Marzio; Goria, Alessandra; Mobrini, Paolo, and Spantidaki, Evi. *Weather Impacts on Natural, Social and Economic Systems (WISE) Part 1: Sectoral Analysis of Climate Impacts in Italy.* Climate Change Modelling and Policy Working Paper. Social Science Research Network Paper Collection. 2004.

Galeotti, Marzio, and Kemfert, Claudia. *Interactions Between Climate and Trade Policies: A Survey.* Climate Change Modelling and Policy Working Paper. Social Science Research Network Paper Collection. 2004.

Gambarelli, Gretel, and Goria, Alessandra. *Economic Evaluation of Climate Change Impacts and Adaptation in Italy.* Climate Change Modelling and Policy Working Paper. Social Science Research Network Paper Collection. 2004.

Government of Canada. Modelling the economic impacts of addressing Climate Change. Report. Internet. Available on-line at *http://www.climatechange.gc.ca/english/publications*. Accessed October 15, 2004.

Hallegatte, Stéphane. *Accounting for Extreme Events in the Economic Assessment of Climate Change.* CCMP – Climate Change Modelling and Policy. Social Science Research Network Electronic Paper Collection. Milano: Fondazione Eni Enrico Mattei. 2005.

Jones, R.L.. Canadian Disasters - An Historical Survey. *Natural Hazards. Vol. (5), pp. 43-51.* 1999.

Shaw, R.W.. *Climate Change and Climate Variability.* Workshop Proceedings. Halifax, Nova Scotia. Environment Canada – Atlantic Region Occasional Report # 9. December 1996.

Shaw, R.W.. *Coastal Impacts of Climate Change and Sea-Level Rise on Prince Edward Island.* Climate Change Action Fund project CCAF A041. Synthesis Report. Dartmouth, Nova Scotia. 2001.

Statistics Canada. *Provincial Gross Domestic Product by Industry.* Catalogue#15-203-XIB. Ottawa. Issues 2000-2002.

Statistics Canada. *Rail in Canada.* Catalogue# 53-216-XIB. Ottawa. Issues 1980-2000.

Statistics Canada. *Trucking in Canada.* Catalogue# 53-222-XIB. Ottawa. Issues 1980-2000.

Yevdokimov, Yuri. *The Chernobyl Accident and Soviet Economy.* Sumy: University Book Publishing, 2001, 142 pages.

Challenges in Measuring the External Costs of Transport

Anthony E. Boardman, David Gillen, W.G.Waters II, Anming Zhang
Centre for Transportation Studies
Sauder School of Business, University of British Columbia

Abstract: One of the major research priorities identified by Transport Canada is to develop measures of the full costs of transport modes. This is in keeping with the growing awareness of the external costs of transportation and the necessity of recognizing these costs in pursuing the national policy interest of efficiency in transportation. The authors have carried out a review of empirical estimates of the various social and environmental costs of transportation. This paper summarizes some of the challenges involved in trying to arrive at reliable measures of the various external cost categories.

Introduction: This paper describes research undertaken in the past two years to examine and evaluate the main methods of estimating the external costs of transportation (Zhang et al., 2004). The research was commissioned by Transport Canada to provide input to their ongoing effort to estimate the full costs of various transport modes, i.e., both the costs provided and borne by transportation users, and the external costs imposed on society at large. This effort draws from a large amount of previous research that provides a basis for developing, and imposing, environmental prices for noise and air pollution on the different modes of transportation, e.g., Gillen and Levinson (1997).

There are many components of the full [social] costs of transportation. A substantial portion are recognized and borne by transportation users, such as the costs of automobiles and their operation. Some costs are subjective but still recognized and borne by users, such as the time and effort expended by automobile drivers. Some costs are imposed on society at large, such as the costs of air pollution and contributions to global warming. Some costs are borne partly by users and partly by society, such as the costs of motor vehicle crashes. In the paper we primarily use the phrase "full costs", but others often use the phrase "social costs" to mean "full" costs.

The report and this paper are confined to externality costs. Externalities are costs or benefits arising from an economic activity that affect somebody other than the people engaged in the economic

activity and are not reflected fully in prices. An externality induces a divergence between social cost and private cost or between social benefit and private benefit. The concern is that market participants only consider their personal benefits and costs in making their decisions; but if externalities are present, then market outcomes (the aggregate of individual decisions) will not be socially optimal.

Based on evidence and experience elsewhere, we develop an array of social cost and externality cost estimates, measured in 2002 Canadian $s, across a range of inter urban and intra urban transportation modes, and for both passenger and freight. Some of the estimates have large variances due to differences in location, and variances in the methodologies employed, the data sources and other characteristics of the studies upon which the estimates are based. Despite our best efforts, considerable uncertainty remains about the correct magnitudes.

Section 1 describes the concepts and terminology used in the literature concerned with external costing, it examines the issues of efficiency and market performance, and describes the importance of distinguishing between marginal and average costs. In section 2, the different categories of external cost are explained, while section 3 provides details on each of the external cost categories as they relate to the various modes of transportation. The section also summarizes issues for each major category -- challenges, insights, and some empirical figures. Section 4 contains a conclusion and identifies directions for subsequent research.

1. Concepts and Terminology: The primary motivation here is a concern about economic efficiency in transportation. This refers primarily to the need to identify the full costs of transport and to compare them to the prices paid for these services. Efficient production and pricing will produce the maximum net benefits to society from available resources.

Efficiency and Market Performance: There are a number of issues in efficiency and market performance. One is the level of competition and information in the market place, which is the traditional focus of economic regulation. Another set of issues concerns infrastructure: do user charges cover the full costs of publicly-provided infrastructure? This paper concerns a third set of issues: the problem of externalities,

i.e., the problem that some of the costs of particular services are not borne by market participants. Resource allocation will not be efficient in even otherwise perfect markets if buyers and sellers do not bear the full costs of transportation services; hence the interest in and importance of identifying and measuring the magnitude of external costs.

Some would advocate including equity concerns associated with transportation. Questions of fairness, equity or justice usually are excluded from pure efficiency considerations. This is not to imply that social dimensions are not relevant, but they require different approaches of analysis or criteria for assessment.

This paper also ignores "monetary" externalities. Construction and operation of transportation facilities can have impacts on the comfort and well-being of nearby residents. Some of these impacts, such as noise and local air pollution, would be included in an efficiency analysis-- so-called "technological" or "real" externalities. However, there may be social impacts via transfers of wealth and/or relocation of economic activities—so-called "pecuniary" or "monetary" externalities (transfers of resources without any change in the underlying productivity of the economy). Standard economic analysis excludes pecuniary externalities because it would be double counting. A final point on terminology, which is of considerable importance, concerns the concepts of average and marginal costs.

Average and Marginal Costs, Internal and External Costs: An investigation of the full costs of transportation can be done for two separate but overlapping purposes. One is a desire to compare the full costs (including external costs) of modes with the costs that are borne by users. This reveals the implicit subsidies received by users and how they differ across modes. For this purpose an *average cost* measure is adequate. The average cost estimate can be multiplied by full output to arrive at a total cost estimate.

A second purpose to identify what policy actions, such as pricing (Pigouvian taxation), would bring about more efficient use of transportation and mix of transport modes. For this a *marginal cost* is required, i.e., the incremental externality cost associated with the current level of transport output. This may differ among localities, e.g., it depends on the degree of modal congestion in a community.

Note that if the marginal costs are rising – which is often thought to characterize transportation externalities – then it would be incorrect to multiply the marginal cost by full output to estimate the full costs. That would be an overestimate. It appears that many overlook this subtlety in estimating the full costs of pollution or other externalities.

Unfortunately, the literature on externality costs in transportation is often unclear about whether marginal costs or average costs are being estimated. Also it does not necessarily sort out which components of unit costs are borne by transport users as opposed to the community. Even if an externality cost is identified as a marginal or average cost, the underlying mathematical form of the cost function is almost never specified. But, without the mathematical function, it is not possible to convert between marginal and average costs. While the relationship between average and marginal externality costs is well recognized for road congestion, the distinction has been discussed less, if at all, for other externality cost categories.

A major consideration in the measurement of the full costs of transport is whether or not users recognize and bear the costs of their actions. The efficiency concern of (negative) externalities is that there are costs resulting from transport users' decisions that are imposed on others. This sounds compelling, but there is more to it. In the case of congestion, users of congested facilities impose delay costs on one another. While these costs are borne collectively by users, there are still externality distortions because each individual user does not account for the externality costs they impose on others—they do not internalize it.

The concept of internalizing an external cost is not identical to determining who bears the cost. Who bears the average external costs might be adequate for an accounting exercise but – as in other economic decisions and analysis – the relevant concept is that of marginal costs, and average and marginal costs often differ in dealing with transport externalities.

2. Categories of External Costs: This paper examines five externality cost categories: congestion delays and their valuation; crash costs; noise costs; the costs of air pollution; and the costs of greenhouse gases (climate change). There are some additional externality categories that were deemed less important and not

reviewed.[1] These include water pollution, vibration, visual intrusion, 'barrier effects', security risks, and situations where market prices in other sectors do not reflect their underlying full marginal costs and transport activities could exacerbate these economic distortions, i.e., 'second-best' issues.

3. **The Costs of Congestion:** Some of the most significant social costs of many transport operations are those associated with congestion. Externalities are involved: users recognize that travel times and costs are higher under congestion, but individuals only consider the costs they themselves incur and do not recognize the delay costs they impose on others. The full marginal social cost of a decision to operate on a congested system is greater than the costs recognized by an individual operator. It is an intra-sectoral externality in that the full costs of delay are borne collectively by transport users and not imposed on the rest of society, but it is an externality nonetheless. The costs of congestion involves two components: 1. measuring the amount of delays; and 2. estimating the value users place on these delays.

The valuation of travel time savings: Theory and empirical evidence confirm that people place value on time savings, usually referred to as the value of travel time savings (VTTS). There is extensive literature on VTTS, primarily from road transport, but also estimates for other modes as well as freight transport.

In recent years there have been improvements in valuation methods and refinements in understanding factors that underlie VTTS, but these have not led to major changes in empirical estimates. The average VTTS is about 50 percent of wages, however values can differ substantially among users. It is preferable to disaggregate to recognize the heterogeneity of VTTS across the population. Income, distance traveled, travel conditions (congestion) and trip purpose can affect it.

One important caveat is the literature on VTTS has been dominated by developing estimates for investment appraisal of transport projects. It does not matter too much if users have different VTTS, an average is adequate to estimate full time saving benefits. In focusing

[1] A useful source exploring externality categories is Litman (updated 2004).

on divergences between private and social costs and the prospect of pricing to correct these divergences, it matters if there is a skewed distribution of VTTS , or if large and small time savings are valued the same. Regrettably, the existing literature has not looked closely at the underlying distribution of VTTS and other valuation attributes. Congestion pricing requires more information about users' VTTS than is adequate for investment appraisal.

The evidence on the valuation of freight travel time savings (VFTTS) is sparse and of wider variance than for passengers. VFTTS varies with the type of goods shipped and other characteristics such as urgency of shipment. There is growing evidence that the most important aspect of freight travel time is the increased reliability of delivery/schedule time. Studies show that the values for reliability and schedule delay exceed the values for freight time savings multiple times.

Estimating Congestion Delays: The pattern of congestion delays is traditionally shown as a relationship between travel times and traffic volumes relative to capacity of the facility.

It is well known that incremental users on a facility increase their own costs of driving and cause the costs of others to increase too. This is generally indicated by a rising marginal social costs expression. This can be calculated by the elasticity or rate of rise of the marginal private costs as traffic volumes increase.

People's choice whether or not to drive results in traffic volume that reflects the privately perceived cost. The interference with other road users means marginal social costs is greater than marginal private cost. Correcting the externality via a congestion tax would require setting it equal to the divergence between marginal private cost and marginal social costs. This results in the optimal level of congestion.

However, many believe congestion is more complex than simple conventional speed-flow relationship. In networks, congestion often arises as bottlenecks and dynamic queuing. If so, the implication is that the marginal impacts of additional vehicles is more complicated. The incremental delay depends on the location *and* timing of when vehicles enter and leave the network, *and* the overall level of congestion prevailing at any location and moment of time. In order to

estimate both the marginal and full costs of congestion, it is necessary to develop models of congestion representative of diverse network sizes and configurations. Congestion pricing is more difficult to formulate accurately than has been commonly suggested.

The Costs of Transport Crashes: The cost of transport crashes is also referred to as the cost of accidents.[2] Crashes may result in loss of lives, injuries, damage to property (including vehicles), clean-up costs and so on. A way to compute the social cost of a crash for a particular transport mode is to estimate the expected impacts or consequences of a crash (lives lost, injuries sustained, etc.) and then multiply each impact by the unit social cost of that impact. Then multiplying the cost per crash by the annual number of crashes is the total annual cost figure.

Usually crash impacts are categorized into three groups: direct costs, indirect costs and intangible costs. Direct costs include property damage such as damage to vehicles, buildings and water supplies; medical costs such as treatment, rehabilitation and counseling; and non-property accident costs such as fire, police and ambulance. Indirect costs include productivity losses at home and work, and legal and insurance costs; it also includes congestion costs at the crash site. Intangible costs are pain and suffering.

A practical difficulty is different researchers estimate different components but do not focus on one component. Aggregating different components from different sources is likely to lead to double counting. The alternative approach is to rely on one source that has been careful to ensure that the measure of accident costs is both comprehensive but does not double count components. [3]

The cost of a crash depends largely on the cost of an injury. Miller (1993) estimated the cost of injuries from non-fatal road accidents for injuries of five different levels of severity, from AIS 1 (minor) to AIS 5 (severe). He gathered monetary costs from a variety of sources and added an estimate of the lost quality of life. The lost quality of life

---

[2] The use of the word "accident" implies the absence of responsibility and blame, which of course is rarely true. It is now common to refer to crashes.

[3] For this reason, we draw heavily on Miller (1993) and Miller, Lestina and Spicer (1998).

was estimated as a proportion of the value of a statistical life (VSL) with proportion increasing with severity of the injury, Miller, 1993).

Miller et al. (1998) estimated the cost of motor vehicle crashes using a variety of methods. These costs reflect injury costs, property damage and other relevant costs. Starting with these numbers and adapting them as above for application in Canada leads to the following estimates: $25,615 (motorcycle), $338,083 (passenger car/van), $72,701 (light truck), $14,185 (medium to heavy truck), $3584 (bus).

These estimates are all based on an estimated value of a statistical life (VSL). The key issue is to determine how much someone is willing to pay for a small reduction in risk; the probability that they will live. One difficulty is that the amount someone is willing to pay (WTP) depends on the initial level of risk. Theory suggests the WTP declines as the initial level of risk declines due to declining marginal utility for safety. The amount also varies with income—people are willing to pay more in rich countries than in poor countries. Furthermore, WTP varies with mode of transport as people appear to be willing to pay more for safety reductions when they are not in control than when they are more in control.

Most recent estimates of the VSL have been obtained from wage-risk studies based on different approaches.[4] Our review reveals plausible estimates for Canada that lie between $ and $7.5 million.[5]

To put this estimate in context, Viscussi and Aldy (2003) report, from a meta-analysis that the median value of U.S. labour market studies is about $8.6 million, which they say is "in line with estimates from studies that we regard as most reliable". In our view, these studies are potentially upward biased. When this is taken into consideration, and figures are adjusted for Canadian income levels, the $4.25 million estimate is more plausible for Canadian policy purposes.

---

[4] Each method suffers from various problems.

[5] The middle of this range is $4.25 million, a figure higher than the current practice of Transport Canada, but lower than the values used by Health Canada and Environment Canada, and also lower than the value of $5 million suggested by Dionne and Lanoie (2004) for transport policy in Canada.

It seems reasonable to think of the crash costs presented above as marginal costs (or average marginal costs). The largest component is injury costs and the main component of that reflects how much someone is willing to pay to avoid a marginal injury sustained in a marginal crash.

These marginal costs can also be thought of as average costs. There should not be a major difference between average and marginal costs until congestion occurs.

Internal vs external costs is more contentious issue than above. In crashes, it is not really clear what an externality is. Fault seems a critical issue. If two drivers are in the middle of the road and slam into each other, they are both partially responsible and the consequences are largely internalized. What about a tourist who is in a park enjoying the view and is killed by a drunk driver who has lost control? That tourist's death seems like an externality. If your actions lead to a crash in which your passengers die, is that an externality? Unlike the tourist enjoying the view, they had the choice of driving with you or not driving with you and had some knowledge of the risks. They are "other users"- part of the user group. However, given that they were not in control, one could the deaths as an external cost.

Insurance costs partially internalize property damage and liability suits are usually shared by users as a whole but this still leaves situations where risk-taking drivers impose risks on others. Insurance premiums vary but this does not fully adjust for crash risks that some impose on others. As automobile purchase prices increase due to mandated safety features, this internalizes part of safety costs although, once again, this is an average cost concept and is not necessarily a guide to the marginal costs of safety associated with different drivers and vehicles and it encourages moral hazard.

The Costs of Noise: Noise constitutes unwanted sound and emanates from all modes of transportation. In order to place a price on noise (or the lack of quiet) two magnitudes must be established, the amount of noise and value which people place on it. The amount of noise is measured in decibels (dB) and can reflect the difference between ambient noise levels at a point in time or over a period of time. The value for noise must be established through secondary markets where noise is bundled with other market products such as housing.

Alternatively, values can be established through stated preference experiments using tools such as contingent valuation or conjoint analysis.[6]

Like other externalities considered in full costing there are two fundamental questions that enter the valuation and pricing decision. The first, is the noise exposure based on average or marginal values and second, is the externality internal or external to the system?[7] Noise damage from road, rail and transit use depends on where and when the noise is generated. Therefore, the full costs of noise should be included in the calculations of the external costs of transport because the noise externality is generated by the components of the transportation system but paid for (through a loss of consumer surplus) by agents outside the system: those beside the airport, the roadway and the rail tracks. This may be less true for public transit.

Is it the average or marginal value that is important for the basis of establishing price? Empirically, the vast majority of the literature measuring noise value is based on cumulative noise metrics. For noise from aviation, there would be no distinction between average and marginal since the noise metric is a cumulative noise measure where the additional or marginal flight is included in the metric. For other modes, rail, road (car, truck and bus) and water, the same argument would hold. The noise contour measured along a road or rail corridor would be based on NEFs where the marginal vehicle has been included in the measure.[8]

From the literature review undertaken we found for aviation that a noise discount for housing in the U.S. was 0.5 to 0.6 percent per dB. Hence a property located in 55 dB would sell for 10-12 percent less if it were located in a 75 dB zone. However, in Canada, the discount

---

[6] Noise measures for a point in time are described as single event measures while those which are measured over time such as NEF (Noise Exposure Forecasts) reflect the amount of noise locations are exposed to over a day with nighttime noise carrying a higher weight than daytime. NEFs are used to create noise contours that define noise exposure of a given level over space.

[7] Another underlying issue concerns the property rights to quiet (lack of noise).

[8] The value of NEF changes in a differing but non-linear way between more noise per event and the number of events.

seems to be higher, from 0.8 to 0.9 per cent per dB. For road traffic noise, for auto, truck and urban bus, noise discounts have been reported that range from 0.08% to 2.22%. A simple mean for these studies is a value of around 0.55, although one would want to use a range of 0.4 to 0.65 in any calculations. Two studies for rail, both stated preference found a noise depreciation of 10 percent but the value dissipated quickly for perpendicular distance from the rail track.

Air Pollution Costs: Transportation produces air pollutants directly from vehicle operations and related infrastructure during construction, and indirectly from vehicle manufacturing, fuel extraction and processing, and manufacturing of construction materials and machinery. Our study focuses primarily on direct emissions. In addition, impacts of air pollutants can be either local or global. We focus on local air pollution whose impacts take place in the area where the emission occurs.

In general, the costs of local air pollution include human mortality and morbidity, reduced visibility, corrosion of materials (buildings, rubber products, etc.), reduced agricultural production and damages to ecological systems. Two broad approaches to the valuation of (non-market) public goods such as environmental quality are used. The "revealed preference" (RP) method infers consumers' valuations from their actual choices in markets affected by them. Hedonic pricing methods, for example, assume that an individual buys goods for their various attributes. The second major approach in valuation is the "stated preference" (SP) approach. The contingent valuation method works by directly soliciting from a sample of consumers their willingness to pay (WTP) or willingness to accept (WTA) a change in the level of environmental service flows, in a carefully structured hypothetical market (Davis 1963). Some unresolved issues are discussed briefly below.

Ecological Systems: Our calculation has focused on the health costs (mortality and morbidity) and material damages of air pollution. The impacts on ecosystem services include natural purification of water, erosion control and habitat for wildlife, which are public goods that have value to society but no relevant markets where their values are expressed (Xu et al. 2003).

Our estimates, as marginal costs, are applicable to small reductions in those levels, the question remains whether they are also applicable to large reduction, or complete elimination of the damage. It is not clear how to use these estimates to compile full national costs of transportation. These two questions hinge on the shape of the damage cost functions. In the literature this is almost never specified. In some cases there may be no difference between average and marginal costs.

There are several reasons to suspect that marginal externality costs increase with the volume of output. First, most studies assume that dose-response functions are linear, but we question this. Second, there is some evidence that for low levels of emissions, damage from air pollutants has lower or even zero values. The Studies by Health Canada (1994) suggest that there are no proven thresholds for the health effects of ozone and particulates, although there is more likely one for CO, and also apparently thresholds for effects of the pollutants on vegetation.

After estimating the social costs of air pollution, it is also important to identify if any of the current "full costs of air pollution" are actually borne by transport users as distinct from borne by society as a whole. All vehicle users will bear some of the costs of air pollutant emissions, which include health bills and some prevention costs (the latter could be substantial, such as the costs of emission controls on automobiles. Thus part of the effects users impose on one another but more important are the effects imposed on non-users. The relative amounts of full damage to users and non-users will depend on their numbers, the concentrations to which they are exposed, and lengths of time they are exposed. We suggest that the internalized part of the social cost might be small.

Greenhouse Gases (Climate Change): Global climate change, often referred to as "global warming," involves an increase in the average atmospheric temperature of the Earth. Predictions of the effects of climate change are tenuous. We do not really know what the consequences are. The extent of global warming may be indicated by the global average surface air temperature, which has risen since Industrial Revolution with an accelerated rate of increase in the last three decades.

For the measurement of GHG costs we review and make use of estimates from elsewhere and discuss their applicability to Canada. In particular, Tol and Downing (2000) provide a set of marginal costs of GHG emissions, based on version 1.6 of the Climate Framework for Uncertainty, Negotiation and Distribution (FUND). The model and scenarios are so far extrapolated that the results are extremely uncertain.

We consider the unit costs estimated to be a marginal cost (as opposed to average cost) concept but this could be debated. As with the case of air pollution, the unit costs are thought to be increasing, at an unknown rate, in the volume of transportation output, i.e., full externality costs are rising at an increasing rate. While the relationship between average and marginal externality costs are well recognized for road congestion, this subtlety in estimating full costs of emissions is little discussed in the GHG emissions literature. We consider this as an important, but difficult, research area in the future.

GHG is an externality cost category, any climate changing effects of GHG emissions are borne by the world at large rather individual transport users. However, insofar as policies are adopted to combat GHG emissions, the costs imposed on transport users would be internalising the external costs to some extent.

4. Conclusions/Directions for Research: Exploring empirical estimates of the external costs of transport is fascinating but frustrating. Despite years of research, the margins of error can be immense, and there are fundamental conceptual issues as well. One overriding issue that emerged is whether we are measuring average or marginal costs. Many studies for various external cost categories are ambiguous. For congestion costs, there is a rigorous mathematical link between average private and marginal social costs, indicated by the speed-flow-density equation. This linkage between average and marginal costs is not well recognized for other external costs. But even for congestion costs, doubts are arising about the predicated stable function between congestion and delay costs.

Congestion may be more complex than as it is portrayed in traditional road pricing illustrations. Incremental impacts of delays imposed on the system may not be a well-defined function of traffic volumes, but also depend on specific location, time periods and

background conditions. Also, the extensive literature on the valuation of time savings or delays is adequate for investment appraisal, but pricing to correct for external costs will require more rigorous understanding of variations in individual VTTS. These latter two themes are important avenues for research.

The valuation of crash costs is becoming well established. There are still controversies over valuations such as VSL, which have increased over the years, but also over the extent to which crash costs are recognized and internalized and borne by transport users. Identifying and correcting for external crash costs in transportation markets requires that these latter controversies be resolved.

The valuation of noise costs has been studied extensively and empirical estimates seem to have a lower variance than estimates for other external cost categories. There are even the beginnings of pricing noise (for aircraft operations in Europe). But there seems to be some ambiguity whether average or marginal costs are being measured.

Local air pollution costs have been a recognized externality category for some time and have led to some serious interventions with attendant internalization of costs (e.g., the expense of emission controls on cars, particularly in North America). There are still substantial debates over the health impacts of various types of emissions, whether the health impacts may be non-linear, and hence whether monetary estimates are average or marginal cost concepts.

The greatest uncertainty about the valuation of external costs are those associated with GHGs and possible climate change. Although the consensus is that transport emissions are causing some cumulative impacts on the atmosphere and weather, there remains substantial scientific debate regarding the importance of human-induced impacts relative to natural causes. Accordingly, trying to estimate the global economic impact of incremental emissions is even more difficult than for the previous external cost categories.

Further over-arching themes for future research arising from our review is the importance of being able to recognize and incorporate technological advances and substitutions. Even well-done empirical studies become dated very quickly. Technology and operating practices have reduced noise impacts and emission rates substantially.

Estimates of the full costs of various transport modes will also require projections of technological trends in order to assess probable future levels and severity of external costs. As policies are adopted to recognize and internalize external costs, there is the need to be able to predict substitutions in transport use both on the supply and demand side.

As with most research, a rich agenda for ensuing research is an important part of our findings. However, good science depends on good information. If there is a single failure it is information to provide a set of policies based on Canadian data. Importing estimates from elsewhere makes strong untested assumptions as to the transferability of parameters and impacts.

References

Davis, R. (1963), "Recreation planning as an economic problem", *Natural Resources Journal*, 3(2): 239-249.

Gillen, D. and Levinson (1996), "The Social Costs of Intercity transportation: A review and Comparison of Air and Highway" *Transport Reviews*, 18(3): 215-240.

Litman, T. (2004): *Transportation Cost Analysis Estimates and Implications*, Victoria Transport Policy Institute, www.vtpi.org

Hansen, J., and S. Lebedeff (1987), "Global Trends of Measured Surface Air Temperature." *Journal of Geophysical Research*, 92, 13345-72.

Hansen, J., and S. Lebedeff (1988), "Global Surface Air Temperature: Update through 1987." *Geophysical Research Letters*, 15, 323-6.

Health Canada (1994), "National Ambient Air Quality Objectives for Carbon Monoxide," Executive Summary, Prepared by the CEPA/FPAC Working Group on Air Quality Objectives and Guidelines, Health Canada, Ottawa, November. (1993): "Costs and Functional Consequences of US Roadway Crashes." *Accident Analysis and Prevention.* 25(5): 593-607.

Tol, R.S.J., and T.E. Downing (2000). "The Marginal Costs of Climate Changing Emissions." Institute for Environmental Studies, Vrije Universiteit, Amsterdam.

Xu, Z., G. Cheng, Z. Zhang, Z. Su and J. Loomis (2003), "Applying contingent valuation in China to measure the full economic value of restoring ecosystem services in Ejina region", *Ecological Economics*, 44(2-3): 345-358.

Zhang, A, A.E. Boardman, D. Gillen, and W.G. Waters (2004), Towards Estimating the Social and Environmental Costs of Transportation in Canada, A report for Transport Canada.

"Measuring Congestion Costs"

W.G.Waters II
And
Markus von Wartburg
Centre for Transportation Studies
Sauder School of Business
The University of British Columbia

ASTRACT

The costs of delay and frustration due to congestion are widely cited as a major social or environmental cost of transportation. There have been numerous estimates of the total costs of congestion borne by local and/or national economies. Many empirical estimates of the costs of congestion compare observed congestion levels against hypothetical free-flow conditions. But free-flow conditions are not necessarily the relevant basis for comparison. Various alternate bases for measuring congestion have been suggested. The economic approach is to compare actual versus optimal levels of congestion, although the latter is more difficult to define and measure than free-flow performance. This paper explains the economic approach, the difficulties of trying to implement this approach, and various other problems that interfere with measuring the costs of congestion. Empirical estimates of congestion costs relative to GDP are cited for a number of countries, mostly European.

Introduction

One of the major research priorities identified by Transport Canada is to develop measures of the full social costs of transport modes. This is in keeping with the growing awareness of the external costs of transportation and the necessity of recognizing these costs in pursuing the national policy interest in efficiency in transportation.

The costs of delay and frustration due to congestion are widely cited as one of the most important environmental costs of transportation.

There are many citations to the economic costs of congestion, including empirical estimates. For example, an annual congestion cost of $1 billion is often quoted for the Toronto region, with projections that may rise to $3 billion over the next decade or two. Annual figures of $500 million or more are cited for the Vancouver region, and there are estimates for other North American cities. There are many factors affecting the level of congestion and their costs, so it is not surprising that estimates of total congestion costs will vary from place to place. This paper first reviews the underlying concept for measuring congestion costs, considers alternative definitions and their implications, and reviews briefly some of the empirical estimates of total congestion costs.

Measuring and Interpreting the Costs of Congestion

The externality of congestion differs in some respects from most other examples of externalities. By definition, externalities refer to costs (or benefits) that are not priced in the market and accrue to third parties as a result of actions taken by individuals. Buyers and sellers in a market place will recognize the costs (and benefits) that they incur personally. But in the absence of property rights or legal frameworks that force individuals to consider such externalities, parties in a market transaction will not consider the costs imposed "externally" by their action. Pollution is a classic example of an externality since neither buyer nor seller of a polluting product has any incentive to consider the pollution costs they impose on others. Insofar as pollution costs are not taken into account in transport users' decisions, it is necessary to add the external costs of pollution to those costs that are recognized by users.

In their decision making process, motorists recognize the higher time costs they face during road congestion, but they do not recognize that they impose delay costs on other motorists on the road. Motorists do not perceive themselves a cause but rather a victim of congestion. Nonetheless, the collective delay costs are borne internally by all the motorists as a group; they are not imposed on a third party. In a sense, the externality is internalized in the aggregate,

a form of collective internalization unlike most other externalities. The question arises whether congestion costs are part of the external social costs to be added to the private costs to arrive at the total social costs of transport, or whether such congestion costs are already internalized by motorists as a group? Both statements are correct. Motorists under congestion do bear the full time costs collectively; nonetheless, there are externalities involved that result in excessive total transportation costs.

There are at least two broad categories of users of congested facilities: those engaged in commercial activity (e.g., trucking) and those traveling for personal reasons including commuting. Some studies concentrate just on the commercial sector, e.g., Lynch, 2004, where reductions in time delays are assumed to translate into increased economic output. Other studies incorporate the value of travel time savings (VTTS) of reduced delays to all travelers, regardless of whether time savings are valued for personal reasons or for economic output. But whether both groups are included or not, there are some underlying issues in identifying what we mean by congestion and its costs, and how we should measure these costs.

Measuring Congestion Costs

There are exists some confusion on the appropriate measure of the social costs of congestion as various alternative measures have been suggested in the literature (Newbery, 1990). If one is measuring the total social costs of transportation, the imputed value of total time spent traveling would certainly be part of the total costs. However, travel time costs exist even in the absence of congestion, but the existence of congestion causes these costs to be higher as users interfere with each other's travel. It is therefore appropriate to separate the costs of congestion from unavoidable time costs of transport, yet there is ambiguity in the proper way to do it.

The classic speed-flow relationship generates the cost curve in Figure 1. Cost are defined as both direct monetary outlays and the motorists' valuation of time and (in)convenience. (Initially, we assume a uniform value of travel time to simply the figure; the implications of this assumption are reviewed later).

Ignoring the possibly backward-bending part of average costs, Figure 1 shows the average costs of travel for different traffic volumes as well as the demand for travel. (The marginal private costs perceived by motorists are the same as the average cost to the group as a whole). In the absence of congestion pricing, all motorists whose willingness to pay (in time and money) exceed their costs will use the road and the market equilibrium occurs at the intersection of the average cost curve and demand, at point *B*. The traffic volume is equal to the distance *OE* with average costs of *OA* and total travel costs equal to the rectangle *OABE*.

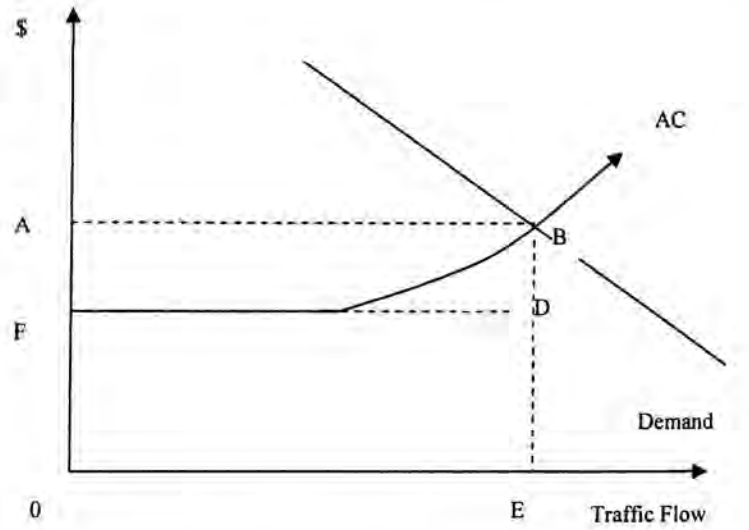


Figure 1: Illustration of Congestion Costs

One possible measure of the costs of congestion frequently cited is the extra cost involved in traveling on congested rather than uncongested roads, i.e. comparing the actual travel conditions with

hypothetical free-flow conditions. In such a hypothetical scenario of zero congestion, the free-flow level of average costs *OF* would prevail for all traffic. Total travel costs would be *OFDE* and the difference from the original market equilibrium *FABD* would be the costs identified with congested versus non-congested conditions.

Defining congestion costs based on a comparison with hypothetical free-flow conditions is unrealistic and misleading. It would clearly be uneconomical to build road capacity up to a point at which congestion is completely absent (Newbery, 1990). Nonetheless, two arguments can be made in favor of defining the costs of congestion in relation to free-flow conditions. Koopmans and Kroes (2004) argue that diagnosis and therapy should not be mixed up and measuring the size of a problem is not identical to implying that this size could, or should, be zero. The problem with this argument is that if congestion costs are not measured appropriately, the indicator is meaningless in directing strategic decisions about the congestion problem. The other argument in favor of the comparison with free-flow conditions is the fact that congestion costs can be calculated with knowledge of the values of time of motorists and the congestion technology without having to estimate travel benefits. Nonetheless, the above mentioned arguments are questionable. Congestion costs estimated based on comparisons to free-flow conditions are misleading and exaggerate the social costs of congestion.

An alternative definition of congestion is to replace free-flow with some "reasonable" degree of congestion that motorists could expect to face, say level of service (LOS) D of the AASHTO Highway Capacity Manual, or some arbitrarily chosen ratio of volume to capacity ratio.[1] In Figure 1, instead of the free-flow cost OF one would adopt some point on the average cost curve between OF and OA as the base for calculating the excessive cost of congestion. The problem, of course, is to agree on the appropriate base from which to measure an excess or social cost of congestion.

---

[1] DelCan et al (2004) reviews alternate engineering measures of congestion, p.23ff.

The economic approach to measuring congestion is to compare the extra costs of congestion in market equilibrium with the efficient situation in which motorists were forced to recognize the marginal social costs of travel. Figure 2 repeats Figure 1 but includes the marginal social costs (MSC) of congestion, i.e., the collective delay costs caused by each incremental addition to the flow of vehicles. As is well known, the theory of congestion pricing calls for motorists to pay a toll equal to the difference between the costs they personally perceive (the average cost curve) and the MSC.

In Figure 2, under the optimal congestion toll, traffic volume is reduced to *OH* and motorists pay a toll equal to *CG* that raises their individual costs of travel up to the marginal social costs. Total social costs of travel have fallen from market equilibrium *OABE* to *OJCH* as there are fewer motorists on the road and the costs to those who remain traveling are lower. The difference *ABCJ* + *HCBE* is the portion of total costs attributable to the congestion externality. This difference can alternatively be calculated as the area beneath the marginal social cost curve, area *HGKE*. Although it is still the case that motorists as a group bear the full social costs of congestion, it is conceptually useful to separate the costs into two components: the total costs under optimal congestion *OJCH* and the excess congestion costs *HGKE*. Note that if the engineering approach of specifying an acceptable base level of congestion would approximate traffic level OH, then the engineering approach would arrive at a measure of congestion costs approximately the same as the above measure of total congestion costs in excess of the economic optimum.

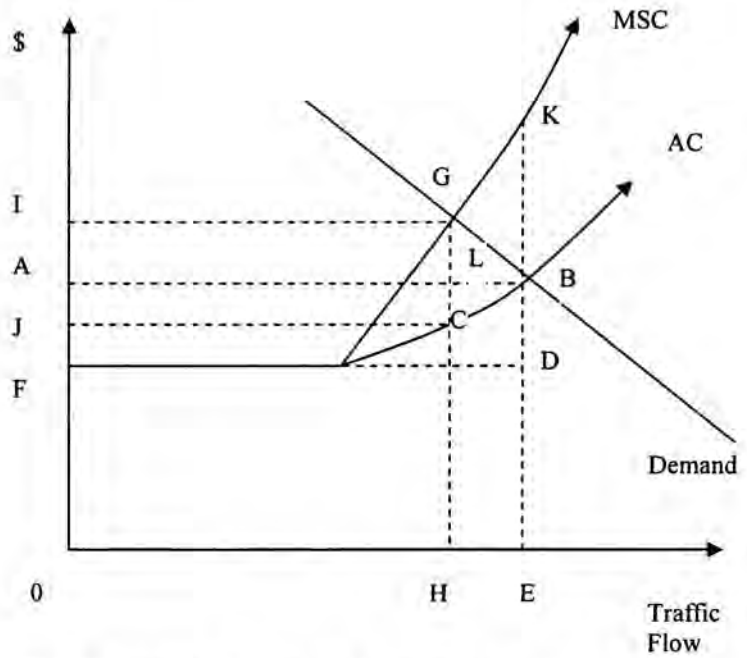


Figure 2: Illustration of Congestion Costs and Pricing

The Economic Welfare Cost Approach

Still another approach to measuring the social costs of congestion is to limit it to the loss in social surplus associated with excessive road use. This is the concept of "deadweight loss" in welfare economics. In the previous paragraph, the incremental congestion costs are represented by the area beneath the MSC curve for the additional traffic volume *HE*, (i.e., HGKE). However, this incremental social cost is accompanied by a benefit of allowing more motorists to travel on the roads. The benefit to the incremental motorists HE is the area beneath demand HGBE. The loss in social surplus due to excessive road use is then equal to the area of the

standard deadweight loss triangle *GBK*. The social costs of congestion are then much less than the gross measure of costs associated with the additional traffic (*HGKE*).

It is the imposition of a congestion tax of amount *CG* that forces motorists to confront the marginal social costs of the decision to drive. This generates toll revenues which are a transfer from motorists to society via the government or tolling authority. There is a net social gain of *GBK*, but motorists will perceive themselves worse off because of tolling. In fact, assuming a constant value of time all types of road users - the tolled and the tolled off (those who avoid the road to shun the toll - are worse off.[2] With the exception of hyper-congested road conditions,[3] the major benefit of congestion pricing lies in the revenue collected.

"... although pricing certainly reduces congestion, the larger part of actual benefit from road pricing does not consist of this congestion relief. The benefit sits, 'locked-up', in the revenue collected, and it is only released when the revenue is used." (Goodwin, 1997).

Public support for road pricing measures will be low unless toll revenues are earmarked and motorists perceive that the money is channeled back in the form of reduced taxes, lower user charges or other transportation improvements (Hau, 1992).

A complication in computing the portion of total costs associated with excessive congestion is that the position and shape of the cost curves depend on the underlying road capacity. Expanding capacity shifts the curves down- and outwards. Conversely, if capacity has not kept pace with the growth of demand over time, congestion and delay costs will be higher. Hence, the level of congestion costs is not only a function of motorists' behaviour, but it is influenced by the actions (or inaction) by road authorities both in pricing and investment policies. The ideal approach to measuring congestion costs requires examining both optimal pricing and optimal investment decisions. Kraus, et al. (1976) first demonstrated this

---

[2] With heterogeneous valuation of time, motorists with very high values of time might be better off despite tolls.

[3] In the presence of hyper-congestion, everyone is better off with congestion pricing.

approach for California freeways; they estimated that the deadweight welfare losses might be on the order of one percent of California GNP at the time.

Some Alternate Approaches to Measuring Congestion

In the traditional speed-flow relationship, measuring the marginal social costs on the system are relatively straightforward. There is a mathematical relationship between vehicle density and flow (hence travel time) which varies with road capacity. But if congestion is associated with bottlenecks and dynamic queuing, then the congestion impact of an additional vehicle depends on both the location and time of entry of the vehicle into the transportation system. Measuring marginal social costs is much more difficult in these circumstances. How congestion is modeled can make a significant difference in measuring the congestion impact of an incremental vehicle entering the network. This makes it all the more difficult to estimate the total costs of congestion.

Because travel demand relative to capacity differs from place to place and hour to hour, there is no single set of congestion cost curves. In order to estimate total social costs of congestion for society, one needs to know degrees of congestion for all locations and times of the day. One might propose a highly aggregative or overall average relationship and an average underlying VTTS. This would enable a very rough estimate of the average level of congestion costs. This has given rise to a literature on developing indices of congestion that can be aggregated and compared from place to place. (See DelCan et al. (2004) for a review and initial estimates of congestion indices and costs for Canadian cities.)

Example Estimates of Total Congestion Costs for Various Countries

Because previous studies have used various definitions for total congestion costs, it is essential to distinguish the different definitions in order to prevent comparing apples and oranges. However, some studies are unclear on precisely how they arrive at their estimates. Therefore the results should be considered and compared with the appropriate care.

Dodgson and Lane (1997) define congestion costs as the extra costs road users incur because the existing British road network is congested. Hence, they estimate total congestion costs based on the comparison to free-flow conditions. They estimate the total costs of congestion (including time and fuel) in Great Britain in 1996 based on the speed-flow model to be C$14.6 billion, (2002 C$) which is about 1% of GDP in 1996.

On the other side of the spectrum of congestion cost definitions, Infras/IWW (2000) define the total congestion costs as the total welfare loss caused by non-optimal pricing (as discussed above: area GKB in Figure 2). This large study analyzes congestion in 17 European countries and estimated total congestion costs (including time but not fuel costs) based on the speed-flow model at around 1.9% of GDP in 1995. Although, the definition of congestion costs is less inclusive than in Dodgson and Lane (1997), the study finds significantly higher cost-GDP ratio estimates. However, the estimates exhibit large variation across countries that might be linked to the presence of large urbanized areas. The Infras/IWW (2000) estimates are presented in Table 1. There is substantial variation in the size of congestion costs relative to GDP, from 0.47% (Ireland) to 2.85% (Netherlands).

Table 1 Infras/IWW Estimates of Congestion Costs by Country

Study	Country	Year	Costs (billion)	Cost/GDP
Infras/IWW (2000)	Austria	1995	3.27	1.29%
	Belgium	1995	5.76	1.91%
	Denmark	1995	1.68	0.95%
	Finland	1995	1.09	0.83%
	France	1995	29.92	1.80%
	Germany	1995	44.53	1.83%
	Greece	1995	3.27	1.60%
	Ireland	1995	0.38	0.47%
	Italy	1995	29.08	2.01%
	Luxembourg	1995	0.35	1.95%
	Netherlands	1995	12.90	2.85%
	Norway	1995	1.17	0.78%
	Portugal	1995	1.61	0.95%
	Spain	1995	15.55	1.90%
	Sweden	1995	1.61	0.69%
	Switzerland	1995	5.08	1.61%
	UK	1995	30.89	2.75%

Source: Koopmas and Kroes (2004).
All figures are in 2002 C$.

There are a few additional national studies which include different components of congestion costs but we were unable to clearly identify the definition for total congestion costs. Hence these estimates in Table 2 should be used with caution. Note that some estimates include crash costs and emissions in the estimates of congestion costs. But also note the order of magnitude difference in the two estimates of congestion costs relative to GDP for the two countries with a comprehensive definition of congestion costs.

Table 2: Various Total Congestion Cost Estimates

Study	Country	Year	Costs (bill.)	Cost/ GDP	Cost Types[A]	M.[B]
BTE (1997)	Australia	1995	13.13[C]	2.32%	TF	?
Schierhackl (1995)	Austria	1995	9.03	3.53%	TFAE	SF
Budget Ministry (2003)	Belgium	1995	1.57	0.51%	TF	?
AVV (1998)	Netherlands	1997	1.13	0.23%	TF	D
Infras (1998)	Switzerland	1996	0.64[D]	0.3%[D]	TFAE	D
Delucchi (1998)	USA	1991	119.27[D]	1.35%[D]	T	SF
TTI (2000)	USA	2000	83.54[E]	0.70%	TF	?

Source: Koopmas and Kroes (2004); cited (along with references) in A.Zhang, et al. (2004) Chapter 2.
All figures are in 2002 C$.
[A] Cost types included in the estimate: T - time, F - fuel, A - accidents, E - emissions.
[B] Measurement method for time cost: D - direct measurement, SF - speed-flow model.
[C] Six large urban areas.
[D] Figure represents mid-point of the range of estimates.
[E] 75 large urban areas.

Other Issues and Complications

What conclusions can we draw? There are substantial variations in estimates of congestion costs. These reflect not only empirical vagries but differences in definition and interpretation of the costs of congestion. Nonetheless, the diverse studies confirm that congestion costs can be substantial.

An issue ignored by the above discussion is the relationship between congestion and the variability of travel times. Road users value reliability (predictability of travel times as opposed to the duration of the travel time), and hence the appropriate road charges should take into account the interaction between congestion and the variability of journey times. It is generally accepted that predictability/reliability declines as congestion rises, but the linkage is not well understood. Such issues probably have to be resolved by relying on traffic simulation models. This is an important research

topic for developing accurate estimates of the marginal social costs of vehicles entering a congested facility.

Another issue not addressed is the implications of vehicles with different performance characteristics and how they affect vehicular flow and congestion.

There are other complications to the above illustration of congestion costs. As noted earlier, the traditional diagrammatic presentation of speed-flow relationships of congestion implicitly assumes a constant value of time. If we allow for heterogenous values of time, the analysis has to be modified but the results are broadly consistent.[4] Instead of the optimal congestion toll being based on the representative driver's VTTS, it is now based on a weighted average of different motorists' valuation of time, weighted by number of trips taken by those motorists using the facility (Hau, 1992). The constant value of time assumed in Figure __ can therefore be re-interpreted as the weighted average value of time. In fact, some motorists now can be better off after the toll is imposed if their value of time is sufficiently higher than average. A companion issue is the skewness of the distribution of VTTS. Many believe that VTTS is "skewed left," i.e., there are a relatively few with very high VTTS balanced by a much larger number with lower VTTS. Even a modest congestion toll might divert a sizeable number of low-VTTS users and result in highly-valued time savings by those with a high VTTS. This suggests that greater knowledge of the distribution of VTTS values and associated willingness to pay for congestion relief should be an important research priority. There has been substantial opposition to congestion pricing because the majority of road users perceive themselves as worse off. But with heterogeneous VTTS there is the possibility of substantial numbers of motorists made better off via tolls.

All these extensions to an analysis of congestion costs can be important. They indicate that there remains a sizeable research agenda regarding the measurement of congestion costs and the impacts of policies to improve it. Given the substantial literature that

---

[4] For a mathematical extension to heterogeneous values of time consult the appendix of chapter 4 in Mohring (1976).

already exists on congestion costs and pricing, it is a bit puzzling – and even daunting -- how many unresolved issues remain.

REFERENCES

DelCan, et al. (2004) "Costs of Congestion in Canada's Transportation Sector," Draft final report, for Transport Canada Sustainable Development Branch, Contract T8080-01-1593.

Dodgson, J. and B. Lane (1997) *The Costs of Road Congestion in Great Britain*, National Economic Research Associates, NERA Briefing Paper, London.

Goodwin, Phil (1997): "Solving Congestion", Inaugural Lecture for the Professorship of Transport Policy University College London, 23rd October, 1997.

Hau, Timothy D. (1992): "Economic Fundamentals of Road Pricing: A Diagrammatic Analysis", *Policy Research Working Papers*, WPS 1070, Transport Division, TheWorld Bank.

Infras/IWW (2000) *External Costs of Transport: Accident, Environmental and Congestion Costs in Western Europe*, Zurich: Karlsruhe.

Koopmans, Carl and Eric Kroes (2004) "Estimation of Congestion Costs in the Netherlands," *SEO Amsterdam Economics Discussion Paper*, No. 28.

Kraus, M., H. Mohring, T.Pinfold (1976) "The Welfare Costs of Nonoptimum Pricing and Investment Policies for Freeway Transportation," *Amer. Econ. Review* (September) 532-47.

Lynch, M. (2004) "Estimating the Delay to Motor Carriers Operating in the Lower Mainland and Fraser Valley," *Proceedings*, Vol. 1, Canadian Transportation Research Forum Annual Conference, Saskatoon: University of Saskatchewan Printing Services, 112-126.

Meyer, M.D. (1994) "alternative Methods for Measruing Congestion Levels," in *Curbing Gridlock: Peak Peri9od Fees to Relieve Traffic Congestion* Vol. 2, Special Report 242, National Research Council, Washington D.C.: National Academy Press, 32-61.

Mohring, Herbert D. (1976): *Transportation Economics*, Ballinger Press; Cambridge, MA.

Newbery, David M. (1990): "Pricing and Congestion: Economic Principles Relevant to Pricing Roads," *Oxford Review of Economic Policy*, 6, pp. 22-38.

Zhang, A., A.E.Boardman, D.Gillen, W.G.Waters II, et al. (2004) "Towards Estimating the Social and Environmental Costs of Transportation in Canada," Report for Transport Canada, Vancouver, UBC Centre for Transportation Studies.

Estimating Traffic Delay Caused by Freeway Crashes

J. H. Noh, F. Saccomanno, and B. Hellinga
Department of Civil Engineering
University of Waterloo
jhnoh@uwaterloo.ca

Abstract

Freeway crashes result in significant traffic delay costs. In the US, these costs have been estimated from aggregate nation-wide data at approximately $20-25 billion dollars annually. The approximate nature of these estimates fails to provide meaningful insight into specific attributes affecting delays that are needed to justify policies and countermeasures. Currently, there is no reliable method to estimate crash-induced delay.

This research seeks to develop a model for predicting traffic delay caused by different types of crashes for different traffic situations and freeway types. The focus of the discussion in this paper is on the estimating traffic delay from freeway crashes using empirical data. The discussion highlights a number of important issues, such as, missing data manipulation, data smoothing, and temporal and spatial extent of individual crash delay. Delay estimates are obtained for a sample of crashes on an instrumented urban freeway in Toronto, Ontario.

Introduction

The performance of freeways is frequently affected by recurrent and non-recurrent traffic congestion. The focus of this paper is on non-recurrent congestion, which occurs when capacity is reduced by any random traffic incident. Lindley (1986) reported that more than 70% of all non-recurrent delay on freeways was incident-related. For urban highways, Schrank and Lomax (2002) reported a value in the range of 52% - 58% of total delay.

According to Lance et al., (1992) crashes represent about 10% of all freeway incidents, but they have a disproportionate effect on traffic delay. This is due to the fact that crashes occur with unexpected rapidity, and frequently involve other vehicles such as, police cars, tow trucks, ambulances or fire trucks. This induces motorists to drive more slowly when passing a crash site. Hence, crashes are an important sub-class of total freeway incidents that cause delay.

In order to establish cost-effective traffic management strategies for alleviating the delay caused by crashes, planners need to better understand the relationship between crash characteristics and associated traffic delay at a given site and set of traffic conditions. An accurate estimate of delay caused by crashes is a necessary first step in a traffic delay prediction model. This paper considers an empirical approach for obtaining accurate estimates of delay caused by crashes on freeways.

Although the delay costs due to crashes on freeways have considerably increased over the past decade, only a few studies have been undertaken to provide accurate and reliable estimates of delay. Much of the existing evidence tends to be aggregate and approximate in nature, much of it based on insurance claims. Blincoe et al. (2002) estimates that traffic delay caused by crashes in the US represented approximately $25 billion/year. These costs involve emergency response, clean-up, medical treatment, property damage, and various administrative and legal costs associated with crashes. According to Vodden et al. (1994), the total cost of motor vehicle crashes in

Ontario in 1990 was approximately $9.1 billion.

These estimates do not normally consider traffic delay caused by crashes, and it is our contention that this omission significantly underestimates the true cost of crashes.

In this paper, we attempt to obtain accurate estimates of crash-induced delays that can be used in the development of reliable delay prediction models. To do this, we introduce an empirical approach for estimating crash-induced delay using instrumented freeway loop-detector data. The paper highlights different methods for estimating traffic delay and underscores their limitations. This serves as a basis for introducing the empirical approach that was adopted in this study. The paper highlights a number of important issues for obtaining an accurate estimate of traffic delay using the empirical approach, including missing data manipulation, data smoothing, estimation of crash footprint, estimation of incident-free baseline traffic conditions, etc.

Methods for Estimating Delay Caused by Crashes

Several methods have been used to estimate traffic delay caused by crashes on freeways. These methods can be classified into four types: empirical estimates, queuing analysis, shockwave analysis, and freeway traffic simulation.

Considerable research with queuing theory has been devoted to analyze the impacts of incidents on freeway. Morales (1986) developed an analytical method using cumulative arrival and departure curves for estimating total delay caused by freeway incidents. Hall (1999) also used cumulative diagrams to represent non-stationary phenomena such as incident congestion. However, the linear arrival/departure curves and the assumption of a vertical stack queue at the incident bottleneck in queuing analysis are clearly unrealistic under peak period conditions.

Based on the shockwave theory, Messer et al. (1973) calculated the

wave propagation rates and travel times under incident conditions. Wirasinghe (1978) used the shockwave theory to compute traffic delay upstream of a crash on a freeway. However, the shockwave theory is based on some key assumptions as follows. 1) All vehicles in a flow state travel at constant speed and density. This assumption is likely violated, particularly within a queue as the speed and density within a queue vary with time and road geometry. 2) Vehicles change their speed instantaneously when they pass a shockwave. This assumption is clearly unrealistic and introduces errors in estimating delay.

Most traffic simulation models estimate traffic delay based on some traffic characteristics such as capacity and maximum link speed reduction caused by a crash. However, in reality there are many other factors influencing freeway crash delay such as crash characteristics, road geometries, weather conditions as well as traffic characteristics. However, there is no traffic simulation model which has the ability to directly model these factors.

As an empirical approach, Al-Deek et al. (1995) proposed a method to estimate incident delay based on the travel time difference between normal and incident conditions. Traffic data were collected from loop detectors on each freeway segment. Incident delay was calculated with respect to a reference average speed that reflects normal situations on freeway segments. To determine the scope of the impact due to an incident on a freeway, Al-Deek et al. performed shockwave analysis in the time-space domain. Then, the delay on each freeway segment was estimated using the following equations.

$$
\begin{cases}
D_j^i = L_j \dfrac{\Delta T}{60} q_j^i \left(\dfrac{1}{s_j^i} - \dfrac{1}{s_j^{i,r}} \right) & \text{for } 0 < s_j^i < s_j^{i,r} \\
D_j^i = q_j^i \left(\dfrac{\Delta T}{60} \right)^2 & \text{for } s_j^i = 0 \\
D_j^i = 0 & \text{for } s_j^i > s_j^{i,r}
\end{cases}
\tag{1}
$$

Where

D_j^i = delay on freeway segment j during time slice i (vehicle-hours);

L_j = length of segment j (km);

ΔT = length of time slice i (min);

q_j^i = flow (from loops) on segment j during time slice i (vehicle/hr);

s_j^i = speed (from loops) on segment j during time slice i (km/h);

s_j^{if} = incident-free speed on segment j during time slice i (km/h).

The total delay on the freeway section caused by an incident is

$$D_T = \sum_{i=1}^{m} \sum_{j=1}^{n} D_j^i \qquad (2)$$

where

D_T = total incident delay on freeway segments upstream of segment k affected by the incident (vehicle-hour);

m = the number of time slices with crash induced delay;

n = the number of upstream freeway segments affected by a crash;

D_j^i = traffic delay on freeway segment j during time slice i (vehicle-hour).

There are two important limitations of the Al-Deek method that should be noted.

- In Equation 1, the same flow rate is used for both the observed crash conditions and the incident-free conditions. This formulation is clearly not valid because the flow for the incident-free conditions should be the observed values on the roadway at a specific point in time in the absence of any incident.

- The method proposed by Al-Deek et al. (1995) does not consider the impact of weather. Consequently, the same reference speed is assumed regardless of the influence of the weather conditions in effect at the time of the crash. This assumption introduces error

into the delay estimation as the incident-free speed overestimated for adverse weather conditions will likely overestimate the delay.

In this study we applied an approach similar to Al-Deek and estimated crash-induced traffic delay using loop-detector data collected from QEW expressway in Toronto (52 directional detector stations in total) for the period January, 1998 to February, 2003. To address the limitations of the Al-Deek method, we used the actual flow rate as extracted from the loop detector in 20-seccond time slices for both crash and incident-free situations. External factors, such as weather, month, day of week, time of day, and location were accounted for in the estimation of crash and non-crash speed profiles.

Empirical Estimation of Crash Delay

The total traffic delay caused by a crash was calculated by the following equations.

$$D_{i,j} = \begin{cases} \dfrac{\Delta T}{3600} \cdot L_j \cdot \left(\dfrac{q_{i,j}}{\bar{s}_{i,j}} - \dfrac{q^f_{i,j}}{\bar{s}^f_{i,j}} \right) & for\ \bar{s}_{i,j} < \bar{s}^f_{i,j} \\ 0 & for\ \bar{s}_{i,j} \geq \bar{s}^f_{i,j} \end{cases} \quad (3)$$

$$D_e = \sum_{i=1}^{n} \sum_{j=1}^{m} (D_{i,j}) \quad (4)$$

where

D_e = estimated total delay caused by a crash (vehicle-hour);

$D_{i,j}$ = traffic delay on segment j during time slice i (vehicle-hour);

$q_{i,j}$ = traffic volume on segment j during time slice i (vph);

$\bar{s}_{i,j}$ = average speed on segment j during time slice i (km/h);

$q^f_{i,j}$ = average traffic volume on segment j for an hourly period under incident-free and same weather conditions (vph);

$\bar{s}_{i,j}^{f}$ = average speed on segment j for an hourly period under incident-free and same weather conditions (km/h);
L_j = length of freeway segment j (km);
ΔT = length of time slice i (20 seconds);
m = the number of upstream freeway segments affected by a crash;
n = the number of time slices with crash delay.

To estimate the only impact on traffic caused by a crash, it should be eliminated the external impacts on traffic from traffic data collected during crash situation. Eliminating the external impacts on traffic was performed by estimating the incident-free traffic flow based on historical data controlled by representative external conditions (i.e. month, day of week, weather, time of day and location). The incident-free traffic flow represents speed and volume that we expect would have been observed if the crash had not occurred.

As an initial step we only consider delay experienced by vehicles traveling the same direction where the crash occurred. The method presented in the paper can also be applied to estimate delays in the opposing direction.

Measuring Delay from Loop-Detectors

Before this method can be applied to estimate crash induced delay, a number of implementation issues must be addressed. First, since the method is empirical, we must have loop data for all time periods and all road segments affected by the crash. Second, the spatial and temporal extent of the crash induced congestion (crash footprint) must be identified so that the delay estimation method is applied to the entire road section for the entire time that traffic is influenced by the crash. Third, loop detector data influenced by missing and fluctuating characteristics must be damped or smoothed before the delay estimation. The process of estimating the crash induced delay is as follows. Here we addressed five specific issues associated with delay estimation:

- Missing data manipulation
- Data smoothing
- Determination of spatial and temporal extent of crash
- Estimation of reference (incident-free) traffic flow
- Estimation of crash-induced delay

Missing Data Manipulation

Traffic data collected from loop detectors on roadways are often stored in the form of a time series. It is not uncommon to face missing values in traffic data as a result of errors in data communication, data collecting, etc. In this study, the missing data are manipulated by linear interpolation.

Figure 1 and 2 show speed and volume data over 20-second interval collected on QEW in Toronto. Figure 1 shows the raw traffic data including missing values and Figure 2 shows manipulated traffic data. Although the missing values in speed data were manipulated by linear interpolation, the speed data are still fluctuated in Figure 2.

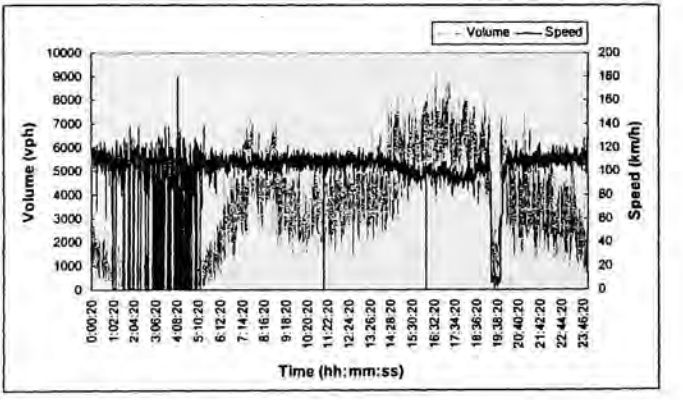


Figure 1. Raw Data

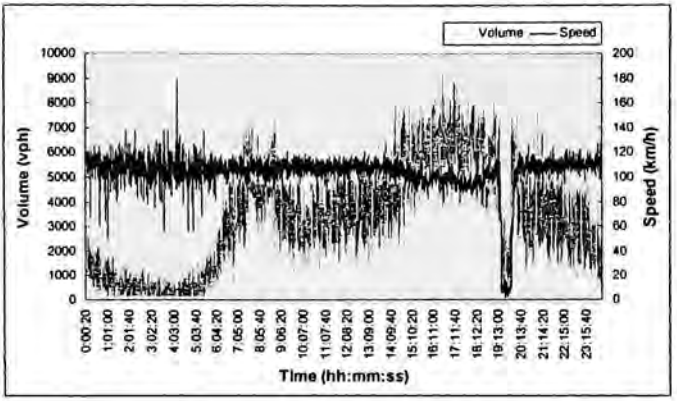


Figure 2. Manipulated Data

Data Smoothing

Raw traffic data are typically averaged over a 20-second polling interval. These data exhibit significant variation, often as large as the underlying trends. Traditionally, transportation engineers have used fixed-time average method, cumulative summation, and moving average method to reduce the variation of detector data.

However, each method has its own limitations for analyzing traffic characteristics. The fixed-time average method is not sensitive to changes in traffic states. Cumulative summation method is not appropriate to exhibit speed data. Finally, although moving average method has been widely used to reduce the fluctuations of detector data, there are some limitations: 1) determination process for the optimal span number is indistinct; 2) as the span increases, a time-lag is introduced into the smoothed data.

Because most traditional smoothing methods smear out the sharp changes in raw traffic data as well as filter out the noise, a nonlinear method, called Tukey smoothing, was used in this study. Particular form of the Tukey smoothing technique is based on a combination of

running medians and linear smoothing. The most attractive properties in this smoothing technique are as follows: 1) filtering out the noise while maintaining the underlying trend in traffic data; 2) minimum time-lag exists when a smoothing method is applied to traffic data. The Tukey technique for smoothing is applied to the manipulated data (Figure 2) that has taken into account missing data, and the results are shown in Figure 3. In this exercise, the Tukey smoothing technique was applied only to the speed data.

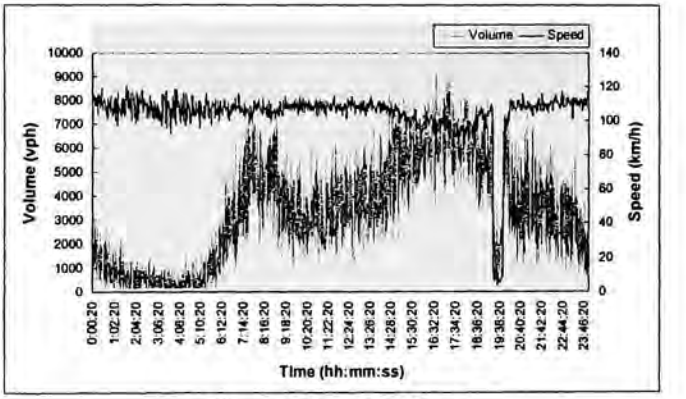


Figure 3. Smoothed Data

Spatial and Temporal Extent of Crash

Freeway crashes affect traffic delay over time and space. To estimate the delay caused by these freeway crashes using traffic data, it is essential to determine the spatial and temporal extent of crash induced congestion. The spatial extent of crash induced congestion represents the freeway segments affected by a crash or crash footprint.

Figure 4 illustrates a contour plot of speed data affected by a given crash. The crash occurred between detector station 450DWS and 460DWS at approximately 19:20. The crash-induced congestion extended upstream from the crash site to the last detector affected

approximately 4.35 Km upstream (520DWS). The shockwave from the crash reached 520DWS at time 20:00, approximately 40 minutes after the time of crash. Crash-induced congestion on the freeway disappeared at about 20:05. The time/space footprint associated with this crash is as follows:

- Congestion time interval: 19:20 to 20:05
- Detector interval: 460DWS to 520DWS

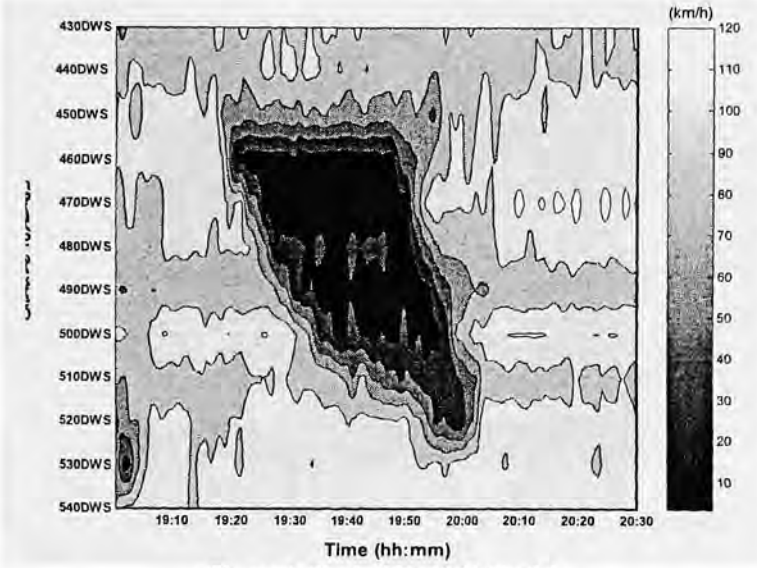


Figure 4. Typical Crash Footprint

The temporal extent of each crash can also be determined by visual inspection using the speed profile for the crash as observed at a given detector. Figure 5 illustrates speed profile for the first upstream detector for this crash at 460DWS. The profile given in this Figure has been smoothed using the method described above.

In Figure 5, the average speed of traffic flow drops dramatically at a

time of 19:17:00 and the congestion caused by the crash is observed to have dissipated by time 19:57:20. Consequently, the temporal extent of the crash (or footprint) at this detector station is 19:17:00 to 19:57:20. We repeat this analysis for all upstream detectors affected by the crash to obtain an estimate of total crash-induced delay. In order to do this we need non-crash speed and volume profile at each detector affected that serves as a basis on which to compare the speed and volume profile for the crash case. This of course is subject to all appropriate external controls as noted above.

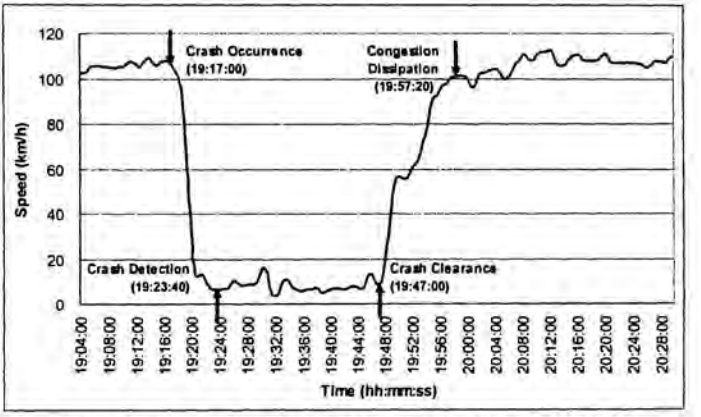


Figure 5. Speed Profile for Detector Station of 460DWS

Estimating Incident-Free Speed and Volume Profiles

The incident-free speed and volume for each detector station affected by the above crash are estimated on the basis of controlled external factors. Figure 6 illustrates the incident-free speed and volume over 24-hour for the same detector station with above Figure. The following factors represent the external conditions when the crash occurred and these factors are controlled to estimate the incident-free speed and volume. The external conditions are controlled on the basis of an hourly period. The external conditions

when the crash occurred are as follows:

- Month: January
- Day of week: Thursday
- Weather: Snow
- Time of day (an hourly period): 19:00 - 19:59
- Upstream detectors affected by crash: 460DWS - 520DWS

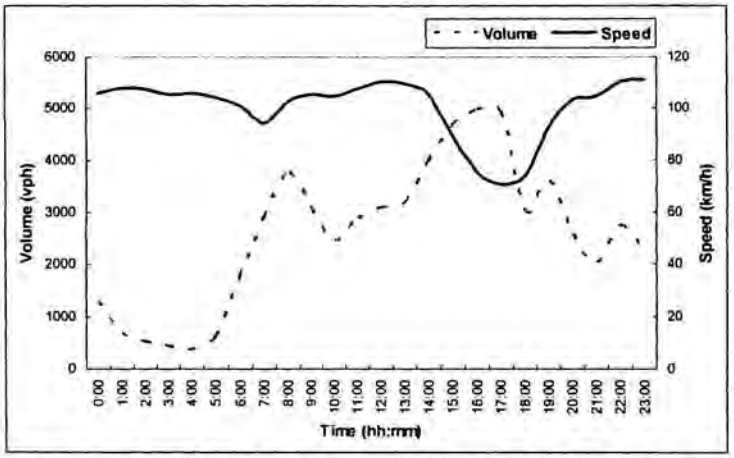


Figure 6. Incident-Free Speed and Volume Profile for Detector Station of 460DWS

Estimating Crash Delay

The temporal and spatial extents were determined by the proposed methods as shown in Table 1. The total crash induced delay was estimated by Equation 3 and 4 applied to the smoothed incident-free profiles for the crash. The total crash-induced delay estimated by the proposed method was 251.1 vehicle-hours for all upstream detectors in the crash footprint and the Al-Deek method resulted in 292.8 vehicle-hours. This result provides evidence that Al-Deek method overestimates the delay caused by a crash.

Table 1. Total Crash Induced Delay

Loop	460DWS	470DWS	480DWS	490DWS	500DWS	510DWS	520DWS
Start	19:17	19:21	19:22	19:26	19:28	19:35	19:50
End	19:57	19:55	20:05	20:05	20:01	20:04	20:03
Proposed Method	127.5	38.5	9.0	16.1	44.2	8.4	7.4
Al-Deek Method	146.6	46.8	11.1	18.5	51.5	9.8	8.4

Conclusions

In this paper, we have proposed an empirical method to estimate traffic delay caused by crashes on instrumented freeways. Used in conjunction with delay prediction models these delay estimates can be used to justify various traffic management strategies for reducing the costs associated with crashes.

The discussion has underscored how the quality of empirical loop detector data can be improved for obtaining more accurate estimates of crash-induced delays. This paper suggests different ways of dealing with problems of missing data and smoothing of observed random responses in speed and volume.

References

1. Schrank, D. and T. Lomax. *The 2002 Urban Mobility Report*, Texas Transportation Institute, The Texas A&M University System, 2002.
2. Lindley, J. A. *Quantification of Urban Freeway Congestion and Analysis of Remedial Measures*, Report RD/87-052. FHWA, US Department of Transportation, 1986.
3. Federal Highway Administration (FHWA), *Freeway Management Handbook*, Report No. FHWA-SA-97-064, Washington, D.C., 1997.
4. Lance, R. G and E. W. Clyde. *The True Cost of Highway Congestion*, ITE Journal 62, March 1992, pp. 16-20.

5. Miller, T., J. Viner et al. *The Cost of Highway Crashes*, Report FHWA-RD-91-055, US Department of Transportation, June 1991.
6. Vodden, K., D. Smith, R. Meng, T. Miller et al. *The Social Cost of Motor Vehicle Crashes in Ontario, Ministry of Transportation Ontario*, Safety Policy Branch, March 1994.
7. Blincoe, L.J., A.G. Seay, E. Zaloshnja, T. R. Miller et al. *The Economic Impact of Motor Vehicle Crashes, 2000*, NHTSA Technical Report HS-809 446, U.S. Department of Transportation, May 2002.
8. Morales, J. M. *Analytical Procedures for Estimating Freeway Traffic Congestion*, Public Roads, Vol. 50, No. 2, 1986, pp. 55-61.
9. Hall, R. W. *Handbook of Transportation Science*. Kluwer Academic Publishers, 1999, pp. 109-150.
10. Messer, C. J., C. L. Dudek, and J. D. Friebele. *Method for Predicting Travel Time and Other Operational Measures in Real-Time During Freeway Incident Conditions*. Highway Research Record 461, TRB, National Research Council, Washington, D.C., 1973, pp. 1-16.
11. Wirasinghe, S., *Determination of Traffic Delays from Shock Wave Analysis*. Transportation Research, Vol. 12, 1978, pp. 343-348.
12. Al-Deek, H. M., A. Garib, and A. E. Radwan. *New Method for Estimating Freeway Incident Congestion*, In Highway Research Record 1494, Highway Research Board, Washington, D.C., 1995, pp. 30-39.
13. Sullivan, Edward C. *New Model For Predicting Freeway Incidents and Incident Delays*, In Journal of Transportation Engineering, Vol. 123, No. 4, July/August, 1997, pp. 267-275.
14. Cohen H., *On the Measurement and Valuation of Travel Time Variability Due to Incidents on Freeways*, Journal of Transportation and Statistics, December 1999, pp. 123-131.
15. Garib, A., A. E. Radwan, and H. M. Al-Deek. *Estimating Magnitude and Duration of Incident Delays*, In Journal of Transportation Engineering, Vol. 123, No. 6, 1997, pp. 459-466.

Biography

Jaehyun Noh is a Ph.D. candidate in the Department of Civil Engineering at the University of Waterloo.

Dr. Frank Saccomanno is a Professor in the Department of Civil Engineering at the University of Waterloo

Dr. Bruce Hellinga is an Associate Professor in the Department of Civil Engineering at the University of Waterloo and an active member of the Transportation Systems Research Group.

THE CANADIAN SHIPBUILDING INDUSTRY, WORLD MARKET DISTORTING PRACTICES AND MECHANISMS TO DEAL WITH IT.

Joseph Monteiro, Gerald Robertson and Keith Dawson*

I. Introduction

A former Minister of Industry Canada stated "Shipbuilding and repair are very much a part of Canada's proud maritime tradition ... So make no mistake: shipbuilding is part of our heritage. It is not and should not be viewed as a 'dying industry'. We can and will compete successfully with the best, but on our terms, and in niches and areas we choose."[1] These are encouraging words but can these hopes be realized as "The worldwide shipbuilding industry is riddled with subsidization and unfair practices."[2] according to the Department of Transportation, Maritime Administration (MARAD). To add to this, the World Trade Organization (WTO) mechanisms to deal with world market distorting practices in shipbuilding have been a dismal failure.[3]

In this paper, an attempt shall be made to cast some light on the latest proposal by the Organisation for Economic Co-operation and Development (OECD) to deal with world market distorting practices in shipbuilding and to review the past mechanisms to deal with it. Before proceeding, the Canadian and World shipbuilding industry will be briefly reviewed. This will provide us some guidance on whether the recent OECD proposal will be successful.

II. The Canadian Shipbuilding Industry

a) The Significance of the Ship Building Industry

The significance of the shipbuilding industry can be seen in various

Table 1 - Index of Real value-added and employment in the Shipbuilding industry (1983 base year)

Year	Real value-added	Employment
1993	145	98
1994	115	75
1995	110	70
1996	100	55
1997	65	55
1998	70	55
1999	67	58
2000	60	53

Source: *A New Policy Framework for the Canadian Shipbuilding and Industrial Marine Industry, Industry Canada*, 2001, p. 29.

* *The views expressed are those of the authors and are not purported to be those of the Commissioner or the Competition Bureau. They would also like to acknowledge the assistant of Scott Barillaro in providing the statistics.*

indicators such: as shipments, real value-added, employment, imports and exports of shipbuilding parts. A snapshot of these indicators for 2000 is: $335 million, $264 million, 4,700, $555 million, and $137million, respectively. An index on real value-added and employment over the last several years are shown in Table 1. These statistics reveal that after 1993, the indices of real value-added and employment fell by 85% and 46%, respectively.

Table 2 provides more aggregate statistics that are available on shipbuilding and repairs for shipments, real-value added, employment, imports and exports. Over the period 1993-2001, shipments, real-value added and employment

Table 2 - Shipbuilding and Repairs (shipments, value-added, employment, imports and exports)

Year	Shipments ($) (M)	Value-added ($) (M)	Employment	Imports($) (M)	Exports($) (M)
1993	1, 178, 151	773, 779	8, 913	39, 434	17, 518
1994	969, 500	610, 012	7, 361	25, 358	44, 602
1995	964, 772	696, 644	6, 568	80, 223	19, 013
1996	730, 494	559, 272	5, 820	10, 078	16, 958
1997	503, 491	378, 006	4, 660	36, 833	17, 993
1998	393, 196	219, 859	4, 892	278, 447	25, 789
1999	573, 801	332, 526	6, 122	66, 549	184, 916
2000	822, 998	587, 604	5, 827	559, 699	42, 164
2001	546, 203	366, 606	4, 284	740, 644	56, 092

Source: *Statistics Canada.* NAICS 33661. *M=Million.*

declined significantly from its peak in 1993 at an annual rate of 5.9%, 5.8% and 5.8%, respectively. In contrast, shipbuilding exports grew by 24.5% annually over the period 1993-2001 which is basically a result of higher value-added products.

b) The Ship Building Industry by Province

The shipbuilding industry in Canada consists of more than 30 shipyards in British Columbia, Ontario, Québec and the Atlantic provinces. Atlantic provinces account for more than 50 percent of employment and 43 percent of

Table 3 - Shipyards and Employment in Shipyards (May 2001)

PROVINCE	NUMBER OF SHIPYARDS	PERCENT	ESTIMATED EMPLOYMENT	PERCENT
Atlantic	15	46.85	2315	49.18
Québec	5	15.62	570	12.11
Ontario	5	15.62	632	13.43
BC	6	18.75	1190	25.28
NWT	1	3.1	NA	NA
TOTAL	32	100.00	4707	100.00

Source: *A New Policy Framework for the Canadian Shipbuilding and Industrial Marine Industry, Industry Canada*, 2001, p. 31.

shipyards. This is followed by British Columbia, Ontario and Québec. The statistics on this are shown in Table 3.

c) Firms in the Ship Building Industry

The majority of the major shipbuilding firms are located in Nova Scotia (8) and British Columbia (6). Newfoundland, Ontario and Québec have 5 firms each and N.W.T., N.B. and P.E.I. have 1 each. These firms are shown in Table 4. The Davies and Saint John yards represent 50% of total Canadian capacity.

Table 4 - Firms in the Ship Building Industry by Province

Nova Scotia	NB	NewFoundland	PEI	Québec	Ontario	BC	NWT
North Sydney	Saint John Shipbuilding Ltd.	St. John's Dockyard Ltd.	East isle Shipyard	Les Industries Davie	Fraser Marine & Ind	Allied Shipbuilding	Northern Transportation Co. Ltd.
A.F. Theriault & Son Ltd.		Terra Nova Alliance/PCL		Chantier Naval Forillon	Heddle Marine	Fraser Shipyard& Ind.	
Lunenburg Industrial Foundry & Eng		Friede Goldman Nfld. Ltd		Chantier Naval Matane	Hike Metalworks	Jenkins Marine	
Pictou Industries		Burin Industries		Les Industries Océean	Metalcraft	Nanaimo Shipyard	
Halifax Shipyard		C&W Industrial Fab. & Marine Equipment		Verreault Navigation Inc.	Port Weller Dry Docks	Point Hope Shipyard	
Dartmouth Marine Slips						Washington Marine Group	
Shelburne Ship Repair							
Steel & Engine Products							

Source: *A New Policy Framework for the Canadian Shipbuilding and Industrial Marine Industry, Industry Canada*, 2001, p. 32.

d) The New Policy Framework For Canadian Shipbuilding

In 2001, Industry Canada announced a new shipbuilding and industrial marine policy that attempts to balance the interests of all parties. "...the result is a future-oriented policy providing new opportunity to the industry. With this policy, we are raising the shipbuilding and industrial marine industry to a new level within government. It is realistic, it is market-oriented, and it is fair."(p. iv) The new policy framework is grouped into five categories: 1) capturing domestic opportunities; 2) looking globally, 3) innovating as key to competitiveness, 4) financing and 5) strengthening partnerships.

The first is on: increasing efforts to secure offshore opportunities, promoting complementary investment, reviewing current restrictions on fishing vessel length, and maintaining the current 'Buy in Canada' policy. The second is on: promoting shipbuilding and industrial marine exports, reducing trade-distorting practices (by

eliminating global subsidies), consulting with industry in trade negotiations (by pressing for amendments to the U.S. *Jones Act* to permit greater access), maintaining current tariff (by addressing distortive practices), and reviewing requests for extensions for temporary vessel imports beyond three years. The third is on: developing guidelines for the application of Scientific Research and Experimental Development tax credits, facilitating access to Technology Partnerships Canada, developing a Technology Road Map, and enhancing training (i.e., assisting small and medium enterprises through Industrial Research Program, and supporting ocean engineering research, maximizing use of Internet through student Connection program, and establishing a sector council that will access skill and training requirements). The fourth is on: maintaining current accelerated capital cost allowance, providing competitive financing to purchasers of Canadian-built vessels who cannot use the former, continuing to provide competitive non-concessionary export financing through the Export Development Corporation, and considering the use of export financing. The fifth is on: establishing an industry advisory committee, organizing an annual stakeholders' forum, and creating a new energy and marine organization to work with stakeholders.

e) Regulations Regrading Water Transport and Imports of Ships

The *Coasting Trade Act* restricts to Canadian ships the transport of cargo and passengers, along all commercial marine-related activities in Canadian waters.[4] It also extends this restriction to the Canadian continental shelf for activities related to exploration and exploitation of non-living natural resources.[5] This restriction, however, has waivers as the *Coasting Trade Act* provides for the temporary importation of foreign and non-duty-paid Canadian flag vessels in specified circumstances and eliminates the uncertainties associated with their operation in Canadian waters.[6] The specified circumstances for foreign ships include the non suitability or availability of Canadian or non-paid duty ships to provide service or perform the activity, the unavailability of an identical or similar adequate service in the case of carriage of passengers by Canadian ships, the payment of duties and taxes by the applicant for the licensed service, the validity of certificates and documents pursuant to a shipping convention to which Canada is a party, and the satisfaction of safety and pollution prevention requirements.[7] Terms and conditions can be imposed on the licence relating to service or activity and place or places to which it relates.[8] The effective period of a licence does not exceed twelve months.[9]

Regarding imports of ships, Canada currently applies a 25 percent Most-Favoured-Nation (MFN) unbound tariff on most non-NAFTA imported vessels.

The main exception to this rate are fishing vessels over 30.5 metres, which do not have any duty applied when they are imported into Canada. The rate on importation of foreign flagged vessels where Canadian vessels are not available is 1/120 of the existing tariff (i.e., a reduced rate) and the allowable period is ten years. In addition, there are duty remissions on various topside modules and sub-sea components that are imported into Canada. It is worthwhile noting that U.S. built ships enter Canada duty-free while Canadian shipbuilders are prevented from selling their vessels in the U.S. under the *Jones Act* (a legislation exempted from NAFTA).

III. The World Shipbuilding Industry

a) The World Shipbuilding Industry

Statistics on shipbuilding are usually aggregated into five major categories: oil tanker, bulk carrier, container ship, LNG carrier and cruise ship. Table 5 provides statistics on the type of ships completed and the shipyards in these

Table 5 - Ship completions and Shipyards (1995-2002)

Type of Ship/Shipyards	1995	1998	2002
Oil tanker/Shipyard	248/118	241/92	274/80
Bulk carrier/Shipyard	274/64	200/52	234/47
Container ship/Shipyard	179/66	285/76	175/39
LNG carrier/Shipyard	5/3	6/5	10/5
Cruise ship/Shipyard	8/5	12/5	14/7

Source: *Dealing With Market-Distorting Practices*, OECD, C/WP6/SNG(2004) 1, p. 11.

categories. The statistics reveal that ship completions over the period 1995-2002 have increased for oil tankers, LNG carriers and cruise ships, whereas they have declined for bulk carriers and container ships. The number of shipyards have declined over this period for the construction of oil tankers, bulk carriers and container ships. This decline has been fairly significant ranging from 27% to 40%. The number of shipyards has increased for LNG carriers and cruise ships.

b) The Market Share of Shipyards and Shipbuilding Nations

The market share of the four largest shipyards and major shipbuilding nations for oil tankers, bulk carriers, container ships, LNG carriers and cruise ships are shown in Tables 6-10.

In sum, the major shipbuilding nations in 2002 are: Korea - oil tankers; Japan - bulk carriers; Korea - container ships; Japan - LNG carriers; and France/Italy -

cruise ships with market shares of 30.6%, 31.7%, 38.2%, 50% and 21.4%/21.4%, respectively. The constituent market shares of various shipyards in each nation

Table 6 - Oil Tanker World Market (2002)

Shipyard	Nation	Number	Share
Daewoo SB	Korea	22	8.0%
Samho	Korea	22	8.0%
Hyundai MIPO	Korea	21	7.7%
Hyundai HI	Korea	19	6.9%

Table 7 - Bulk Carrier World Market (2002)

Shipyard	Nation	Number	Share
Oshima	Japan	25	10.7%
Imabari	Japan	21	9.0%
Tsuneishi	Japan	15	6.4%
Mitsui	Japan	13	5.6%

Table 8 - Container Ship World Market (2002)

Shipyard	Nation	Number	Share
Hyundai HI	Korea	23	13.1%
Samsung	Korea	21	12.0%
Hanjin	Korea	16	9.1%
China SB	Chi. Taipei	7	4.0%

Table 9 - LNG Carrier World Market (2002)

Shipyard	Nation	Number	Share
Mitsubishi	Japan	4	40%
Daewoo	Korea	2	20%
Hyundai HI	Korea	2	20%
Mitsui	Japan	1	10%

Table 10 - Cruise Ships World Market (2002)

Shipyard	Nation	Number	Share
L'Atlantique	France	3	21.4%
Fincantieri	Italy	3	21.4%
Litton Systems	USA	2	14.3%
Meyer	Germany	2	14.3%

Source: Based on Lloyd's Fairplay. OECD Paper C/WP6/SNG(2004) 1.

rarely exceed 10% except for the LNG carrier and cruise ship markets.

c) Canadian Share of the World Market

Ships built in Canada in 2002 are shown in Tables 11 and 12. The statistics reveal that Canada's share of the world market is .03% which is insignificant. Canada did not manufacture any ships in the major categories i.e., tankers,

Table 11 - All Ships World Market (2002)

Nation	Number	Share	Gross Tonnage
PRC	613	30.85%	2, 836, 100
Japan	401	20.18%	12, 201, 926
Korea	208	10.46%	12, 771, 078
Norway	75	3.77%	286, 100
Canada	6	0.30%	6, 598

Source: Compiled from *Sub-Group on Supply and Demand*, OECD , C/WP6/SG(2003)6.

bulk carriers, container ships, LNG carriers and cruise ships in 2002. Its production was confined to passenger ships and other non-cargo vessels. An examination of Canada's market share in the areas of its production reveal that in 2002 it was marginally better accounting for 2.6% and 2.2%.

Table 12 - Passenger Ships and Other Non-Cargo Vessels World Market (2002)

Nation	Number of Passenger ships	Share	Number of other non-cargo vessels	Share
Norway	10	26.3%	39	17.3%
Germany	9	23.7%	5	2.2%
PRC	7	18.4%	89	39.5%
Italy	4	10.5%	4	1.8%
Japan	1	2.6%	31	13.8%
Spain	0	0.0%	22	9.8%
Canada	1	2.6%	5	2.2%

Source: Compiled from *Sub-Group on Supply and Demand*, OECD , C/WP6/SG(2003)6.

In sum, Canada accounts for an insignificant share of world shipbuilding.

d) Recent Developments in World Shipping Policy (EU) and Issues they Raise

In 1997, the European Union updated its *State Aid Guidelines*.[10] The guidelines authorized: fiscal treatment (i.e., it is aimed at creating fair competition with third flag ships and supporting EEC flags); treatment with respect to reducing labour related costs to shipowners; accelerated depreciation; deferred taxation; tonnage tax; social security cost relief; seafarers income tax relief; relief in relation to crew repatriation; aid to safety, environmental standards and training; etc.[11]

These developments raise a number of issues[12]: Will Canadian participation

and involvement in international shipping enhance Canadian know-how, expertise, safety and environmental protection? Is it necessary to provide a fiscal environment that is essentially 'zero tax' (i.e., levelling the competitive playing field) to remain competitive. To what extent has Canada's initiative to provide state aid to achieve a 'zero tax' for certain international shipping activities been successful? Under what circumstance is a 'tonnage tax' preferable to a 'zero tax' regime? Since Europe has extended various forms of aid to most forms of domestic maritime transport, should Canada adopt the same measures to level the competitive playground? Finally, special treatment of shipping raises issues of the need for similar treatment with other modes of transport?

IV. Distorting Practices in the Shipbuilding Industry

The shipbuilding industry has been profoundly affected by the support of governments to assist their industries and by pronounced cyclicality which has led to market distorting practices. The term 'support' is used as a general expression for subsidies, state aids and national procurement. The term 'market distorting practices' is used as a general expression for any measure which diminishes competition, interferes with the commercial workings of the market, and deals with instances when firms may abuse their market position together with pricing practices.[13] These distorting practices will be briefly reviewed.

a) Subsidies and State Aid

Subsidies and state aid interferes with the workings of the global market and could diminish or eliminate competition. The Department of Transportation, Maritime Administration (MARAD) states that "The worldwide shipbuilding industry is riddled with subsidization and unfair practices. MARAD reveals that dozens of countries provide subsidies in one form or another to their shipyards."[14] Not surprisingly, the Agreement[15] lists eleven measures of support that should be eliminated. These include: export subsidies; grants; loans on terms and conditions more favourable than those obtainable on the market; loan guarantees on the above absent the government guarantee; forgiveness of debt; provision of equity capital inconsistent with usual investment practice; provision of goods and services at less than the adequate remuneration; tax policies and practices benefiting the shipbuilding and repair industry; assistance to suppliers of goods and services to the shipbuilding and repair industry if it provides benefits to the shipbuilding industry; official regulations and practices that discriminate in favour of domestic shipbuilding; and, other assistance. There are, however, five exceptions to the

support measures: export credits; home credits; research and development; assistance to workers; and restructuring support. The first four of these are permanent and the fifth was not to be permitted on new programmes.

b) National Procurement

National procurement practices which limit the pool of potential suppliers have an important negative effect on the level of competition and on the efficiency of such purchases. National procurement is a common practice in many industries. The EC estimates that procurement by public authorities alone account for 11% of the EU's GDP. In shipbuilding, national procurement policies of many nations are designed to counter foreign subsidies and unfair trade practices. For example, the existing policy in Canada states that any additions or replacements to the Canadian fleet will be sourced in Canada, similarly the *Jones Act* effectively closes the domestic market to foreign-built vessels. Eliminating such practices besides correcting distortions in competition and enhancing efficiency of purchases will minimize attempts to collude, to rig bids and to control corruption and political favouritism.

c) Market Dominance[16]

Abuse of dominance or monopoly usually results in high prices and reduced output, extortionate profits, and anti-competitive activities designed to destroy weaker competitors. In addition, monopolists are often accused of the inability to innovate or continually innovate.[17] The mere fact of holding a dominant position or a joint dominant position is not considered per se *abusive*. Just certain types of anti-competitive practices. These anti-competitive practices reduce or eliminate competition thereby affecting the optimal production of goods in the market. For Canada, however, under the abuse provision of the *Competition Act,* three criteria are required to be met: 1) a substantial or complete control of a class or species of business (i.e., the definition and control of the relevant product); 2) a practice of anti-competitive acts (i.e., these acts would also include rebates and tie-in agreements), and; 3) a preventing or lessening competition substantially in a market (i.e., whether the above practices have had, are having or are likely to have this effect).

d) Market Distorting Practices[18]

Five pricing practices are proposed by the OECD as a form of price discipline in the shipbuilding market. The possible reasons behind this proposal are

considered together with its relation to the *Competition Act*.

1) Predatory Pricing: Predatory pricing is the use of low prices to inhibit or eliminate existing or potential competition for the purpose of dominating or monopolizing a market. Such a strategy may be profitable in the short-term if losses can be recouped due to supra-competitive profits after the exit of the competitor or target firm. If the predator is successful in establishing a monopoly it could not only result in reduced output and higher prices but could also lead to a reduction in innovation. Current thinking emphasizes the scope for recoupment, the roles of reputation and signalling in facilitating recoupment by the predator. Distinguishing predatory from non-predatory price cuts are not always easy. Various tests have been put forth.

Most earlier tests in predatory analysis (Areeda and Turner, Williamson, Baumol, Joskow and Klevorick, Scherer, Haden, and the Canadian *Predatory Pricing Guidelines*[19]) considered pricing below marginal cost predatory even though they did not use the marginal cost test. Their analysis (other than Baumol and Williamson) used an average variable cost test as information on marginal cost was not available. Baumol used a price reversal test and Williamson used an output expansion test. According to the Canadian *Predatory Pricing Guidelines* prices 'above average total cost' are not unreasonably low. Prices between 'average total cost' (ATC) and 'average variable cost' (AVC) (ATC lies above the AVC) requires consideration of surrounding circumstances e.g., excess capacity, demand strength, etc. Prices 'below average variable cost' are unreasonably low unless clearly justified. Therefore average variable cost forms a boundary as to when selling below average total cost becomes market-distorting. The Competition Bureau's is revising its *Guidelines* to keep abreast of recent development in economic thinking in two areas: the use of avoidable cost (to take account of multi-product firm costs); and the role of recoupment.[20] When these considerations are applied to the shipbuilding industry with its excess capacity, the company would incur a loss and sales would be irrational, unless recouped in the future at any sales below marginal cost near the profit maximizing output.[21]

2) Rebates and tie-in agreements: Rebates and tie-in arrangements are a normal part of business but they can also hinder the maintenance of effective competition. These practices are considered to be a type of abuse of market dominance and are addressed by most national competition or anti-trust laws for reasons provided under abuse of dominance. In Canada, the three criteria that are required to be met under the abuse provision of the *Competition Act* have been described under market dominance. Such practices are common in shipping conference agreements.

3) Collusive pricing and price fixing agreements: Collusive pricing and price fixing agreements restrict the extent of price competition. By coordinating their pricing and production activities members of the agreement or cartel increase their collective and individual profits by restricting output and raising market price. This reduction in output and increase in price not only results in a transfer of consumer surplus from consumers to the cartel but also a loss in total welfare as potential demand is reduced. Collusive arrangements could take various forms. In Canada, under the *Competition Act,* two basic elements must be satisfied for an offence: an agreement or intent to form an agreement; and an undue prevention or lessening of competition (caused by the agreement). In the economic analysis of such matters once it has been established that the accused entered into an agreement to lessen competition two steps are used. First, one begins by determining if the parties to the agreement possess market power or are they likely to obtain market power by virtue of the agreement. In determining this matter, market share alone is not sufficient to demonstrate market power. It is also important to determine if barriers to entry exist as it enables one to identify whether the accused have market power. Second, the agreement must be or must have been likely to injure competition. All agreements may not do this eg. exchange of product standards, or defining trade terminology, etc. Whether there is over capacity or under capacity or the potential for supplying substitutes is not relevant.[22]

4) Rigging of bids: Bid-rigging has a negative effect on the level of competition and on the efficiency of such purchases. It raises the cost to the purchaser or the price that the tenderer has to pay for the product. It not only results in a transfer of consumer surplus from consumers to bid-riggers but also a loss in total welfare if potential demand is reduced. As a result, bid-rigging is prohibited by most jurisdictions. In Canada, bid-rigging is a *per se* offence under section 47 of the *Competition Act*. The bid-rigging provision contains three sub-sections relating to: 1) definition, 2) offence, and 3) exception. The two elements of the definition are: a) proof of an agreement or arrangement {i.e., i) not to submit bids, or ii) in the submission of bids}; and b) the agreement or arrangement was not made known to the person calling the bids or requesting the bids at or before the time when the bids were made. The exception is to show that the agreement or arrangement is not between affiliates. Whether bid-rigging unduly lessens competition or is in the public interest or has a purpose is not relevant. Bid-rigging in the international shipbuilding market should be prohibited making it consistent with the laws of most nations.

5) Cross-subsidisation in large conglomerates: Cross-subsidization is generally

not prohibited. In certain circumstances, cross-subsidization could diminish competition and interfere with the market. It could eliminate competitors in the cross-subsidized market; it could prevent other competitors from entering these market or alternative products from being developed; it could prevent users from obtaining lower prices in which profits are made to cross-subsidize other markets; and it could prevent these products from being competitive with alternatives. However, under certain circumstances it could have beneficial effects for instance when used to enter the market or develop products or achieve economies of scale by increasing demand. Cross-subsidisation 'by itself' should not be considered a market distorting practice. It is therefore not surprising that the *Competition Act* does not contain any specific provision designed to deal with cross-subsidization.

V. Mechanisms To Deal with Market Distorting Practices

a) WTO Anti-dumping Code

The best known current mechanism for dealing with pricing practices that are considered to distort the world market is through the WTO anti-dumping discipline. Two criteria are needed: sale at below normal value and injury to the industry.

There are problems in applying the anti-dumping code to shipbuilding. First, the WTO anti-dumping discipline requires a series of goods to be produced. Ships are rarely produced in series and exported to buyers in the same country. The purchase is one-off orders and is sufficiently different to treat them as a series. Furthermore, it is not possible to apply anti-dumping duties to future sales as buyers rarely purchase multiple units. Second, an essential element of a WTO-type anti-dumping action is that it affects goods for on-sale. Since there is generally no 'import for resale' this may not cause injury to local competitors selling domestically produced equivalents. Further, since new ships are generally not registered under the flag of the owner's country but in an 'open register', the notion of 'importing' the vessel does not easily apply.

In sum, the WTO's anti-dumping approach is not ideally suited for shipbuilding, as the industry does not readily meet the requirements normally necessary for an anti-dumping action to be successful. In addition, applying a penalty is difficult. The anti-dumping code relies on placing a penalty on imports as the means of curbing future imports of goods at dumped prices. This type of remedy creates serious difficulties. Since the remedy would have to apply to that shipbuilder or ship buyer rather than to future production. This is something that the WTO model anti-dumping code cannot do. Finally, there are differing views on the economic

rationale for anti-dumping. Proponents, argue the need to deal with pricing actions that injure, and opponents argue that it is not unusual for products to be sold below their normal value from time to time.

b) The 1994 Agreement - Injurious Pricing Code

The 1994 Agreement contains a pricing discipline (the Injurious Pricing Code) which seems appropriate for the shipbuilding industry. First, the non-serial production of ships was recognised by making the IPC (Injurious Pricing Code) apply to the specific shipbuilder that caused the injury. This was done by imposing a charge on the shipbuilder in the first instance, and eventually denying on-loading/off-loading privileges to vessels constructed by the same shipbuilder if the charges were to be ignored. Because the penalty is applied to the seller, it avoids the problem of vessels generally not being imported but simply sold to end users. These specific features of the 1994 IPC would appear to address all the deficiencies implied by a traditional anti-dumping code. This suggests that the IPC was effectively designed to achieve the specific result of providing a remedy for the sale of ships at below normal value. The IPC was not endorsed by Canada though it was supported by the U.S. but not ratified. It is therefore not yet in force.

c) A Proposed Alternative Mechanism - National Competition Laws

A recently proposed alternative mechanism is based on national competition laws. Actions based on competition laws could enhance competition and offer a viable form of price discipline in the absence of some kind of injurious pricing code. Competition and anti-trust laws would operate independently outside any new shipbuilding Agreement. Competition laws of each country are different and the extent of protection would differ from one country to another. This divergence may cause some countries from participating in the Agreement. It is also worthwhile noting that competition law is not designed to deal with low prices unless there is an anti-competitive element. Anti-competitive behaviour associated with prices fall into three categories: *price collusion*; *predatory pricing*; and *abuse of dominant position*. The difficulty is whether any form of legal action could be brought against the seller. Based on one transaction it would be very difficult to prove anti-competitive behaviour even if it were possible to bring action. Further, in the absence of a specific enforcement mechanism, it would be quite difficult to enforce decisions of antitrust authorities across different jurisdictions.

Though it has been acknowledged that a strict WTO style anti-dumping scheme would not work in shipbuilding because of the peculiarities of that industry, the

question is whether this problem could be resolved by adopting an antitrust model? First, a pricing discipline could be abused by administrations intent on protecting their industries rather than preserving competitive conditions in the industry.[23] Second, there are fundamental differences (philosophies, definitions, evidentiary requirements, remedies, etc.) between the antitrust laws of each nation. Third, it is unlikely that each nation is going to add provisions to its competition or antitrust legislation that are specific to one industry. Fourth, establishing an international competition agreement subject to a dispute settlement or specifically including a list of pricing practices is likely to be objectionable to world competition agencies. Finally, since the market share data of individual shipyards are quite low, an antitrust approach would not be a practical mechanism that would be successful in dealing with market-distorting practices.

d) Canada's Existing Shipbuilding Policy to Combat Foreign Practices

Given the difficulty of getting a mechanism to deal with these undesirable practices, Canada's has adopted a shipbuilding policy to combat them consisting of the following key elements [24]: 1) A domestic federal fleet procurement; 2) A twenty-five percent tariff on non-NAFTA vessels imported into Canada; 3) A accelerated capital cost allowance i.e., a four-year full write-off; 4) An export financing plan of up to 80 percent over 12 years; and 5) Access to the Institute for Marine Dynamics. It has been suggested that this policy falls short for several reasons: there are exceptions to the domestic procurement of the federal fleet; the 25% import tariff is ineffective against subsidies of 40% as a tariff of 67% would be required; the accelerated capital cost allowance is restricted to too few Canadian investors; the export financing available is ineffective in battling foreign subsidies and in competing with Title XI financing from the U.S. (i.e., the repayment term is up to 25 years; the level of financing is up to 87.5% of the cost of the vessel; the interest rate is negotiated on a case-by-case basis; and the term of interest rate can be fixed for the entire term of the loan); and a high percentage of the graduates from the Institute are lost to U.S. employers.

As a solution, a number of proposals were made[25]: 1. Eliminating exceptions to the existing Canadian shipbuilding policies such as tariffs and federal procurement; 2. Strengthening industry and government partnerships - focussing more on high prospect areas; 3. Improving the planning process on federal procurement; 4. Using non-concessionary financing mechanisms such as extending terms to better meet buyer's needs; 5. Making tax transfer and withholding tax-exempt financing options available to Canadian-built vessels; 6.

Allowing accelerated capital cost allowance to be used by either Canadian owners or Canadian operators; 7. Pressing for elimination of foreign subsidies and unfair pricing practices; 8. Negotiating for relaxation of the restrictive conditions of the *Jones Act*; 9. Making federal and provincial government programs more accessible; and 10. Designating the industry as a national priority and promoting an environmentally friendly alternative.

VI. Concluding Remarks

First, Canada accounts for a very small portion of the world shipbuilding market. The realities of the marketplace dictate that it would be better for Canada to concentrate its resources on niche markets (i.e., high-tech ships, etc.) where it has a comparative advantage. It is more likely to succeed in these niche markets especially with the WTO's lack of success in dealing with world market distorting practices. "A more cynical view may be to consider Darling's recommendations and the Coasting Trade Act actually conceding shipbuilding to other countries, with a view that if Canada cannot build them, at least it can profit in terms of a duty."[26]

Second, restrictive regulations that impede competition should be eliminated bilaterally or multilaterally if possible. Examples include, negotiations with the U.S. to eliminate the restrictive conditions under the *Jones Act* regarding imports of ships, and elimination of cabotage restrictions under the *Jones Act* which deny access to Canadian ships resulting in an elimination of competition.

Third, every effort should be made to provide or negotiate conditions that will level the playing field for Canadian shipyards in the world shipbuilding market. A number of proposals were made in the report *Breaking Through* described earlier. Emphasis should be on a positive approach, for example attempts at obtaining concessions similar to those provided by Canada rather than eliminating opportunities for competition such as reducing waivers to the importation of ships, increasing tariffs, eliminating exceptions and eliminating or reducing government support.

Fourth, the realization by the OECD that the existing mechanisms are inadequate to deal with market distorting practices in the shipbuilding industry is noteworthy. Further, the proposal to deal with them using an antitrust model is an innovative idea, however, it is doubtful whether such an approach would be successful or would be widely supported by antitrust agencies in most nations. This view is supported by market shares which rarely exceed ten percent for individual shipyards. An alternative should be sought, perhaps extending the 1994 Agreement by phases.

Finally, recent developments in international shipping policy raise a number of issues that merit attention. To remain competitive, we believe that Canada should attempt to negotiate with other countries to remove subsidies and state aids where possible rather than providing state aid and subsidies to provide a zero tax regime. Distorting the pricing mechanism, under certain circumstances, may be beneficial if it will provide long term benefits. Fostering technological change may sometimes call for it. But we believe that a policy of distorting the market mechanism, if used at all, should be used with great caution only for a short period of time.

Bibliography

1. *A New Policy Framework for the Canadian Shipbuilding and Industrial Marine Industry*, Focusing on Opportunities, 2001.
2. *Breaking Through*, Canadian Shipbuilding Industry, October 20, 2000.
3. *Distortive Pricing Practices*, OECD, (24-Sept-2002) C/WP2/SNG(2002)3.
4. *A Pricing Discipline For The Shipbuilding Agreement An Economic Perspective*, OECD, (20-Mar-2003) C/WP6/SNG(2003)4.
5. *Dealing With Market-Distorting Practices*, OECD, C/WP6/SNG(2004)1.
6. *Sub-Group on Supply and Demand*, OECD, C/WP6/SG(2003)6.

Footnotes

1. The Honourable Brian Tobin, Speaking at the Marine Industrial Forum, St. John's, Newfoundland, *Breaking Through*, Canadian Shipbuilding Industry, October 20, 2000, p. 4.
2. *Breaking Through*, Canadian Shipbuilding Industry, October 20, 2000, p. 20.
3. *Distortive Pricing Practices*, OECD, (24-Sept-2002) C/WP2/SNG(2002)3.
4. *Coasting Trade Act*, s. 3. Under section 663(1) and (2) of the former *Canada Shipping Act*, trade in Atlantic waters was open to Commonwealth registered ships and the change in the *Act* to permit only Canadian registered ships resulted from the proposal in the discussion paper *FREEDOM TO MOVE*. This change instead of opening entry did just the opposite and was in contrast to the general deregulation policy in the above paper.
5. *Id.* s. 2f.
6. *Id.* ss. 4(1) and 5(1).
7. *Id.* s. 4 (a)-(e).
8. *Id.* s. 6(1).
9. *Id.* s. 6(2).
10. It recognized that: free market means access to all; shipping is able to seek out the lowest cost environment; low cost can result in reduced safety, environmental quality, etc.; reduced involvement leads to loss of employment, know-how, etc.; aid measures are needed to close the gap; and maritime measures are long term. See Mary Brooks, Dalhousie University, *Transport Canada Seminar*, Oct. 3, 2003.
11. The EU conditions are: the strategic and commercial management of shipping must be in the State, the activity must contribute substantially to EU economic activity and employment; the aid must be

essential to achieve the above; the beneficiaries must be liable to payment of corporate tax; and the aid must benefit the shipping sector. See Mary Brooks.

12. See Mary Brooks, Dalhousie University, *Transport Canada Seminar*, Oct. 3, 2003.

13. *Dealing with Market-Distorting Practices*, OECD, C/WP6/SNG(2004)1, p. 3.

14. *Breaking Through*, Canadian Shipbuilding Industry, October 20, 2000, p. 20.

15. Agreement Respecting Normal Competitive Conditions in the Commercial Shipbuilding and Repair Industry.

16. *Dealing with Market-Distorting Practices*, OECD, C/WP6/SNG(2004)1, p. 4.

17. Jewkes, J, "Monopoly and Economic Progress," *Economica*, August 1953, p. 208.

18. *Dealing with Market-Distorting Practices*, OECD, C/WP6/SNG(2004)1, p. 3.

19. This has been recently revised to take into consideration changes in economic thinking concerning low-pricing behaviour. See *Enforcement Guidelines for Illegal Trade Practices*: Unreasonably Low Pricing Policies, March 8, 2002.

20. Recently, changes have been made to the Guidelines, "..., in carrying out the cost-revenue analysis to determine below-cost selling, the Bureau will use 'avoidable cost' as opposed to average variable cost and average total cost used in the previous guidelines. It is now recognized that average variable cost is not appropriate for the analysis of a firm producing multiple products. Accordingly, avoidable cost is the appropriate standard which will be used in the Bureau's analysis addressing both single-product and multi-product firms." *Enforcement Guidelines for Illegal Trade Practices:* Unreasonably Low Pricing Policies, March 8, 2002, p. i.

21. It is worthwhile noting that the 'avoidable cost' test is used to determine below-cost selling in the airline industry. See *Competition Bureau Clarifies Enforcement Approach in the Airline Industry*, September 23, 2004, www.ic.gc.ca.

22. It is believed that the existence of over-capacity is likely to create greater incentives to collude so as to prevent prices from falling which is a consideration relevant to the shipbuilding industry. 23. *A Pricing Discipline For The Shipbuilding Agreement An Economic Perspective*, OECD, (20- Mar-2003) C/WP6/SNG(2003)4, p. 9.

24. *Breaking Through*, Canadian Shipbuilding Industry, October 20, 2000, p. 21.

25. *Id.* pp. 9-10.

26. Hanson, Trevor, R., *The Effect of the Darling Report on Coasting Trade Legislation*, CTRF 2004 Proceedings, p. 639.

Midstream Opportunities and Challenges for Canada's West Coast Ports.

1. Introduction

Sletmo (1999) observes that ports display a discernible life cycle consisting of early growth, development, and maturity phases. If ports fail to capture new business opportunities, decline sets in and a port ceases to be part of a modern logistics network. Outside of the transportation sector, the 2002 *Canada Marine Act Review* fostered little public debate about what role our ports could and should play in a modern logistics network. The Federal Transport Minister provided a rather narrow range of instructions to the review panel, and the government subsequently failed to act on many of the recommendations. The implied maritime policy of the federal government is that existing port foundations (i.e., dry bulk cargo, inter-modal traffic, and cruise ship travel) are sufficient to maintain Canada's trade competitiveness.

More recently (2005), the *British Columbia Ports Strategy* was released, highlighting the infrastructure needed by Canada's West Coast ports to fully capture the opportunities in containerized cargo flows generated by Chinese and Asian economic growth. This represents a significant accomplishment when considered in light of the federal government's implied policy direction. However, the B.C. strategy document is almost silent with respect to the planning and infrastructure implications that may arise from new trade opportunities generated by industrial growth in the ports' immediate hinterland.

Michael Ircha has argued persuasively that the newly commercialized Canadian Port Authorities need to be engaged in strategic planning.

Ircha's (2000) work highlights the need for port planners to focus both on the needs of existing customers, and to evaluate new business lines. The purpose of this paper is to use growth vector analysis to evaluate two particular new business lines: that arising from Alberta's oil sands production, and a potential B.C. offshore oil and gas industry. In doing so, it will explore the logistics, transportation, and infrastructure challenges that await ports wishing to participate in this specialized segment of the maritime sector.

Port executives must choose from a wide variety of strategic and competitive analysis tools. One particularly useful tool is growth vector analysis (GVA). GVA provides a disciplined approach to framing market strategies. The term growth vector refers to the strategic direction a port is taking in relation to both its cargoes ands markets. The result of the GVA analysis can be presented in a simple two row and column table that lists the market penetration, market development, cargo development or cargo diversification options available. By clearly identifying the market (expanded, or new) and cargo (improved, new) options available the vessel, cargo handling, upland terminal, and port infrastructure requirements can be readily identified

In the energy industry, the upstream sector is where oil and natural gas are found and extracted. The downstream sector is where oil is refined and processed, and where natural gas and finished petroleum products are distributed to industrial or retail markets. The midstream stream sector is the logistics system that link upstream oil and gas production with downstream refining and distribution. For example, the function of a bulk liquid crude terminal is to transfer cargo between an inland transmission pipeline or coastal transport mode, and ocean going vessels.

Canada's West Coast ports generally have sufficient water depth for ships alongside the berth and navigation channels, wharves that can handle the loads traversing them, cargo-handling equipment, storage facilities, and commercial processes to handle the dry bulk, break-bulk, and container traffic that dominate the cargo mix. Truck and rail are the dominant inland transport modes used to move cargo to and from the port. However, a liquid bulk petroleum terminal, or an offshore oil and gas port, has its own unique requirements. For example, the offshore oil and gas logistics system consists of three major components: the offshore platform, the port, and an inland transportation network. The transport option chosen to move oil or gas from the offshore (shuttle tanker or pipeline) location will result in the need for specialized port infrastructure. Offshore oil and gas ports also serve as industrial sites for large shipyards, equipment fabrication/repair, and value-added processing activities (Jayawardana and Hochstein, 2004). Ports wanting to participate in the midstream sector thus need to offer advanced handling facilities, access for large liquid bulk carriers, and sufficient upland property to support industrial activity. The provision of this infrastructure type is capital intensive and requires long lead times for planning and construction. Thus, it is important for port strategists to carefully consider all the relevant options.

2. Alberta Oil Sands

Canada is the only Group of Seven nation with the resource base to support significant growth in crude oil production. The vast majority of Canada's oil reserves are found in Alberta. Alberta has 176.1 billion barrels of oil, or about 14 percent of world reserves. The Alberta Energy and Utilities Board (EUB 2004) determines the level of established reserves based on what is recoverable under current technology and present and anticipated economic conditions as proved by drilling, testing, or production, plus a portion of contiguous recoverable reserves.

2.1 Midstream Opportunities

Alberta's current crude oil production of 1.6 million barrels/day accounts for 66 percent of Canada's total oil production. Alberta's total crude oil production is forecasted to grow to 2.6 million barrels/day in the next eight years, a 60 percent increase (Alberta Energy and Utilities Board forecast period 2003 to 2013). Since 1994, marketable oil sands production has more than doubled, surpassing 1 million barrels/day in 2004. Dunbar (2004) estimated the oil price required to justify new oil sands investments, and concluded that Alberta's oil sands are a relatively high-cost source of oil compared to oil from the Middle East. Investors could expect attractive returns on projects when the price for West Texas Intermediate oil was in the range of US $25/barrel. Substantial oil sands production growth would be expected under this moderate crude price scenario (6.8 percent per year from 2002 to 2017). Under a high price scenario of US $32/barrel, oil sands production growth would be 8.7 percent per year over the 2002 to 2017 period.

In 2000, three-quarters of Canada's oil sands production was delivered to domestic refineries, all of which are operating at or close to capacity (National Energy Board, 2000 p. 66). Traditionally, the largest export markets for Alberta crude oil are the Midwest and Rocky Mountain regions of the United States. American refineries in these areas are capable of absorbing a substantial increase in crude supplies from Alberta. However, other potential market regions include the US west coast, and the Far East (Alberta Energy and Utilities Board, 2004 p.2-16). While crude oil production growth in Alberta is not without its risks, it nevertheless represents a significant to opportunity for a few West Coast ports to expand their trade.

2.2 Midstream Challenges

Crude oil exports are generally moved by pipeline to the coast, where the bulk liquid cargo is loaded into tankers at a terminal. The most basic port terminal requirements would be a tank farm for temporary oil storage and a jetty, or single buoy mooring, projecting into deep water where large tankers can load cargo. From the port, a tanker would deliver directly to a refinery, or to a crude oil terminal linked to refineries by a pipeline (Stopford 1995).

The steps in the bulk crude logistics process illustrate that pipeline economics and vessel characteristics are the two primary factors that determine whether crude oil exports are a potential business opportunity for a port. Ports that can demonstrate that their physical location represents the lowest cost expansion, or new pipeline route, will be at a competitive advantage. To understand why this is so, it is necessary to consider the logistics system used to move oil to market.

In Alberta, the various feeder pipelines that gather and move oil sands production from northern Alberta converge at two main hubs in the Edmonton region. The Edmonton hub has two main transmission pipelines. Kenbridge Pipelines Inc. is the major carrier of crude oil to eastern Canadian and U.S. markets, while Teresan's Trans Mountain Pipeline System moves crude oil to the west coast. The Alberta Energy and Utilities Board (2004 p. 2-13) and the Canadian Association of Petroleum Producers (2005) both report that the expected increase in oil sands production will result in pipeline availability constraints and additional pipeline capacity would be required.

Crude oil is a globally traded commodity. Producers do not set the price, rather crude oil flows to markets that provide the highest value.

Canada's traditional oil trade pattern reveals that it flows to the domestic market first, then to the U.S. because these markets have the lowest transportation costs. Therefore, resource developers are provided with the highest net revenue. In the oil and pipeline business, this is referred to as the 'highest netback.' Increased oil sands production not only brings with it the possibility of existing pipeline capacity constraints but also increasing supply to existing markets, thereby depressing the price. Bulk liquid crude exports using a Canadian West Coast port represent an opportunity to solve the pipeline capacity problem, and simultaneously diversify the export market. However, it is important to note that as the distance to new markets increase, the pipeline shipping costs borne by oil sands producers also increases. As a result, the netback is reduced. Port planners thus need to carefully assess the economic factors driving new pipeline investment decisions.

The 1,150 km Treason Pipeline Trans Mountain pipeline transports crude oil and refined products from Edmonton to marketing terminals and refineries in Greater Vancouver and Puget Sound in Washington State. The $2.1 billion Trans Mountain Pipeline Expansion (TMX) represents a cargo and market extension for the Vancouver Port Authority because currently the Terasen Westridge Marine Terminal is the only facility on Canada's West Coast that can ship crude oil by ocean-going vessel. The company has indicated that they believe the markets for Canadian crude oil on the West Coast are Washington State and California.

Terasen's three-phased approach to their TMX project first involves constructing a 178 km-section of 32-inch pipe looping the east end of the Trans Mountain system between Hinton and Jackman, B.C., several pump stations, and tank facilities. The second phase will bring future increases in pipeline capacity by adding new pumping stations through to the port of Vancouver.

Terasen may also offer a third expansion option. This would be a northern leg to the port of Prince Rupert or Kitimat. These ports are capable of accommodating a Very Large Crude Carrier. The first phase is planned for completion in 2008, with full expansion capacity available in 2011.

The proposed 1,200 Enbridge Inc. Gateway pipeline proposal would transport crude oil and refined products from Edmonton to a marketing terminal at either the Port of Kitimat or Prince Rupert. The $2.5 billion Gateway proposal represents both a cargo and market diversification opportunity for both ports because neither one currently has a marine terminal that can ship crude oil by ocean-going vessel. Enbridge's CEO has indicated that China, Japan, South Korea, and California would be the target markets. The project would require that a new 30-inch crude oil pipeline, bulk liquid crude storage facility and marine terminal be built. The company is aiming for an in-service date between 2009 and 2010.

Gilje, Dinwoodie, & Challacombe (2002) observe that the longer the voyage distance, the more economical efficient larger vessels will be until geographical and design constraints are encountered. High value cargoes suit smaller vessels due to increased financial penalties associated with holding large stocks, and conversely, lower value cargoes suit larger vessels. Port planners must therefore carefully consider economic the trade-offs involved in economies of large cargoes at sea against the storage, handling, and port costs incurred.

There are three main sizes of oil tankers. Popular Aframax vessels were initially 80,000 dwt (dead weight tons) but are moving to a standardized 105,00-107,000 dwt today. Very large crude carriers are over 200,000 dwt and require a water depth of about 20 meters. The larger Ultra Large Crude Carriers (ULCC) is over 300,000 dwt, with a loaded draught of 22 meters.

However, the size of the ULCC can be up to 550,00 dwt, 308 meters in length and 40 meters in breadth. While most oil tankers discharge with their own pumps, the large vessels require a complete dedicated port infrastructure, both in exporting areas and in the importing countries. Their deep draught also places restrictions on the use of key shipping lanes (Gilje, Dinwoodie, & Challacombe, 2002). The Westridge Marine Terminal in the Port of Vancouver can only accommodate the smaller Aframax tankers up to 100,000 dwt and barges for crude oil due to navigational restrictions arising from the terminals location deep inside Burrard inlet. A new bulk liquid terminal at either Kitimat or Prince Rupert ports would face no such restrictions.

3. B.C. Offshore Oil and Gas

3.1 Midstream Opportunities

According to the Geological Survey of Canada, the Queen Charlotte Basin could contain up to 734 billion cubic meters of natural gas and 1.56 billion cubic metres of crude oil. The Winona, Tofino, and Georgia basins are areas of sedimentary rocks off the B.C. coast and have a smaller resource potential, primarily natural gas. The entire Pacific Margin region is estimated to contain about 4 percent of Canada's total conventional hydrocarbon resources (Canadian Centre for Energy Information 2004 p. 13). Yet, in 1972, Canada made a policy decision not to approve any new exploration permits or programs in the West Coast offshore area and to suspend all work obligations under existing permits. This decision is known as the federal moratorium. In the midst of negotiating the Pacific Accord in the late 1980s, oil spills in Washington and Alaska led the British Columbia government to announce that it would not allow offshore exploration for at least five years. This policy decision is known as the provincial moratorium.

The area under provincial moratorium includes Johnstone Strait, Georgia Strait, Juan de Fuca Strait, Hecate Strait, and Queen Charlotte Sound.

The election of a new B.C. provincial government in June 2001 brought a renewed interest to examine the potential for offshore oil and gas development. The Liberal Government's 2003 throne speech outlined the goal of seeing exploratory drilling underway by 2010. A number of environmental issues (species and habitats at risk, sea-floor stability, currents, tides, earthquake risk) need to be considered before this goal can be achieved. Thus, the B.C. government appointed an independent panel to examine whether offshore oil and gas could be developed in a scientifically sound and environmentally responsible manner. The Scientific Review Panel (2002 p. 51) panel concluded that there was "no inherent or fundamental inadequacy of the science or technology, properly applied in an appropriate regulatory framework, to justify retention of the B.C. moratorium" on offshore oil and gas activities."

During the last four years, the Government of Canada has also been reconsidering its offshore oil and gas policy options. For example, it appointed an independent panel to facilitate and coordinate a public and First Nations review process. The review panel noted that many participants in the public and First Nations process believed that significant socio-economic information gaps existed with respect to the potential impact offshore oil and gas development. The Report of the Public Review Panel on the Government of Canada Moratorium on Offshore Oil and Gas Activities in the Queen Charlotte Region of British Columbia (2004 p. p. 106) surprising concluded that the firmly held and polarized views they received did not provide a basis for reaching a policy compromise regarding keeping or lifting the moratorium. The final report was rendered after the 2004 federal election that resulted in a minority federal Liberal government.

The report's authors were no doubt aware that a minority government would probably prefer not to deal with the complex policy issues associated with offshore oil and gas development in a region where they hoped to garner future electoral success.

The application of the growth vector technique to Canada's west-coast ports reveals that the offshore oil and gas sector represents both market development and service diversification opportunities since there are no ports currently participating in this market segment. Jayawardana and Hochstein's (2004) research reveals that ports can become a centre of the midstream sector if they effectively connect the onshore logistics network with the offshore segment. The network requires liquid bulk cargo handling terminals, industrial sites for large shipyards, equipment fabrication and repair, and value-added processing activities. Given the nature of these activities, any port seeking to participate in a future B.C. offshore oil and gas industry would be well advised to assess how they could accommodate the offshore oil and gas industry's logistics requirements. Unless these logistics requirements are correctly identified at the outset, they will not be able to be considered in the preparation of any long-range strategic plans, land-use plans or infrastructure-upgrading decisions. However, port officials with this type of planning information will be in an ideal position to contribute to the public policy debate regarding the economic impacts associated with offshore oil and gas development.

3.2 Midstream Challenges

The five sequential phases in typical offshore oil and gas development are seismic and geophysical surveys; exploration; development; production, and decommissioning. An effective logistic system to support each of these phases consists of the offshore platforms, port infrastructure, and inland transportation networks.

Offshore oil and gas exploration and development may be made cost prohibitive if there are deficiencies in any of these three areas. This paper will now briefly examine typical logistics issues that may need to be incorporated into certain west coast ports' strategic plans.

The Scientific Review Panel (2002 p. 33) report stated that significant advances in drilling technology have enabled the industry to conduct offshore exploration and production in much deeper waters than when the federal moratorium was first introduced. Given the relatively deep waters off the B.C. coast, the Panel's observation that it was now technically feasible to contemplate the extension of offshore operations beyond the narrow continental shelf raised public awareness of the potential scope for exploration and development activity. This in turn increased the information needs of coastal communities and First Nations residents as they sought to make informed policy opinions.

In the event that an economic offshore oil deposit is identified, there are a limited number of offshore drilling and production platforms to consider. The type of offshore platform used depends on water depth, as well as the economic characteristics and production life of the resource. For example, drill rigs on piles can be used in water three meters in depth, submersible rigs in waters up to 20 meters, jackups in 100 meter water, anchored semi-submersible platforms in waters up to 200 meters, and dynamically positioned drill ships in water over 2,000 meters or more (Canadian Centre for Energy Information 2004).

The Scientific Review Panel (2002 p. 34) concluded that offshore oil and gas production options in B.C. include a moored semi-submersible, a tension leg platform, a floating production storage and offloading unit, or possibly a gravity-base structure. Floating facilities are less costly and are suited to a smaller resource with a shorter production life.

The Panel also concluded that it should be possible to require that the lines from the production platforms be tied into pipelines as opposed to offloading the oil into shuttle tankers by way of buoys. In contrast, the Terra Nova oil field uses shuttle tankers to carry the crude oil from Terra Nova directly to market or to the same transhipment terminal used by the Hibernia project.

If the offshore resource being developed is natural gas, there are a few basic logistics options to transport natural gas production to market. For example: (a) using an undersea gas pipeline; (b) cooling the natural gas until it liquefies (LNG), and then transporting it by ship; (c) compressing the natural gas (CNG), and then transporting it by ship.

Offshore platforms used to extract an oil and gas resource require a fleet of dedicated vessels to service their needs. For example, anchor handling tugs, cable laying vessels, supply vessels, construction vessels, dredgers, floating storage vessels, shuttle tankers, mobile drilling units, mobile production systems, remotely operated vessels, standby and rescue vessels, survey vessels, and tugs.

There are also significant differences at each stage of the onshore logistics process for an offshore port compared to a typical cargo-handling terminal at a conventional port. While conventional ports attempt to maximize throughput productivity, the functions at an offshore port mean that reliable and uninterrupted supplies of inventory to the offshore platforms are a major focus of activity. This requires a significant landside area for staging product, fabrication, and consolidation supplies. For example, warehousing infrastructure may be related to the need for safety inventories and thus may be in storage for longer periods. Transporting goods (personnel and cargo) offshore by either air or water is also an important element of the logistical system.

Common cargo include food and provisions, water for drinking and industrial use, drilling mud, lubricants and fluids, waste disposal, and turbines and other instrumentation (Jayawardana and Hochstein, 2004). The exact nature of the activities in an upland port area will depend on offshore requirements in exploration, development, and production. However, it is not hard to imagine the volume of rail and truck traffic required to support such activities. Port engineers and planners will no doubt note that the vessel berthing areas, repair and maintenance services, and landside services at the offshore oil and gas port are generally inconsistent with what would be found at a dry bulk, break-bulk, or container terminal. They would be difficult to incorporate in ports pressured by urban encroachment such as Vancouver Port Authority of the Fraser River Port Authority where there are significant traffic volumes and congestion outside the port's property limits. Ports on Vancouver Island would potentially incur more expensive inland transportation costs given their location, unless barge service could be established for transporting heavy equipment or supplies.

4. Matching West Coast Port Capacity With Oil and Gas Production: Some Conclusions

Increased production in the Alberta oil sands over the coming years offers the potential for a cargo and market extension for the Port of Vancouver, or a cargo and market diversification opportunity for Kitimat or Prince Rupert port, if the appropriate logistics networks are put in place at the right time.

Port strategists have a number of items to consider when evaluating the Terasen and Enbridge proposals. Terasen's focus on the California market suggests an Aframax-sized vessel might be deployed for the ocean going portion of the trip to market where the shipping advantage of using ULCC may not be advantageous.

This is important since the location of Terasen's Vancouver terminal prohibits a VLCC or ULCC from berthing because they require a deep draught anchorage, wide approach channels, and turning basins. It appears that the Terasen's short-term strategy is based on physical and operational constraints at their terminal, as those influencing pipeline economics.

There are more factors to consider before deciding whether increasing oil sands production represents a cargo and market diversification opportunity for either Kitimat or Prince Rupert port. A key item is the capital cost of a new pipeline corridor. Given B.C.'s coastal geography and the fact that Kitimat is closer to the Edmonton hub, a crude oil terminal at Kitimat may result in significant pipeline cost savings compared to Prince Rupert. A northern port option represents an advantage compared to a Vancouver terminal location because of the economies of scale in ocean shipping. Locating a liquid bulk crude terminal at either port would make it accessible to the largest ocean going vessel. This would extend the potential market reach of Alberta oil sands producers. Given the fact that both crude pipeline proposals feature a northern port option, the long-run choice may not be a southern or northern port location. Rather, it may between the two northern competitors.

It is difficult to anticipate if any of B.C.'s offshore oil and gas resources will be developed. Nevertheless, port engineers and planners can readily appreciate the fact that deep-water oil and gas exploration and production generates a greater need for more extensive, and capital-intensive port infrastructure. The geographical location of B.C.'s offshore oil gas resources and water depths may indicate that developing one main offshore oil port would probably be the most economical option. Based on the logistic needs identified in this paper, the Vancouver Port Authority

and Fraser River Port Authority do not appear to be ideally positioned to assume this role. The logistics requirements seem to favour either Kitimat or Prince Rupert. Kitimat may be able to leverage its competitive advantage for a bulk liquid crude terminal for Alberta's oil sands exports to become the location for a future B.C. offshore oil and gas port.

References

Alberta Energy and Utilities Board (May 2004) 'Alberta's Reserves 2003 and Supply/Demand Outlook 2004-13' Statistical Series (ST) 2004-98.

The B.C. Government (2005) '*British Columbia Port Strategy*' prepared by: Colledge Transportation Consulting Inc., Seaport Consultants Canada Inc., Seabulk Systems Inc., COMFAC Services Ltd. 31 pp.

CAPP (2005) 'Crude Oil Pipeline Expansion Summary', Canadian Association of Petroleum Producers. 17 pp.

Canadian Centre for Energy Information (March 2004) *Canada's Evolving Offshore Oil and Gas Industry*, Canadian Centre for Energy Information. 41 pp.

Dunbar Bob (2004) 'Oil Sands Supply Outlook', *CERI Energy Insight*, Canadian Energy Research Institute, Issue No. 2, October 2004.

Gilje, T., Dinwoodie, J., & Challacombe, J. (2002) 'Crude Carrier Consolidation and Capital Cost', *International Journal of Maritime Economics*, vol. 4, #1. pp. 35-45.

Ircha, Michael (2000). 'Port strategic planning: Canadian port reform'.

Jayawardana, J. & Hochstein, A. (2004) 'Supply Network for Deepwater Oil and Gas Development in the Gulf of Mexico: An Empirical Analysis of Demand for Port Services; Final Report', OCS Study MMS 2004-047. 98 pp.

Sletmo, Gunnar (1999). 'Port life cycles: policy and strategy in the global economy', *International Journal of Maritime Economics*, Rotterdam. Selected pages.

Stopford, M. 1995, *Maritime Economics*, Routledge, London.

Report of the Scientific Review Panel (2002)– British Columbia Hydrocarbon Development Chapter Four – Engineering and Technology.
Access October 1, 2004 http://www.offshoreoilandgas.gov.ab.ca

CANADA-US PORT COMPETITION AND CARGO DIVERSION

Joseph Monteiro*

Introduction

Empirical studies have shown that there is a positive relationship between growth and trade. Trade liberalization reduces the effective rate of protection provided by tariffs. Tariff costs are gradually becoming less than the costs of moving goods from one destination to another thereby increasing the competitiveness of goods. It is therefore not surprising that all aspects of transport costs have come under increasing scrutiny, including port costs. Port costs and port competitiveness have recently attracted considerable attention as allegations have been made that the playing field between Canadian and US ports are not even. The matter of port competition is also of interest as it can affect the movement of cargo via Canadian or US ports.

This focus of this paper is to examine whether ports in North America compete and to emphasize an analytical approach in reviewing this matter. The methodology used will follow the jurisprudence established in Canadian antitrust cases. This paper will also examine the statistics on cargo diversion, one partial though imperfect indicator whether ports compete. However, first the statistics on traffic at North American ports will be reviewed to broadly isolate those ports which are likely to compete. Finally, a few concluding remarks are made.

I. A Brief Review of Ocean Traffic at North American Ports

Containerized cargo statistics arriving at North American Ports by ocean in 2001 are shown in Table 1. These statistics reveal that the major (i.e., more than 7 m. tonnes) trade lanes are the West Coast-Asia & Oceania, the East Coast-Europe and the South Coast-Cent. & S. America, each being nearly twice as important as the next in

* *The views expressed are those of the author and are not purported to be those of the Commissioner or the Competition Bureau.* *The author acknowledges the data provided by Ruth Abbott and Paul Kochhar.*

tonnage. The statistics also reveal that the ten major ports in North America are Long Beach, LA, NY/NJ, Vancouver, Charleston, Montreal, Oakland, Houston, Norfolk and Savannah. The first three of these ports are nearly double the size of the other ports in total tonnage.

The total tonnage by itself is often a misleading indicator of competition as ports on the east coast of North America rarely compete with ports on the west coast. As a consequence, if one examines the East Coast-Europe trade lane, the ranking changes (see highlight in column 1), placing Montreal and NY/NJ as the major ports. Similarly, for the West Coast-Asia & Oceania trade lane, the rankings of the major ports are Long Beach, Los Angeles, and Vancouver (see highlight in column 2). As can be expected, the important ports to the Cent. & S. America-US trade lane are Miami, NY/NJ and Port

Table 1- Containerized Cargo Handled at North American Ports in the Ocean Trade (2001)

Port (Rank)	East Coast-Europe	W. Coast Asia & Oceania	Middle East & Africa Coast	Cent. & S. Coast America	Total
	M. Tonnes (%)	M. Tonnes (%)	M. Tonnes (%)	M. Tonnes (%)	M. Tonnes (%)
NY/NJ (3)	**7.91 (19.3)**	7.01 (7.5)	**1.55 (23.1)**	**2.90 (11.5)**	**19.38 (11.7)**
Montreal, PQ (6)	**8.02 (19.6)**	0.05 (0.1)	0.29 (4.3)	0.04 (0.1)	**8.40 (5.1)**
Charleston, SC (5)	4.69 (11.5)	2.68 (2.9)	0.86 (12.8)	1.66 (6.5)	**9.89 (6.0)**
Houston, TX (8)	4.10 (10.0)	0.51 (0.5)	1.00 (14.9)	2.17 (8.5)	**7.78 (4.7)**
Norfolk, VA (9)	3.42 (8.3)	2.55 (2.7)	0.62 (9.2)	0.90 (3.6)	**7.49 (4.5)**
Savannah, GA (10)	1.49 (3.6)	4.39 (4.7)	0.55 (8.2)	0.53 (2.1)	6.96 (4.2)
Miami, FL. (12)	1.20 (2.9)	0.74 (0.8)	0.13 (1.9)	**3.83 (15.1)**	5.91 (3.6)
Halifax, NS (14)	2.18 (5.3)	0.95 (1.2)	0.38 (5.6)	0.40 (1.5)	3.91 (2.4)
Port Everglades (15)	0.14 (0.3)	0.02 (0.0)	0.00 (0.0)	**2.89 (11.4)**	3.05 (1.8)
Long Beach, CA (1)	0.81 (2.5)	**22.19 (24.0)**	0.09 (1.3)	0.67 (2.6)	**23.76 (14.3)**
Los Angeles, CA (2)	1.16 (2.8)	**19.92 (21.5)**	0.29 (4.3)	0.69 (2.7)	**22.06(13.3)**
Vancouver, BC (4)	0.13 (0.3)	**9.57 (10.3)**	0.08 (1.2)	0.30 (1.2)	**10.08 (6.1)**
Oakland, CA (7)	1.12 (2.7)	6.91 (7.5)	0.10 (1.5)	0.16 (0.6)	**8.29 (5.0)**
Seattle, WA (11)	0.17 (0.4)	6.23 (6.7)	0.04 (0.5)	0.10 (0.4)	6.53 (3.9)
Tacoma, WA (13)	0.01 (0.0)	4.68 (5.0)	0.10 (1.5)	0.00 (0.0)	4.79 (2.9)
Other USA	4.28 (10.5)	3.87 (4.2)	0.63 (9.3)	7.65 (30.3)	16.43 (9.9)
Other Canada	0.06 (0.1)	0.34 (0.4)	0.01 (0.1)	0.39 (1.5)	0.80 (0.1)
TOTAL	40.89 (100.00)	92.61 (100.00)	6.72 (100.00)	25.28 (100.00)	(165.50)

Source: Statistics Canada, "The Future for Canada-US Container Port Rivalries," Revised D. O'Keefe, 54F0001XIE.

Everglades; and the most important port in the East Coast-Middle East & Africa trade lane is NY/NJ.

II. Do Ports in North America Compete?

Ports compete among themselves, if the service and related services provided by the ports, the quality of these services, the price of these services and the markets that these ports intend to serve are comparable. In other words, if the services that ports offer are substitutable they can be considered to be competitive.

Substitutable services or products for antitrust purposes have been defined by Canadian Courts. For instance in *Canada (Director of Investigation and Research) v. Southam Inc.* the court held that *direct evidence of substitutability* includes both statistical evidence of buyer price sensitivity and anecdotal evidence, emanating from buyers on past or hypothetical responses to price changes. *Indirect evidence of substitutability* may be garnered from certain practical indicia, such as functional interchangeability and industry views/behaviour. Since direct evidence of substitutability is generally not readily available one generally relies on functional interchangeability and industry views. Even in the face of evidence disclosing low-demand elasticity (i.e., direct evidence that products are not substitutable), product substitutability may nevertheless be shown to exist, based upon indirect evidence such as functional interchangeability and industry views/behaviour.[1]

A. Direct evidence of substitutability

There are well known statistical tests for determining substitutability to determine whether ports compete. The two easiest tests are price similarity and price equality.

The price similarity tests proposed by Stigler and Sherwin holds that two products or geographic areas are in the same market when their relative prices maintain a stable ratio. Measures of price similarity are: price correlation (i.e., R^2), correlations of their first differences, etc. This test has limitations: it may not reflect economic forces or prices may suffer from inertia.

The price equality tests were proposed by Shrieves and Horowitz. Geographical areas could be grouped into the same market if price differences between the port services tend to be zero over time. The test is performed by regressing $[\log p_x - \log p_y]_t = a + b[\log p_x - \log p_y]_{t-1} + w_t$. If x and y are in the same market, one would expect the coefficient 'a' to be 0. This test may not be necessary or sufficient for either the services or areas to be or not in the relevant market.

Data were not readily available to perform the statistical tests. However, if the tests were performed a high R^2 would mean that the ports should be considered to be in the same geographic market. Similarly, if the price equality tests were performed and the coefficient 'a' was zero then the ports are in the same geographic market.

It should be remembered that the demand for port services or transport services is a derived demand. It is worthwhile noting that the lower the elasticity of demand for the final transported product the lower the elasticity of demand for transport or port services. Further, the lower the proportion of port costs or transport cost to total cost the lower is the elasticity of demand for transport.[2]

B. Indirect evidence of substitutability

Indirect evidence of substitutability that is provided by: 1) functional interchangeability; and 2) industry views will be examined.

1) Functional interchangeability

Functional interchangeability of service in the case of ports depends upon: *natural endowment factors* (i.e., navigating and landside access and space); *infrastructure factors* (i.e., docks, cranes, intermodal connections); *human factors* (i.e., labour harmony, productivity, and legislative environment); and *locational factors* (i.e., closeness to markets and markets to be served).[3]

Before considering functional interchangeability, one would have to identify the markets these ports are likely to serve. An examination of the traffic patterns at North American ports and cargo diversion indicates that there are two basic markets that are of particular interest to Canadian ports: a) the Europe - US North Central Region; and b) the

Asia & Oceania - US West Coast. In the former market, the two important ports that compete are Montreal and NY/NJ. In the latter market, the three important ports that compete are Vancouver, Seattle and Tacoma. The above observations should not imply that there are no fringe competitors. For example, Halifax from the North may compete with Montreal; and Charleston, Houston and Norfolk may provide competition to NY/NJ from the South.

a) Europe - US North Central Region

The Europe - US North Central Region (Wisconsin, Michigan, Illinois, Indiana and Ohio) market is served by several east coast ports. To determine whether Halifax, Montreal and NY/NJ compete the factors that determine functional interchangeability of service at these ports indicated above will be examined.

Halifax: Halifax is a natural deep sea water port whose natural endowments enable it to accommodate large container ships. The port has channel depths of 60 feet making it the deepest Atlantic sea port giving it the potential for handling container ships of 6000 TEU. Baltimore and Norfolk are the only two other North American east coast ports that have 50 foot channel depth. The infrastructure at the Halifax port enables it to handle containers and it is served by CN providing it with rail intermodal connection to the US Midwest through the opening of the Sarnia tunnel in 1995. The port has purchased cranes capable of handling super-post Panamax ships. It has a significant locational advantage being the closest North American port to Northern Europe (eg. Antwerp, Amsterdam, etc.) with an ocean-distance advantage of 1150 kilometres.

Montreal: The natural endowments of the Port of Montreal place certain restrictions on it though it handles the largest volume of containers in the Europe-US East coast trade. It is better suited to handle vessels of less than 2000 TEUs which contrasts with the trend of emerging large container ships. Further, the winter season results in navigation being closed to this port for a few months of the year. The infrastructure at the port is excellent since it is served by two major railways CN and CP. The dockside rail and the improvements in

intermodal access have enabled it to reduce rail times from Montreal to Toronto, Detroit and Chicago by 10, 25 and 30 hours, respectively. These advantages have been further enhanced by the double stack service, the widening of the tunnels and the opening of the Sarnia tunnel. Being 1600 kilometres inland, its locational advantage provide it with the closest rail access to the US Midwest.

NY/NJ: The natural endowments of the Port of NY/NJ have enabled it to become the largest port on the North American east coast, though it ranks second in the Europe-US East Coast trade. It has depth limitation (45 feet) which is not suited to the trend towards large containers of more than 4000 TEUs. The infrastructure at the port is excellent since it is served by two major railways CSX and NS. The human factor such as the US harbour maintenance tax on imports has in the past had the effect on routing via Canada. The ability of the port to obtain financing at more favourable terms due to its ability to raise tax-free and low-interest revenue bonds provide it with an advantage over Canadian ports.

In sum, with regard to the Europe - US North Central Region market, the service provided by the ports of Halifax, Montreal and NY/NJ are functionally interchangeable for ships of less than 2000 TEUs (excluding winter shipping for Montreal). For ships of more than 2000 TEUs, the service of Halifax and NY/NJ are functionally interchangeable. Canadian ports have locational advantages over US ports in the movement of traffic to the US Midwest. Its infrastructure advantage relating to intermodal service has improved considerably due to developments such as widening of tunnels, double stack service, mergers or alliances by Canadian railways. This should provide it with a competitive advantage. However, its natural and human endowments which provided them a competitive advantage in the past could be affected in the future. This is because the natural endowment (depth) disadvantage that NY/NJ has had is changing due to its dredging program; and the human endowment disadvantage could be affected by the 1998 amendments to the *US Ocean Shipping Reform Act* which enable shipping lines to offer individual confidential service contracts. On the other hand, NY/NJ will continue to have the advantage of a

larger market it serves. Further, its ability to raise capital at more favourable terms could provide this port with an additional competitive advantage.

b) Asia & Oceania - US West Region

The Asia & Oceania - US West Region (Seattle, Washington and California) market is served by several west coast ports. The market share of the west coast ports in the Asia & Oceania -US west Coast region is shown in table 1. The data reveal that the two west coast ports that are likely to be potential competitors with Vancouver are: Seattle, and Tacoma. Functional interchangeability of service at these ports will be further examined.

Vancouver: The Port of Vancouver has natural deep channel entrances with berth depths of 50 feet or more. It has three terminals (Vanterm (76 acres), Centerm (72 acres), Deltaport (100 acres) the former two in Burrard Inlet and the latter at Roberts Bank) whose capacity recently increased which previously not only acted as a constraint to its growth but also resulted in diversion of Canadian cargo to the US. The infrastructure at the port is served by three major railways CP, CN and BNSF. BNSF has rail links between Deltaport to the US Southwest, Northwest and Midwest. BC Rail also serves this port. However, since the port is designed for rail intermodal shipments, the trucking industry that has developed there is inadequate to handle intermodal service which is especially needed for shorter distances such as Seattle/Tacoma over which trucking has economies compared to rail. The Canadian railways can provide cross-continental services without interlining with eastern railroads. The three terminals have 13 dock-side gantry cranes (with 2 on order) and 7 container berths. The Deltaport and Vanterm ports have large rail-mounted gantry cranes capable of handling super-post Panamax ships. It averages 25-26 container moves per hour. Vancouver's container trade has increased from 616, 692 containers (20 TEUs) in 1996 to 1, 163, 178 containers in 2000 and their annual container capacity is more than 1.2m TEUs. The port has the natural advantage of being the nearest port of call for inbound traffic from Asia. It is 1,092 km closer to Hong Kong than the

port of Long Beach in California and is nearer than Seattle or Tacoma.

Seattle: The Port of Seattle has natural deep channel entrances and berth depths of 50 feet that can accommodate the super-post Panamax container ships. The infrastructure at the port is excellent being served by two major railways BNSF and UP. The trucking industry that the port has is far superior to that in Vancouver. The human factors such as the labour relation problems and the US harbour maintenance tax on imports place it at a disadvantage to the Port of Vancouver. However, the taxing authority of the port to raise money places it in a favourable position compared to the Port of Vancouver. The port is about 150 miles south of the Port of Vancouver.

Tacoma: The Port of Tacoma has natural deep channel entrances and berth depths of 50 feet that can accommodate the super-post Panamax container ships. The infrastructure at the port is served by two major railways BNSF and UP. The human factors are both favourable and unfavourable. On the one hand, it has taxing authority similar to the Port of Seattle and on the other it is faced with a harbour maintenance tax on imports. The port is about 180 miles south of the Port of Vancouver.

In sum, with regard to the Asia & Oceania - US West Region market, the service provided by the ports of Vancouver, Seattle and Tacoma are functionally interchangeable. Vancouver has a slight advantage as far as the locational factor is concerned being about 150 to 180 miles nearer to Asia and Oceania. As far as the infrastructure factor is concerned all three ports have docks, cranes and railway intermodal connections that make them comparable. Vancouver may have a marginally better rail intermodal system but the trucking industry serving it is inferior compared to those at the other two US ports. Until recently, the Port of Vancouver was seriously handicapped due to the lack of terminal capacity but this appears to have changed due to the increase in capacity of Deltaport. The human factors gives the Port of Vancouver a slight advantage as far as labour is concerned. However, with regard to taxes the US ports have the disadvantage of having the harbour maintenance tax on imports but have the advantage of being able to raise finance because of its taxing authority.

2) Industry and Government Agency Views

Industry views on whether ports are competitive are perhaps best indicated in various sources such as: the Transportation Act Review Panel, the Report on Competitiveness, the Canada Marine Review Panel and in articles published in the newspapers or trade journals. These views will be briefly summarized.

Industry Views

i. *Canada Transportation Act Review Panel (CA Panel)*

Interveners' concern about the future competitiveness to the Panel on the marine industry centre on: 1. the need for adequate funding for marine and port infrastructure and fiscal regulatory requirements; 2. the governance requirements on newly formed Canadian Port Authorities which could place Canadian ports at a certain disadvantage; 3. the increase in competition faced by Canada's container ports from US ports and the need to upgrade facilities to meet it; 4. the US announced government funding of marine infrastructure. All these disadvantages faced by Canadian ports exacerbate an already unlevel playing field.

The Panel states "On the issue of transborder competition, although it is true that Vancouver competes with Seattle and Tacoma for container cargoes and Halifax and Montreal compete with New York, Baltimore and Philadelphia, it is probably too early to determine whether the *Canada Marine Act* has had a negative effect. On the evidence to date, Montreal, Halifax, and Vancouver have enjoyed strong growth in both containerized and non-containerized cargoes. More to the point, Vancouver has succeeded consistently in recapturing Canadian container cargo from Puget Sound ports in Washington, while Montreal continues to enjoy a large volume of US traffic."[4]

"The Panel believes that US government expenditures on marine infrastructure do, in the longer term, represent a competitive threat to Canada's largest ports. About half of all US ports receive some form of funding assistance, while others have the ability to levy a municipal tax on local citizens to fund port improvements. Still others can use revenues generated from non-port activities. The Panel suggests that the government should continually scrutinize the performance of

Canadian ports relative to US competitors and be prepared to take policy action if US government funding seriously distorts competitive traffic patterns."[5]

ii. *Report on Competitiveness of Canadian Ports*[6]

This report considers the Canadian port system's competitiveness with that of the US by making comparisons in the following areas: governance; recent and anticipated capital expenditures; and funding, taxation and public charges. First, Canadian ports are regarded as sources of revenue for the federal government or municipality, and since payments to them are a first claim on their revenues, it increases their risks and reduces access to funding from investors. Second, capital investment on ports lags in comparison with investments in the US. Third, centralization of governance structure in Canada leads to uncertainty and delays in comparison to the US. Fourth, funding of port security in the US has generated greater response than in Canada.

All the above factors have an impact on the competitiveness of Canadian ports which is summarized as follows in the report. "The expansion taking place in the U.S. ports, supported by public funding, has to be regarded as a serious competitive issue for the Canadian port system and its users. ... In the long run, this could lead to erosion of traffic from Canadian ports, particularly containerized cargo (where most of the U.S. port investment is being dedicated). Ultimately this will cause spiralling downward effects on revenue and ultimate loss of competitive position for Canadian ports."[7]

iii. *Canada Marine Act Review Panel (CMA Panel)*

The Panel Report and interveners expressed their belief that Canadian and US Port compete. Evidence of this is found in the following statements. "...Canadian marine transportation routes and infrastructure, including the St. Lawrence Seaway and Canadian ports, compete fiercely with the U.S. marine routes and port facilities."[8] "Some presentations stated, and the Panel believes, that Canada is losing hard-won international traffic to competing U.S. ports."[9] "They [ports] were particularly concerned about issues of financial viability

... and competitiveness with U.S. ports."[10]

iv. *Trade Journals/Newspapers*

The Canadian Gateway - It all comes down to cost. This paper is on recent cargo transhipments. It states "statistics indicate that cross-border transshipments are now an established fact. 'It all comes down to cost,' ... 'If you have access to an alternate routing that saves you bucks, your cargo is going to move that way. It's a matter of cost – cost, cost, cost'."[11]

Northern Competitor - Canada's Port of Vancouver targets US cargo.[12] This paper briefly examines recent developments at the Port of Vancouver and its competitiveness with US west coast ports. It points to the following factors that favour imports through Vancouver: time saving being the first port of call, weak Canadian dollar, aggressive pricing by Canadian railroads and their intermodal service, and the US harbour maintenance tax. Two factors against it are: the inadequate trucking industry and the lack of port capacity (more so in the past).

US Exports and Imports Transhipped Via Canada and Mexico 2000.[13] MARAD has provided statistical data for 2000 on US transshipments (exports and imports) via Canada in terms of tons, value, commodity, trade area and by US Custom District. Europe and the Mediterranean remains the leading origin or destination of US transhipments (79% in terms of tons or value) via Canada.

Government Agency Views

i. *North American Trade Diversion: A New Perspective*

This article examines: the factors that have affected Eastern Canadian ports; the reasons behind cargo movement; and the economic factors behind cargo diversion and intermodalism. Intermodalism has had a significant impact on the traditional definition of hinterland captive market, trade diversion and distribution systems. The study concludes by indicating that "trends in the development of international intermodal transport systems, ... will clearly have a major direct impact on the movements of Canadian and American trade as well as the

viability of our ports."[14]

ii. *Canadian International Marine Cargo Flows: A Profile of Transhipments Through US Ports, 1989-1991*

The purpose of this article is to assess the quality of PIERS data and provide a profile of Canadian transhipment information through US ports. While noting the shortcomings of PIERS data it indicates that it should play an important role in "Canada's national container trade perpetually challenged by highly competitive American ports."[15] It indicates that between 1989-1991, NY/NJ accounted for 75% of all Canadian TEUs exported and imported through US east coast ports; and Tacoma and Seattle accounted for 77% of all Canadian TEUs exported and imported through US west coast ports (with Tacoma nearly double that of Seattle).

iii. *Canada's East Coast Container Ports. Do they Compete With or Complement One Another In the Race For North American Container Traffic?*

This article addresses the issue of competition of Canadian vs US ports. It examines the North American container market, compares Canadian and US east coast ports, and then focuses on Montreal vs. Halifax. It provides an in depth comparison of the advantages and disadvantages of the above ports. It indicates that international shipping is highly competitive and that ports competing for this container traffic stand to gain or lose. Their central hypothesis is that the Port of Montreal will continue to be a niche port in the ocean container market despite trends towards scale economies. It states "...while there appears to be some competition for containerized cargo between these ports [Halifax and Montreal], the competition appears to be directed more at U.S. Atlantic coast ports. The large size of U.S. markets makes them very attractive in an era of free trade and Canadian ports are well suited to compete within these markets."[16]

iv. *The Future for Canada-US Container Port Rivalries*

This article examines the trends in container shipping, container

shipping in North America, competition with US ports and the future for Canadian ports. It concludes by stating "Canadian container ports have successfully competed with US container ports in the past for North American container traffic. The keys to their success have been a combination of natural endowments such as superior locations and deepwater ports, investments in intermodal facilities including good rail access to all US markets and competitive pricing due to a favourable currency exchange. These factors are likely to continue into the future, however, the competition among container ports is likely to intensify in the future as industry consolidation continues and publicly funded US intermodal terminal and corridor projects come to fruition."[17]

In sum, industry and government agency views suggest that certain ports compete for traffic to certain geographic markets. The concept of geographic markets over time has changed with the development of containerism and intermodalism. In the 1960s and early 1970s, geographic markets closely approximated the area surrounding the ports. Today, markets such as Chicago which can be served by ports on the East, West and Gulf coasts can all be considered competitive, though in reality much depends on the trade lane. Part of the testimony that ports compete lies in the concerns and interest expressed in cargo diversion. This becomes evident if any of the ports are provided a competitive advantage and industry reaction is usually swift to indicate the changes and the potential for loss of traffic.

III. Cargo Diversion

If all other factors discussed under functional interchangeability are equal, other than port costs, then one could attribute any cargo diversion to the latter. Cargo diversion is a term used to describe a country's overseas trade through foreign ports. Diversion is the process whereby carriers serving country A solicit cargo from shippers in country B, then transport it overland from country B to A for final shipment overseas to country C. In the context used here, it simply means cargo destined to the US from Europe or elsewhere that uses Canadian ports for transhipment to the US, or vice versa.

Canadian Container Traffic Transshipped Through US Ports: Canadian container traffic transhipped through US ports from 1990-2000 is shown in Table 2. Over this period, Canadian exports via US ports have increased from 1,241 to 1,799, and Canadian imports via US ports have increased from 1,076 to 2,470. However, as a percent of total exports, these transhipped exports have decreased by 7%; and as a percent of total imports, these imports have increased by 7%.

Table 2 - Container Traffic Trans-shipped via Canada and US (000 tonnes)

Year	Exports* (Canada)	Exports# (US)	Total Export Imbalance	Imports* (Canada)	Imports# (US)	Total Import Imbalance	Total Imbalance
1990	1, 241	2, 284	1, 043	1, 076	1, 771	695	1, 738
1991	1, 220	2, 274	1, 054	1, 124	1, 667	543	1, 597
1992	1, 214	1, 853	639	1, 059	1, 821	762	1, 401
1993	1, 782	2, 091	309	1, 611	2, 212	601	910
1994	2, 006	2, 381	375	2, 112	2, 841	729	1, 104
1995	2, 382	2, 514	132	1, 970	2, 737	767	899
1996	1, 998	2, 683	685	1, 862	3, 067	1, 205	1, 890
1997	1, 990	2, 633	643	2, 037	3, 521	1, 484	2, 127
1998	2, 125	2, 489	364	2, 276	4, 012	1, 736	2, 100
1999	2, 038	2, 463	425	2, 321	4, 397	2, 076	2, 501
2000	1, 799	2, 358	559	2, 470	4, 666	2, 196	2, 755

*Canada's Container Traffic Trans-shipped Through US Ports #US Container Traffic Trans-shipped Through Canada's Ports

US Container Traffic Trans-shipped Through Canadian Ports : US container traffic transhipped through Canadian ports from 1990-2000 is also shown in Table 2. Over this period, US exports via Canadian ports have increased marginally from 2,284 to 2,358, and US imports via Canadian ports have increased from 1,771 to 4,666.

Balance of Container Traffic Trans-shipped: Export imbalance (i.e., Canada's container export traffic trans-shipped through US ports minus US container export traffic trans-shipped through Canada's ports) over the period 1990 to 2000 has decreased from 1,043 to 559 (46%) and import imbalance (similarly defined) has increased from 695 to 2,196 (216%). The total imbalance of exports and imports has risen from 1,738 to 2, 755 (58%) in favour of Canada. Notwithstanding this favourable imbalance in absolute terms, in contrast, Canada's

transhipments via US have increased marginally from 36.35% to 37.80% and US transhipments via Canada have decreased from 63.65% to 62.20%.

Diversion Pattern. West Coast Transhipments In 2000, Tacoma and Seattle were by far the most dominant ports on the US west coast in terms of Canadian transhipment, accounting for **72.3%** of all west coast Canadian TEUs exported and imported. By imports Tacoma and Seattle accounted for over 72.5% of TEUs imported (i.e., 47.7% and 24.8%). By exports Tacoma and Seattle handled 71.9% of TEUs exported (i.e., 45.4% and 26.5%). The overall percentage has decreased by 5% since 1991. *East Coast Transhipments*: In 2000, NY/NJ and Philadelphia were by far the most dominant on the US east coast in terms of Canadian transhipment, accounting for **73.2%** of all east coast Canadian TEUs exported and imported. By imports NY/NJ and Philadelphia accounted for over 78.6% of TEUs imported (i.e., 67.7% and 10.9%). By exports NY/NJ and Philadelphia handled 68.2% of TEUs exported (i.e., 61.8% and 6.4%). The overall percentage has decreased by 10% for NY/NJ since 1991.

Is There a Rationale for Diversion? : In the literature on this subject, several reasons have been attributed for diversion of cargo. The three reasons most frequently mentioned are: 1) the availability of more efficient and integrated intermodal services; 2) the lower costs (resulting from natural endowments such as distance and transit time, conference services at non-conference prices to Canadian ports, labour costs, port charges, favourable exchange rates, and union regulations); and 3) the reduced service or lack of service resulting from lower volume of trade. Other reasons that have also been given are: 1) historical factors eg. investments by companies that have led to certain routing patterns; 2) service availability such as port terminal capacity, trucking services and highway development; 3) tax advantages and regulatory advantages; and 4) custom gateways and related services. Many of these factors responsible for cargo diversion determine the competitiveness of ports which are usually more important than port costs in determining specific routes.

IV. Conclusion

An attempt was made to examine whether ports in Canada and US compete. First, the major trade lanes - West Coast-Asia & Oceania, the East Coast-Europe and the South Coast-Cent. & S. America, - were examined to determine which ports are likely to compete.

Second, once it was determined which ports are likely to compete, the antitrust methodology established in Canadian cases was used to review the matter. That is ports compete among themselves if there is direct or indirect evidence that the services offered by them are substitutable. In the absence of direct evidence of substitutability (price factors such as cross-elasticity and anecdotal evidence on hypothetical price response), indirect evidence such as functional interchangeability and industry views are relied upon. The latter suggest that the ports of Montreal, Halifax and NY/NJ compete among themselves (excluding the winter season or ships of more than 2000 TEUs in the case of Montreal) for the Europe - US North Central Region; and the ports of Vancouver, Seattle and Tacoma compete among themselves for the Asia & Oceania - US Western Region. One factor that should be considered when examining the direct evidence of substitutability is the favourable exchange rate which unquestionably favours Canadian ports.

Third, cargo diversion was examined, one imperfect indicator of competition if all other factors are the same. From 1990-2000, Canadian exports via US ports have increased from 1.2m tonnes to 1.8m tonnes, and Canadian imports via US ports have increased from 1m tonnes to 2.5m tonnes. Over the same period, US exports via Canadian ports have increased from 2.3m tonnes to 2.4m tonnes, and US imports via Canadian ports have increased from 1.7m tonnes to 4.7m tonnes. The dominant ports in 2000, were Tacoma and Seattle on the US west coast and NY/NJ and Philadelphia on the US east coast in terms of Canadian transhipment. The three most important reasons attributed for diversion are: more efficient services; lower costs; and lack of services.

In conclusion, some ports may have a small advantage over others with whom they compete. The presence of cargo diversion that has occurred in the past with the advantage being held by Canadian ports on the East coast and by US ports on the West coast provides further

testimony in part that these ports compete.

Footnotes/Bibliography

1. *Canada (Director of Investigation and Research) v. Southam Inc.*, November 25, 1997.
2. Let $P_t = P_d - P_s$. Where $P_d = f(q)$ and $P_s = g(q)$. The elasticity of demand for transport of commodity 'i' is: $E_t = [aE_d (E_s/E_s - E_d)^{(1-a)}]$ where $a = P_t/P_d$. If E_s is very large it tends to infinity as the second bracketed term becomes unity then $E_t = aE_d = (P_t/P_d)E_d$. Since $(P_t/P_d) < 1$, it follows that $E_t < E_d$. Where P=price; t=transport; d=retail; s=supply; E=elasticity. So P_t means price of transport services and E_t means elasticity of demand for transport, etc. See *Competition and Regulation in the Railway Freight Industry*, CTC., WP 10-81-11, May 11, 1981, Appendix 2 and p. 2.
3. See characteristics discussed by O'Keefe, D, "Canada's East Coast Container Ports. Do they Compete With or Complement One Another In the Race For North American Container Traffic?", *CTRF Conference* 1998, 154-69.
4. *Vision and Balance*, Report of the Canada Transportation Act Review Panel, June 2001, p. 137.
5. Id, p. 137-8.
6. *Report on Competitiveness of Canadian Ports*, Association of CPAs, September 23, 2002.
7. Id, p. 43.
8. *The Canada Marine Act: Beyond Tomorrow*, TP14107B, June 2003, p. 19.
9. Id, p. 8.
10. Id, p. 29.
11. Bonney, Joseph, "The Canadian Gateway - It all comes down to cost", *JOC WEEK*, June 10-16, 2002, p. 19.
12. Mongelluzzo, Bill, "Northern Competitor - Canada's Port of Vancouver targets US cargo", *JOC WEEK*, October 29, November 4, 2001, pp. 30-31.
13. *US Exports and Imports Transhipped Via Canada and Mexico 2000* , US Department of Transportation, Maritime Administration.
14. Archambault, Michel, "North American Trade Diversion: A New Perspective", *CTRF Proceedings* 1985, p. 857.
15. Wardlow, Katie and David Dahm, "Canadian International Marine Cargo Flows: A Profile of Transhipments Through US Ports, 1989-1991", *CTRF Conference* 1993, p. 245.
16. See footnote 3, p. 168.
17. O'Keefe, Doug, *The Future for Canada-US Container Port Rivalries*, Statistics Canada, Cat. No. 54F0001XIE, p. 15.
18. Mongelluzzo, Bill, "Seattle, Tacoma Ports Face Challenges: Vancouver to Grab Business", *Journal of Commerce*, April 26, 1999, pp. 1-3.
19. "Is the labor dispute scaring away container shipments from Port of Vancouver?", *Canadian Transportation and Logistics*, August 26, 2003.
20. O'Keefe, Doug, *The Future for Canada-US Container Port Rivalries, Revised March 2003*, Statistics Canada, Cat. No. 54F0001XIE, ISBN: 0-662-34376-X, pp. 1-12.

The Effects of Railway Investments in a Polycentric City

Ghebreegziabiher Debrezion*
Eric Pels
Piet Rietveld

Free University, Department of Spatial Economics
De Boelelaan 1105, 1081 HV Amsterdam

March 2005†

Abstract
This paper analyses the effect of railway investment on land prices and land use in a polycentric city under various regulatory regimes of land markets. The introduction of a faster mode of transport (train), accessible in discrete locations leads to an extended city size. The stations of the "fast" mode attract dense residential settlements. As a result, the average residential and commercial land rents increase in both the competitive and segmented land market situations, as compared with the "slow" unimodal transport case. When rail investments only serve one particular centre, this leads to the growth of the advantaged centre at the expense of the other centre. Generally speaking, investment in the fast mode results in city growth and increase in rent receipts. However the effect of the investment for individual centres and their corresponding residential areas depends on the underlying land market assumptions and the level of investment. Distorted land markets lead to increases in commercial rents, but this is more than offset by the decrease in residential land rent.

Key words:
Bimodal urban model, bicentric city, commercial land rent, residential land rent, railway station, railway development.

---

* PhD student at the Free University of Amsterdam and Tinbergen Institute. E-mail gdebrezion@feweb.vu.nl.

† Prepared for Canadian Transportation Research Forum (CTRF) Hamilton, Canada.

1. INTRODUCTION

The car has gradually become the dominant transport mode in most cities of developed countries. However, there still are cities, such as London, Paris and New York, where a large part of the workers use public transport. Therefore, to make a proper analysis of land and labour markets in such cities, both transport modes should be considered. Many cities started with a clear monocentric structure. During the course of time, however, a gradual deconcentration process took place, leading to a decreasing dominance of the original centre. But, in some cases, edge cities have developed, implying the emergence of additional centres of commercial activity in a metropolitan area (Garreau, 1988). In other cases, the gradual growth of small- and medium-sized cities led to the evolution of large metropolitan areas consisting of overlapping urban areas that were formerly independent. In both these cases of city evolution, the original monocentric urban model no more applies. This paper sets out to study both phenomena in an urban model which deal with the combination of multiple transport modes and multiple centres of economic activity.

In the urban modelling, in relation to the type of centre and mode of transport assumptions, we can logically distinguish four families of models: monocentric-unimodal transport; monocentric-multimodal transport; polycentric-unimodal transport; and polycentric-multimodal transport. Most of the literature is in line with the first category, although recently, more and more studies that make the polycentric assumption have appeared. But these are still dominated by the unimodal transport assumption. We deal with an urban model of multiple centres and multiple transport modes, thus putting the paper in the last family. The additional mode of transport addressed here (railway) is more in line with public transport. It operates from discrete locations that are represented by stations. By developing an urban model, we aim at assessing the effects of an additional transport system (rail) on urban grow the in general, and the growth of particular sub-centres and competitive position of the sub centres.

2. LITERATURE REVIEW

The relation between land values and proximity of land to employment centres has been addressed extensively in the literature. The monocentric circular city has received the most attention. However, in many parts of the world, especially in Western Europe, Japan and the U.S., metropolitan areas are increasingly assuming polycentric structures. The Randstad Holland,

Rhine-Ruhr metropolitan area in Germany, the Flemish Diamond in Belgium, the Glasgow-Edinburgh region in UK, the Padua-Treviso-Venice region in Italy, Southern California in the U.S. and the Kansai area in Japan are probably the most frequently mentioned polycentric structures (van der Wusten and Faludi, 1992; Dieleman and Faludi, 1998; Batten, 1995; Musterd and van Zelm, 2001).

A common conjecture in the literature is that transportation is believed to be the prime factor in shaping the urban structure (Clark and Kuijpers-Linde, 1994). However, the equilibrium location of businesses and households has always been an outcome of the trade-off between production economies and transport. The transportation side favours dispersed business locations because of congestion, whereas the production side favours compact business locations due to agglomeration (Glaeser and Kahn, 2003). The trade-off between agglomeration effects and transportation costs explains the coexistence of multiple centres in a city. The increasing use of private cars is believed to be instrumental in shaping the present dispersed urban structure (Glaeser and Kahn, 2003). In addition to the use of cars, Sivitanidou (1997) showed that the recent information revolution is also contributing towards the weakening of spatial links between commercial activities and large business locations, thereby leading to increasingly dispersed business locations.

Several studies have addressed the effect of urban spatial structure on property values. The value of a centre is capitalized in the form of land rents. In addition to the predetermined centre in the urban models, other studies, without explicitly referring to the centre(s), have concluded that the rent gradient peaks around the most valuable location in the urban spatial structure (Anas and Kim 1996). Indirectly, these peaks are also used to identify the centre(s) of the city. However, in this sense, the monocentric assumption is in reality a very simplistic assumption. Therefore, over the years, attempts have been made to develop urban models in the context of polycentric situations (pre-specified and non-pre-specified locations).

A comprehensive general equilibrium polycentric urban model was developed by Anas and Kim (1996). Without scale economies of shopping, production is dispersed in the city with rent, wage, and commodity price and density gradients peaking in the centre of the space. One of the models on property values in a bicentric city was developed by Sivitanidou and Wheaton (1992). Special attention was given to the centres' production cost difference and commercial land market regulation. The main finding of the paper was that cost advantages are capitalized in commercial land rent and

wages (and wages in their turn, in residential land rents). The level of capitalization of production cost advantages in commercial land rent becomes higher in the regulated commercial land market compared with the competitive market.

3. MODEL DESCRIPTION: Bicentric-Bimodal urban structure

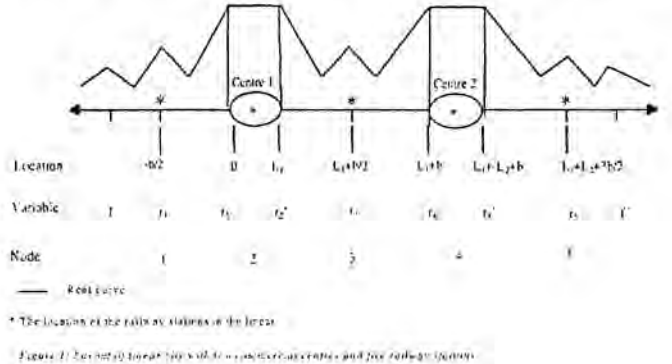
The city in our model has two area occupying centres, b miles apart from each centre's edge. They both function as employment and production centres. Labour is employed from households living in the residential areas of the city. We assume one member households who deliver identical labour. The inputs in the production process constitute labour and commercial floor space. In the production process of the centres, we assume a fixed ratio between labour and floor space. Floor space is prepared in a cost minimising fashion from commercial land rented at the commercial land rent rate and capital rented at some market rent of capital. The output follows a constant returns to scale technology, and is exported at a given unit price in a fully competitive market.

The households have a well-behaved utility function. Residential land and non-land consumption goods make up the utility-bearing bundle. The price of non-land consumption goods is taken as a numéraire (unity). By travelling to one of the employment centres, households acquire an endogenously given wage. The residential land rent follows a bid nature. All commercial and residential rents are absorbed by absentee landowners.

Two modes of transport operate in the open city: a "slow" mode (car) that is accessible from anywhere in the city, and a "fast" mode (train) accessible from certain fixed stations. The distinction between the "slow" and "fast" modes is related to generalised transportation cost structure. The fast mode results in some sort of cost saving, and is thus termed "fast". In our linear city model, in addition to stations at the two centres we assume three more stations at a distance of $b/2$ miles from the edges of the centres (see Figure 1 below). Transportation inside the centres is assumed to be zero.

We further assume that the city we deal here with is open: workers are free to migrate into or out of the city. The households enjoy the national utility level u which is bounded from below by the supreme utility level (a level of utility that guarantees the existence of the city). Thus, all households in the city enjoy a given utility level that is equivalent to the level of utility enjoyed by the households outside the city in the economy.

Centre 1 is assumed as a reference (located at 0) for all nodes. In Figure 1 below, the first row gives the variable representing the location of some critical points in the linear city: namely, the fringes of the city, the edges of the centres and the station locations. The left edge of centre 1 is taken as the origin of the linear city. L_1 and L_2 represent the size of the two centres, respectively. This is dealt with in detail in the next section.



The exogenous parameters of the model can be seen from three sides. On the consumer side: we have national utility level, price of non-land consumption goods and transportation costs of the modes. On the producer side: we have cost advantage of one centre over the other, floor space requirement per worker and cost of capital. Lastly, on the spatial structure of the city: we have width of the city and the distance between the nodes. The description of the exogenous variables used in the simulation is given in Table 1.

4. EQUILIBRIUM CONDITIONS

4.1 Households' equilibrium: utility maximization

Here we give the derivation of a general case residential bid rent function. It is defined by the maximum rent per unit land that a household, at a location r, which travels to a specified employment centre to get an income Y, can pay while achieving a certain utility level u. The bid rent function in the city is a function of the distance and the utility level enjoyed by the households (Fujita, 1989), which can be written as:

$$\Psi(r,u) = \max_{z,s}\left\{\frac{Y - T(r) - z}{s} \mid U(z,s) = u\right\} \quad , \tag{1}$$

Where $\Psi(r,u)$ is the residential land bid rent, for a household at location r enjoying a given exogenous level of utility u. $U(z,s)$ is the utility function, z is the composite consumption good of the household that has a unit price, and s is the land lot size per household. The household incurs transportation costs $T(r)$ which is a function of the location r in relation to the location of the employment centre, and receives Y level of income.

Equation (1) can be rewritten by expressing the amount of composite consumption goods of the household in terms of lot size of land and utility:

$$\Psi(r,u) = \max_{s}\left\{\frac{Y - T(r) - Z(s,u)}{s}\right\} \quad . \tag{2}$$

For a fixed utility level u, the FOC for maximizing the right hand side of equation (2) occurs at the point where the marginal change of the function with respect to s is zero. This leads to the relation:

$$-\frac{\partial Z(s,u)}{\partial s} = \frac{Y - T(r) - Z(s,u)}{s} \, . \tag{3}$$

At the optimal choice of s the right hand side of (3) equals the bid rent $\Psi(r,u)$:

$$-\frac{\partial Z(s,u)}{\partial s} = \Psi(r,u) \, . \tag{4}$$

For simplicity and ease of derivation, we assume that the utility function of the household has the following functional form:

$$U(z,s) = \alpha \ln z + \beta \ln s; \quad \alpha + \beta = 1 \, . \tag{5}$$

The equation of the indifference curve is given as $Z(s,u) = s^{-\beta/\alpha} e^{u/\alpha}$. Solving the maximization problem in (2) using the condition in (3), we have:

$$\Psi(r,u) = \alpha^{\alpha/\beta}\beta(Y - T(r))^{1/\beta} e^{-u/\beta}; \tag{6}$$

$$S(r,u) = \beta(Y - T(r)) / \Psi(r,u) = \alpha^{-\alpha/\beta}(Y - T(r))^{-\alpha/\beta} e^{u/\beta} \, . \tag{7}$$

Artificial income at nodes

In order to relate the above general bid rent analysis to our bimodal multi-nodal model, we introduce a special income variable. The multiple nodes situation in our case implies that these nodes can be described as centres. Some of these centres are real production and employment centres, and some are artificial centres with zero production. Thus, an income variable can be attached to each of these centres which we term "artificial income" (Y_j^*), defined by equation (7) below. The artificial income at any node is the wage at the employment centre of choice minus the transportation costs by the fast mode from the node to this centre. As there are five nodes (including the centre) in our model, Y_2^* ($\geq w_1$) and Y_4^* ($\geq w_2$) represent the artificial income at centre 1 and centre 2, respectively. The equality of these incomes with the real wages offered at the corresponding centres represents the coexistence of the centres as production and employment centres. However, the case when one of the artificial incomes is higher than the real wages offered at the centre implies that the centre ceases to be a production centre. Households find commuting to the dominant centre profitable, making it the sole production centre. The introduction of the artificial income here helps to reduce the whole analysis of the model to a one-transport case, by eliminating the second mode of transport at this level, and thus we have:

$$Y_j^* = \max_i (w_i - T_f(r_{ji})) \qquad j \in J \text{ and } i \in I, \tag{8}$$

Where $T_f(r_{ji})$ represents the transportation costs from node j to centre i by the fast mode. In the simple linear cost function $T_f(r_{ji}) = k_f \times \|r_i - r_j\|$. For the rest of our analysis, we adopt this linear transport cost function.

(a) Residential land rent, lot size and settlement density:

Ones the artificial income at nodes is determined the bid rent, max bid lot size, and density functions corresponding to each of the nodes are given by can be given by equations 9, 10 and 11 respectively.

$$\Psi_j(r,u) = \alpha^{\alpha/\beta}\beta(Y_j^* - k_s\|r_j - r\|)^{1/\beta}e^{-u/\beta}; \tag{9}$$

$$S_i(r,u) = \beta(Y_i^* - k_s\|r_i - r\|)/\Psi_j(r,u); \tag{10}$$

$$\rho_j(r,u) = 1/S_j(r,u). \tag{11}$$

The prevailing rent and density in the city are the maximum envelopes whereas the prevailing lot size function is the minimum envelope of the above variables corresponding to the nodes.

(b) Commercial land rent

In the production process of the firm(s) operating at the employment centre, land is one of the inputs in the production of floor space. Because transportation costs inside the centre are assumed to be zero, a uniform land rent for commercial use is obtained. The commercial land rent at the centres (RC_i) is determined by the nature of the land market. We introduce the two markets: competitive and segmented market in section 5.

(c) Prevailing land rent and land use

In the model we distinguish three types of land use: commercial; residential; and agriculture. Thus, the prevailing rent curve at any point in space is the maximum of the residential, commercial and agriculture land rents, which can be written as:

$$R(r) = \max_{i \in I, j \in J} (\Psi_j(r,u), RC_i, RA) \cdot \tag{12}$$

4.2 Producers' equilibrium: cost minimization

At the production side, the model incorporates the assumption of Sivitanidou and Wheaton (1992), in which the two centres make products that utilize labour and floor space and that are sold outside city. Floor space at the centres (Qs_i) is produced by making use of capital (K_i) and land (L_i) with constant returns to scale Cobb Douglas production function, as given below:

$$Qs_i = K^{\delta} L^{1-\delta} \tag{13}$$

Centres are assumed to have constant floor space demand per worker $a_i = \frac{Qs_i}{N_i}$. Thus, given the price of capital (r_c), rent for floor space is given by:

$$RF_i = \frac{RC_i L_i + r_c K_i}{a_i N_i}, \quad i \in I \cdot \tag{14}$$

The cost minimizing input bundle is given by:

$$K_i = a_i N_i \left(\frac{\delta RC_i}{(1-\delta) r_c}\right)^{(1-\delta)}; \quad L_i = a_i N_i \left(\frac{(1-\delta) r_c}{\delta RC_i}\right)^{\delta}, \quad i \in I \tag{15}$$

Therefore, the long-run cost-minimizing floor space rent is:

$$RF_i = r_c^{\delta} \delta^{-\delta} (1-\delta)^{(\delta-1)} RC_i^{(1-\delta)}, \quad i \in I. \tag{16}$$

The production equilibrium condition between the centres thus states that the costs per worker at each of the centres should be equal. The element of cost advantage here can be regarded as internal cost savings in the production process, or otherwise as external subsidy:

$$w_1 - w_2 = RF_2 a_s - RF_1 a_s + CA. \tag{17}$$

Centres operate at a zero profit. They are price takers for the goods or services supplied on the national market:

$$prt + CA - w_1 - RF_1 \times a_s = 0; \; prt - w_2 - RF_2 \times a_s = 0 \tag{18}$$

4.3 Indifference location

The number of households in the city (N) is an important element in the determination of the equilibrium conditions. It is assumed that the city does not continue beyond the fringes of the residential areas. The total number of households in the city is calculated as the integral of household density between the city fringes. The left (f^-) and right (f^+) fringes of the city are locations where the residential rent corresponding to the nearest node and the agricultural rent intersect, which can be written as:

$$\Psi_1(f^-, u) = RA = \Psi_S(f^+, u) \tag{19}$$

$$N = \int_{f^-}^{f^+} \rho(r, u)\, dr. \tag{20}$$

But, in order to determine the number of households commuting to each of the centres, we need another critical location, i.e. a point in the linear city where households are indifference between the centres. However the presence of a station can leave a range of house holds to its left and right indifferent opening the possibility of cross commuting. In such cases we assume the expected number of households in that range commuting to the centres will be distributed equally. This leaves the indifference point necessary for the analysis still effective. The rent curves corresponding to the

two centres and the middle station ($j = 2,3$ and 4) interact in determining the indifference location (f).

$$\begin{aligned} &\Psi_3(f,u) = \Psi_4(f,u) \quad \text{if} \quad CA > 0 \\ &f = r_3 \qquad\qquad\qquad\quad \text{if} \quad CA = 0 \\ &\Psi_3(f,u) = \Psi_2(f,u) \quad \text{if} \quad CA < 0. \end{aligned} \tag{21}$$

Given that (f) is the indifference location between the centres, the number of households working in each centre is given by:

$$N_1 = \int_{f^-}^{f} \rho(r,u)\,dr \; ; \quad N_2 = \int_{f}^{f^+} \rho(r,u)\,dr \tag{22}$$

5. ALTERNATIVE LAND MARKETS

5.1 Competitive land market:

In this model we assume households and firms freely bid against each other for land. At the edges of the centres the commercial and residential rents are equal. The competitive land bidding ensures that landlords will eventually extract the maximum saving that the consumers may enjoy, given the utility level. Because, by assumption, there are no transportation costs inside centres, the commercial rent curve assumes a uniform pattern. Centres situated at pre-specified locations make no profit from production processes. The equilibrium condition for the competitive market contains 13 equations in 13 unknowns.

5.2. A segmented land market

In this second model, we impose a binding restriction on the commercial land area for one or both of the centres, such that $L_i = \bar{L}_i < L_i^*$, where L_i^* is the land area occupied by centre i if no restriction is imposed on it. This restriction affects the commercial land rent, and the relative cost of land versus capital. Thus, at the centre with the commercial area restriction, the land rent is no longer determined by competitive bidding from residential land rent, but is instead based on the supply of commercial land rent. The equilibrium condition in this market situation has 11 equations in 11 unknowns.

6. NUMERICAL SOLUTION AND DISCUSSION

On the basis of the model above, we now simulate the results of three railway investment levels for the two land market conditions. With respect to the transport mode, we have looked at the unimodal case and, two bimodal cases: namely, partial and full bimodal. In the partial bimodal case, only centre 1 is served by the fast mode from two stations b/2 miles away from its edges, in addition to the slow mode. On the other hand in the full bimodal case where feasible, both centres are served by the faster mode from the pre-specified stations. In all cases the slow mode alternative is available

Table 1: Value of exogenous parameters

Parameters	Values
Distance between nodes (b/2 mile)	10
Width of the city (mile)	1.00
β in utility function	0.50
Agricultural rent ($/acre)	7500
Annual cost of transport	
- Slow mode k_s ($)	350
- Fast mode k_f ($)	200
δ in floor space function	0.77
Floor space per worker a_i (sq. ft.)	250
Annual rent of capital ($/sq. ft.) r_c	7.00
National utility level u	1.20
Price of non-land consumption	1.00
Price of production output (p)	1.00
Productivity (unit) per worker prt	22,371
Cost advantage (CA) for centre 1 ($)	0; 1,000
Commercial land restrictions	
- Centre 1 (sq. miles)	1.80
- Centre 2 (sq. miles)	1.80

As an extension of the model presented by Sivitanidou and Wheaton (1992), our model uses the same values for some of the exogenous parameters that

are used in their simulation. The remaining variables that relate to the new extension to the model are selected in a way that facilitates comparison. The values are given in Table 1 above.

Due to the space limitation in this session we could not include the graphical and tabular illustrations. It can be available from the authors at request.

6.1 Effect of land market distortions

The competitive land market makes it possible for the production centres to acquire the required amount of land input for their production process at the competitive land rent. As long as the wage-paying ability is not affected, the competitive land rent level remains the same. This leaves unchanged the relative cost of land and capital, and the capital to land ratio in the production technology of the centres. In contrast, in the segmented market, the level of land restriction (the supply of commercial land) imposed in the centre(s) affects the commercial land rent and residential land rent. The restriction has a direct effect on the relative cost of capital and land. This in turn affects the wage-paying ability of the centres. Therefore, segmentation of the market has an effect on the production technology of the centres. The higher the commercial land restriction imposed on the centres, the higher the relative price of land to capital. Centres will then be increasingly capital-intensive in their production process. This leads to higher commercial land rents in the centres. In real life, this fact is visible in the form of high-rise buildings in central areas of cities. On the other hand, the commercial land restriction reduces the wage-paying ability of the centres. Thus, at each location, residential land rents will be lower than the corresponding locations under the competitive land market situation.

The occurrence of restrictions on commercial land in the centres leads to an *increase* in total commercial land receipts *(.277 versus .273* in billions of Dollars*)*; the increase in rent per unit dominates the *decrease* in area. This seems to indicate that restrictions on commercial land use improve opportunities to use land rents as a source of finance for infrastructure. However, the *increase* in commercial land rents is more than off-set by a *decrease* in residential land rents *(3.758 versus 3.786* in billions of Dollars*)*.

6.2 Effect of transportation investment

The main focus of this paper is to determine the effect of investment in a fast mode on the land market. To trace the effect of the transportation investment,

we compare three cases: namely, 1) the base case (i.e. the unimodal case); 2) partial bimodal case, where only one centre is served by the fast mode from two stations; and 3) the complete bimodal case, where the two centres are served by the fast mode from three stations. Investment in the fast transport mode makes commuting to the centres possible from a wider range of locations. Hence, the city size is enlarged by claiming land from agricultural use outside the city. In addition, the fast mode stations attract denser residential settlement, which contributes to the population increase in the city. In general, the effect of the investment in fast transport has a different effect for the two land-market structures. In the competitive market, the investment does not affect the level of wage paying ability of the centres. However, the average residential land rents increase as a result of an increase in the density of settlement around the stations. On the other hand, because of the unchanged wage level, the commercial land rent is not affected. The capital to land ratio that represents the technology of production at the centres remains the same. However, the size of the centres increases due to the increase in the number of employees (jobs) arriving at the centres. In the segmented market case, some effects occur on both the residential and commercial sides. The additional labour made possible by the fast mode has the same direction of effect on the residential and the commercial land rents in the city (as was discussed earlier in this paragraphs). The important feature here is that, due to the increase in the demand of commercial land (as a result of the increase in labour), coupled with the limited supply of land, the commercial land rent increases. This makes the non-labour input into the production process costly. At the break-even level, the wage-paying ability of the centres declines. As a result, the average residential land rent declines compared with the case of the competitive market (1.34 versus 1.36 in $/sq ft).

Under a distorted land market, total commercial rent increase in the city, as a result of the introduction of rail, is higher than in a competitive market (.302-.273 versus .314-.277 in billions of Dollars). However, total residential rents decrease substantially due to segmentation, and thus the net effect on total rent receipts is clearly less favourable under distorted land markets than under competitive land markets. Hence, if capturing rents as a means to finance infrastructure is feasible at all locations, the competitive market offers the best opportunities. But, if these opportunities are only possible at commercial locations, the conclusion may change.

(a) Competitive position of centres

We can see the effect of railway investment on the competitive position of centres by allowing rail connection to one (centre 1 in this case) and controlling the other. In our linear city setting we call it partial rail investment. Our numerical analysis indicates that the advantaged centre grows at the expense of the disadvantaged centre by claiming more of the households residing in the area between the centres.

(b) Effects of mixed land market

We can also allow a mixed land market for the two centres in the city and see what effect this has. Generally, as expected, the outcome is in between the out comes of the two uniform land market situations. In the unimodal and complete bimodal transport cases, alternating the land market situation between centres results in perfect symmetry. However, the partial bimodal transport case has some special features. Higher land rent receipts are achieved when the centre served by the railway has a competitive land market.

7. CONCLUSION

From the discussion above, we can conclude that each area of change on the model has a unique effect on the model equilibrium result. Due to possible transport cost saving, railway stations serve as centres with an artificial wage and a corresponding residential land bid curve. Generally, investment in the fast mode results in: city growth both in terms of area size and population; increase in rent receipts; and denser residential settlements. However the effect of the investment for individual centres and their corresponding residential areas depends on the underlying land market assumptions and the level of investment.

When land rents are captured as a source of investment for railway infrastructure, the increase in total rents is highest in the situation of a competitive land market. But it is important to realise that the rent increases are spread widely in the urban metropolitan area, which may make them difficult to collect in real-world situations. Of course the most focussed increases of rents take place near the railway stations. Of special importance is the finding that, in the case of segmented markets, the total commercial rent receipts are higher than that of a competitive land market. Thus, as long as rent capturing is limited to commercial land use, the segmented land

market is not as unfavourable as one might expect. The issue of distortions of land markets is important because these distortions may have decisive impacts on long-run changes in transport demand in response to changes in the transport system. The present paper shows how such distortions can be addressed in an appropriate way.

In the present paper, we have analysed the impact of a second transport mode on the dynamics of centres in a metropolitan area, under the assumption that the additional infrastructure may reinforce or weaken the existing commercial centres. However, transport investments may also have far reaching effects on spatial structure, since they may stimulate the emergence of new centres. This theme of the formation of new centres has not been addressed in the present paper. Instead, we have focussed explicitly on the demand for commercial land and the implications of distortions for the land market. Analysing the possible emergence of additional centres falls outside the scope of this model, but is certainly a promising extension. In order to achieve this aim, the model should be developed in the direction of a more explicit treatment of production processes and agglomeration economies.

References:

Anas, Alex and Ikki Kim (1996). "General Equilibrium Models of Polycentric Urban Land Use with Endogenous Congestion and Job Agglomeration", *Journal of Urban Economics* ***40***, 232-256.

Arnott, Richard (2001). "Urban economic aggregates in monocentric and non-monocentric cities". Unpublished manuscript.

Batten, D. F. (1995). "Network Cities: creative urban agglomerations for the 21th century", *Urban Economics* ***32***, 313-327.

Clark, William A. V. and Marrianne Kuijpers-Linde (1994). "Commuting in Restructuring Urban Regions", *Urban Studies* ***31***, 465-483.

Dieleman, F.M. and A. Faludi (1998). "Randstad, Rhine-Ruhr and Flemish Diamond as one polynucleated macro-region", *Tijdschrift voor Economische en Sociale Geografie* ***89***, pp. 320-327.

Fujita, Masahisa (1989). "Urban Economic theory: Land Use and City Size". New York: Cambridge University press.

Fujita, Masahisa, Paul Krugman and Anthony J. Venables (1999). "The Spatial Economy: Cities, Regions and International Trade". London: the MIT Press .

Garreau, Joel (1988). Edge City: Life on the New Frontier. New York: Doubleday.

Giulliano, Genevieve and Kenneth A Small (1991). "Subcenters in the Los Angeles Region", *Regional Science and Urban Economics* ***21***, 163-182.

Glaeser, Edward L. and Matthew E. Kahn (2003). "Sprawl and Urban Growth". Handbook of Urban and Regional Economics Vol. 4.

Helsley, Robert W. and Arthur M. Sullivan (1991). "Urban Subcentre Formation", *Regional Science and Urban Economics* **21**, 255-275.

McDonald, John F. (1989). "Econometric Studies of Urban Population Density: A Survey", *Journal of Urban Economics* **26**, 361-385.

McMillen, Daniel P. (2001). "Polycentric Urban structure: The Case of Milwaukee", *Economic Perspectives* **25**, 15-27.

Modarres, Ali (2003). "Polycentricity and Transit Service", *Transportation Research Part A* **37**, 841-864.

Musterd, Sako and Ingrid van Zelm (2001). "Polycentricity, Household and the Identity of Places", *Urban Studies* **38**, 679-696.

Roback, Jennifer (1982). "Wage, Rent and the Quality of Life", *The Journal of Political Economy* **90**, 1257-1278.

Sasaki, Komei (1989). "Transportation System Change and Urban Structure in Two Transport Mode Setting", *Journal of Urban Economics* **25**, 346-367.

Sivitanidou, Rena (1996). "Do Office-Commercial Firms Value Access to Service Employment Centres? A Hedonic Value Analysis within Polycentric Los Angeles", *Journal of Urban Economics* **40**, 125-149.

Sivitanidou, Rena (1997). "Are Center Access Advantages Weakening? The Case of Office-Commercial Markets", *Journal of Urban Economics* **42**, 79-97.

Sivitanidou, Rena and William C. Wheaton (1992). "Wage and Rent Capitalization in the Commercial Real Estate Market", *Journal of Urban Economics* **31**, 206-229.

Wheaton, William C. (1977). "Residential Decentralization, Land Rents, and the Benefits of Urban Transportation Investment", *The American Economic Review* **67**, 138-143.

Wheaton, William C. (2002). "Commuting, Ricardian Rent and House Prices Appreciation in Cities with Dispersed Employment and Mixed Land Use", MIT Department of Economics Working Papers Series.

Wusten, H. van der and A. Faludi (1992). "The Randstad: playground for physical planners", in F. M. Dieleman and S. Musterd (Eds.) *The Randstad: A Research and policy Laboratory*, pp. 17-38. Dordrecht: Kluwer Academic Publishers.

AUSLINK: A NEW APPROACH TO INTERGOVERNMENTAL ARRANGEMENTS FOR LAND TRANSPORT FUNDING

Raluca Raicu, Derek Scrafton and Michael Taylor
Transport Systems Centre
University of South Australia

Introduction

The Australian federal government, through its 'AusLink' initiative, has developed a new approach towards national land transport policy. AusLink encourages a new emphasis on strategic, multi-modal transport planning at the national level, aimed at overcoming perceived weaknesses in the current intergovernmental funding and decision making processes. Amongst these weaknesses are: a focus on short term and *ad hoc* solutions, lack of coordination between transport modes and between the public and private sectors, and insufficient recognition of the potential of technological solutions.

A significant element of AusLink is the idea that the federal government should 'sponsor the development of a new planning system for the land transport infrastructure network', rather than the oft-suggested development of a prescriptive national transport infrastructure plan. AusLink also encourages a multi-modal focus for transport funding, in place of the current system of separate funding programs for road and rail, and stresses the importance of freight transport in strategic policy considerations.

The AusLink initiative was first announced in 2002, but the proposed changes to transport funding arrangements between the federal and state governments are contentious, particularly as the states believe there will be no increase in the quantum of funds and diversion of road grants to finance rail projects. Discussion on the proposal centred

on a Green paper issued by the Federal Government in 2002, followed by release of a White Paper in 2004.

The AusLink initiative is predicated on a need for more and better evaluation of transport plans, a move to needs-based assessment rather than formulae-based allocations, and more flexibility in the realisation of those plans. Amongst the results could be improved pricing practices, more consistent and rigorous transport planning, and renewed recognition of the value of transport research in generating ideas. The paper explores these issues, discusses lessons learnt from previous national transport policy initiatives, and progress towards implementation of the AusLink program.

Background: road funding in Australia

Responsibility for the management and funding of road construction and maintenance in Australia is shared by all three levels of government (federal, state and local) and the private sector. Under the current inter-governmental Road Funding Agreement, dating from 1991, the federal government accepts financial responsibility for the designated National Highway System, 18500 route kilometres of main road linking the major cities. The six state and two territory governments are responsible for the operation of major arterial roads within their boundaries, including those in towns and cities (plus roads in unincorporated areas, mainly in outback Australia), whilst the many local governments are responsible for the extensive urban and rural local road networks.

Financial outlays on roads by governments vary year on year around what were once approximately equal one-third shares by each level of government, though in recent years the states and territories have contributed up to almost half total outlays, e.g. in FY1997-98 $A3.4bn out of a total of $A7bn. (1). The Commonwealth's road expenditures also include special programs, e.g. the Roads to Recovery program to assist local governments, and the Black Spot program that has a safety focus (DoT&RS, 2002).

Private sector financial involvement in road infrastructure is a relatively new element that to date has been into major projects in Melbourne, Sydney and Brisbane, where traffic densities are such that returns can be achieved from toll roads, bridges and tunnels. While in the total picture private investment will be a small proportion of total expenditure on roads, it is important in removing some high cost projects from calls on the public purse (EPAC, 1995). Public acceptance of tolls, at least in the big cities, raises hopes that the traditional revenue raising for roads through fuel excise and vehicle registration charges (whether hypothecated or delivered through consolidated funds) might one day be replaced, at least in part, by charges more closely related to road use. In Australia, as in North America and elsewhere, there is gradual acknowledgment that mass-distance charges are appropriate for heavy vehicles, though changing from existing charging systems is a slow process (NRTC, 1993).

Constitutionally, the responsibility for roads in Australia is vested in the state governments, but the Commonwealth has several constitutional powers under which its involvement in roads can be justified, including roads in the Territories (2) and on other Commonwealth property, interstate trade and commerce), and grants of financial assistance. The Commonwealth's exclusive power to collect fuel excise is contained in section 90, the customs and excise clause (Dowcra, 1993). Successive Commonwealth governments have attempted to use road funding to further the national interest, in the way that the Trans-Canada Highway was seen as a nation-binding project in Canada, e.g. the Australian Bicentennial road program.

The current Commonwealth Government decided that more comprehensive reform of road funding was required in association with improved national transport planning. In May 2002 the Minister of Transport & Regional Services announced a new federal initiative – AusLink – to reform planning and financing of transport infrastructure. The nine-point proposal of inter-related activities which constitutes the AusLink program is to address the need for 'better advice, better planning and better infrastructure'(Anderson, 2002). Then in November 2002 the Minister released a Green Paper,

AusLink: Towards the National Land Transport Plan (DoT&RS, 2002), based on the nine-point theme:

- Development of a National Land Transport Plan;
- Integration and improvement of the National Land Transport Network;
- Establishment of a National Transport Advisory Council;
- Generation of ideas from a wide range of sources;
- Increased emphasis on, and funding of, innovative solutions;
- Application of a consistent approach to funding;
- Joint development and funding of projects;
- Emphasis on improved information and consistent evaluation; and
- Negotiation of a new intergovernmental agreement on land transport.

The AusLink Program Proposals

The Green Paper consists of five chapters, with the body of the AusLink program of initiatives outlined in Chapter 3. The first chapter, 'Current Demand and Future Challenges' sets the background to the proposals for change, by describing the transport task in Australia and the current framework within which transport planning and funding operates. Then follows a listing of a number of weaknesses in the current transport system and its decision-making processes: a focus on the short term and on *ad hoc* solutions; a lack of coordination between modes and a lack of cooperation with the private sector; and insufficient recognition of the potential of technological solutions. In short, the Green Paper described existing approaches as inadequate and suggests incremental change needs to be supplemented or replaced by major reform.

In Chapter 2 the Green Paper set out some 'better responses' to the weaknesses described earlier: improving long term planning of transport and land use; increased private sector investment; identification of a strategic national network of railways and roads (to which investment will be targeted); consistent pricing; and strategic application of technology. Chapter 4 accentuated the 'regional' in the

Minister of Transport & Regional Services portfolio of responsibilities, with a focus on local government and on non-metropolitan issues and initiatives, and Chapter 5 covered process-related matters such as the intended decision-making timetable, a sample planning and funding cycle, and the legislative changes and transitional arrangements that would be required to implement the AusLink package.

Most of the nine inter-linked initiatives described in Chapter 3 were changed to varying degrees between the Minister's speech and publication of the Green Paper some six months later. Whereas the nine points in May 2002 could all be classified as either planning or funding initiatives, the first of the Green Paper's proposals – 'Integrating and improving the National Transport Network' – became a new introductory element, clarifying the aims of AusLink, outlining what the program seeks to achieve in funding the combined national road and rail system, and reflecting the Commonwealth's priorities through investment in transport.

Three of the remaining eight points were planning-related: the development of a National Land Transport Plan; establishing a national advisory body; and 'embedding continuous improvement' in the planning process, including more rigorous assessment and evaluation. All three changed in tone, and one of them in substance, between the May statement and the Green Paper (Scrafton, 1993).

The most significant change is from an initial idea that the Commonwealth would 'sponsor the development of an indicative 5-10 year *transport infrastructure plan'* [our emphasis], to 'sponsor the development of *a new planning system* for the land transport infrastructure network', involving the 'development of a 5-year rolling plan for development and maintenance of the national network'. There was also a positive change in approach to the advisory committee. In his initial announcement the Minister had stated the Commonwealth was 'open to the idea of a national advisory body to assist in developing the national infrastructure network'. The Green Paper's incorporation of the Australian Transport Council's support for a

National Transport Advisory Council as a body to provide advice to AusLink is much more ambitious and appropriate, adding inter-modal matters and pricing to any responsibilities for assessment of infrastucture. The NTAC proposal emerged from a statutory review of the National Road Transport Commission, and ensures the revamped and now renamed National Transport Commission's predominantly regulatory responsibilities are not complicated by such politically sensitive issues as investment and evaluation. Until recent years the Australian Transport Council, the twice-yearly meeting of Commonwealth and State Ministers responsible for transport (railways, roads, ports, aviation), was itself called the Australian Transport Advisory Council, a title that accurately reflected its discursive and consensual approach to national transport matters. In the 1990s, when the NRTC's recommendations were required to be voted on and disapproved (or otherwise automatically approved) by the roads/transport Ministers acting as the Ministerial Council on Road Transport, ATAC and MCRT were merged as the ATC.

The proposed improvements in advice on national transport planning include greater 'consistency, transparency and accountability in decision-making'. The section of the Green Paper entitled 'Embedding continuous improvement' would, if implemented successfully, see a return to the quality of transparent evaluation processes used by the former Commonwealth Bureau of Roads and the Bureau of Transport Economics in the late 1960s and early 1970s. One clause does, however, leave an out for those not too impressed by rigorous evaluation of their favourite project: the methodology must be 'able to be tailored to fit the Government's objectives (economic, social safety, environmental and regional)'. A valid inclusion from a policy viewpoint, maybe, but any one or more of that list of objectives could be invoked to give the nod to almost any project. It would allow, for example for economic criteria to be set aside in the interest of regional development, which may be a euphemism for creating employment, in which case the so-called transport project is more accurately a construction and/or make-work project. Aside from this concern, any move to more and better evaluation has to be supported, as does transparency, which is increasingly necessary, including the

removal of the ability to invoke commercial confidentiality, to be able to assess whether taxpayers funds are being applied effectively and honestly.

One of the key points in the AusLink proposal is within the Federal Government's own domain (amalgamating Commonwealth land transport funding), two require the agreement with state governments (renegotiating the 1991 Roads Agreement, and moving beyond the National Highway System to allow joint funding of land transport), and two are process-related innovations (project bidding and competitive leveraging). Bidding and leveraging could be applied within the Commonwealth's own programs, in the same way that private sector involvement is promoted, but there is greater scope for success if state and local governments would agree to include some of their projects in the process.

The broad scope of AusLink is demonstrated in the sections of the Green Paper entitled 'Funding the best solutions', where the opportunity will be given to all levels of government, 'the private sector, user organisations and regional development bodies' to submit proposals or projects, and for innovative solutions to problems to be put forward for funding. The incorporation of non-engineering solutions into funding programs could lead to improved pricing practices, to more consistent and rigorous transport planning, and to renewed recognition of the value of research in generating ideas.

The reactions to the AusLink funding proposals have been mixed. Some groups, including some state governments, support the Commonwealth moving beyond road funding to cover investments in rail and even urban transport (RTBU, 2002). Others in the community tend to favour retention of the *status quo*, whatever its weaknesses. Which side of the fence particular parties occupy is largely determined by their assessment of whether a change from the existing funding arrangements to a more needs-based process would be to their advantage. The quantum of funds under consideration is obviously important, and there is a concern that unless more funds are available

from all sources than are currently available for roads, spreading those funds more widely across other modes will not necessarily result in improved transport service.

The AusLink proposal to require bidders to support their projects with comprehensive economic, environmental and social appraisal should help to reduce the 'shopping list' approach to transport investment. Consistent and transparent evaluation should be applied to all projects, including public-private partnerships. The transport scene around the world includes several failed PPPs. The Green Paper did not clarify to what extent the Commonwealth Government is willing to be the guarantor of last resort, though some indication might be read into its intention to encourage competitive leveraging of proposals, with those having a greater level of investment by the proponent likely to be given higher priority for funding support.

AusLink 2004: The White Paper

The White Paper, *AusLink: Building our National Transport Future,* was released in June 2004 and represents the federal government's latest policy statement on land transport. Its structure is similar to that of the Green Paper, with the message from the public and industry (after some 18 months of discussion on the Green Paper's proposals) that 'governments need to better coordinate and resource their roles in infrastructure development while delivering national solutions to national priorities'. (DoT&RS, 2004)

Chapter 1 of the White Paper revisits the first two chapters of the Green Paper, with a small additional section on airport infrastructure, as passenger numbers in 2020 at major airports will be at least double levels in 2000. The chapter reiterates the Commonwealth's position on urban transport, stressing the fact the public transport is a responsibility of state governments.

Chapter 2 is important in redefining the AusLink National Network. The Green Paper proposed an initial national network, but proposals were advanced in consultations and submissions for inclusion of

additional corridors and links. The result is that the AusLink National Network is now more extensive than previously proposed, though the federal government's intent to invest in projects that are clearly of national priority and have great national benefits is reiterated. There is a definite expectation that state governments will invest in those projects on the national network that will bring benefits to the states.

Integrating land use and transport planning is mentioned in Chapter 2 as being seen by the federal government as necessary to improve the performance of the national network, but there is no further reference to the topic elsewhere in the White Paper. Similarly, the Green Paper's emphasis on private sector investment is maintained, yet Chapter 3 of the White Paper states that funds for the Thoresby Highway and other projects will not be forthcoming unless existing plans for tolls are dropped.

Chapter 3 is the main body of the White Paper, spelling out projects and funding, though there is little detail on any of the new elements. In particular, an extensive list of 'national investment priorities' is presented, but it is not clear whether the list is in priority order. Strategic directions to improve long-term infrastructure planning and to guide investment priorities are described, including 'links of most concern' such as those to major ports, airports and intermodal centres. Major metropolitan corridors slated for investment are clearly outlined, while the addition of specific rail projects for federal funding support represents a new and significant addition to existing funding proposals.

Spending of 1% of the AusLink budget on research, development and technology is significant, with the federal government committing to encourage applied R&D, adoption of new technologies and development of national standards and protocols.

Chapter 4 of the White Paper, covering local and regional transport infrastructure, is similar to the regional material in the Green Paper, and Chapter 5 describes the processes associated with the implementation of AusLink: legislative, intergovernmental, and

institutional arrangements. The White Paper states that the National Transport Advisory Council will be established to provide strategic analysis and advice to Ministers. However, the NTAC is still not in place, nor is it clear whether the NTAC will be responsible for applying the project assessment methodology to all the projects listed in Chapter 3, or whether 'grandfather' provisions will be invoked to allow some projects to proceed without further evaluation.

The new assessment methodology is a major advance, and the concept of a new 'national transport data framework' is a necessary element of such an assessment process, providing access to data on which to base national strategic plans and infrastructure investment decisions (Ockwell *et al,* 2004). The quality of the data compiled will be critical to its utility.

A new Chapter 6 on complementary policy issues covers related matters including transport safety, security, regulation reform priorities, environmental outcomes and energy efficiency, technology, mobility, and freight logistics.

Potential Benefits of AusLink

National transport policies and plans are desirable, to ensure, *inter alia,* that infrastructure development responds to needs, and not merely satisfies local political aspirations. However, developing national transport policy in a federal nation is a daunting challenge. States or provinces appreciate receiving federal funding for projects or programs that put pressure on their budgets, though they are usually reluctant to cede any powers, authority or control in exchange for the assistance.

Most Australian states have privatised much of what were previously primary functions of their governments, such as power, water supply and railways, and transferred other activities to local government, so we might postulate a long-term future with only two levels of government: national and regional, the former evolving from the federal government, and the regions being built from stronger, larger

and more effective local governments than exist at present. In the immediate future, and within the constraints of the existing institutional context, improved national transport policy formulation and implementation are necessary. AusLink provides a framework for current processes to be improved.

A major change proposed by AusLink is the extension of the present arrangements whereby federal funds are limited to roads, to provide federal financial assistance to development of other transport modes. The main beneficiaries are likely to be railways. Just as the main federal interest in roads is in National Highways, so any funds for rail will be directed to a designated network of main lines, mainly the interstate standard gauge routes.

As noted above, state governments' main problem with AusLink is the uncertainty about the quantum of federal funds to be available under the program. They are also concerned about the emphasis on the national network that, other than the National Highways city mileage, and road and rail links to seaports and airports, excludes urban transport. Given that many other former State and Territory activities have been privatised, urban public transport is one of the remaining elements of the transport system still owned and paid for by the States.

Australian federal government interest in national transport policy is not new, but the extent of federal involvement in the modes varies and is slow to change. The Commonwealth Government, and staff in its departments and agencies, has for many years sought to improve the evaluation and performance of its transport programs, and from time to time has contemplated influencing more directly land transport and funding. A major issue has been the extent to which Commonwealth policy on roads 'should attempt to directly influence road investment and maintenance priorities' (Dowcra, 1993).

A major change in policy occurred with the election of a federal Labor Government in 1972 that sought much stronger federal involvement in transport and urban affairs. Prime Minister Whitlam told the State Premiers in 1973: 'We believe that it would be irresponsible for the

National Government *(sic)* to content itself with simply providing funds without being involved in the process by which priorities are set, and by which expenditures are planned....'(quoted in Maclean and Hickman, 1976). The implementation of this policy created a great deal of opposition and debate, yet when the Liberal-led coalition was returned to power there was only minor relaxation of Commonwealth controls over road investment decisions, with a sweetener in the form of the Bicentennial Roads program introduced in 1982.

Successive governments at all levels can posit different approaches to transport planning and funding, but the problems remain to be solved on the ground. It became obvious in the 1980s in Australia that while the *status quo* of formula-based road funding plus occasional special programs and intermittent financial support for railways and urban transport might be good for political kudos, the combination was having limited effect on congestion, pollution, subsidies and other problems.

The 1990s were therefore characterised by a succession of formal inquiries and less formal attempts at improved national transport planning, regulation and investment. Inter-governmental agreement on National Competition Policy became a driver for reform, with the NCP directly responsible for the establishment of the National Road Transport Commission, and though the Commission was not to be involved in road funding, it was empowered to recommend road charges for heavy vehicles.

Two of the transport inquiries were undertaken by the Industry Commission, into rail (IC, 1991) and urban transport (IC, 1994); one by the Economic Planning Advisory Commission, into private investment in economic infrastructure (EPAC, 1995); and a follow-up on the progress into rail reform by the Productivity Commission (PC, 1999). In addition the National Transport Planning Taskforce reported to the Federal Government (NTPT, 1995), and the Tax Research Foundation completed a report on Taxing Cars for the Australian Automobile Association (Pender, 1999).

Although the topics were different, there were common themes running through the reports of these inquiries: the need for efficiency, the need for transparent evaluation, and the need for a national perspective to be adopted and/or federal leadership in tackling national transport issues. Peter Abelson, in a commissioned paper for the National Transport Planning Taskforce, pointed out that the Commonwealth and most State governments had guidelines on the use of economic analysis, but they were rarely used (Applied Economics, 1995). All these themes reappeared in the Green Paper, and AusLink offers hope for more rigorous application of these or similar guidelines to road and transport projects in future.

The delay in progressing from the Green Paper to a White Paper was beneficial in providing time for the development of project evaluation standards and procedures that are acceptable to all parties. Officers of the Commonwealth Department of Transport & Regional Services have chaired working groups including state and local government representatives to develop draft national transport investment methodologies and appraisal guidelines, and to assess the availability and quality of data required to enable analysis of alternative transport modes in inter-city corridors.(Ockwell *et al,* 2004)

Related institutional changes should also contribute to improved performance from Australian transport. The conversion of the former NRTC to the National Transport Commission, in January 2004, added technical and regulatory functions for rail to the Commission's road transport responsibilities. NRTC has demonstrated that state governments can work together, given the right environment for issue resolution and inter-governmental decision-making. The next step, a rather more contentious one, is to set up the National Transport Advisory Council, to develop a parallel program in the areas of transport planning, pricing and investment.

Conclusions

The Australia of the future to be served by the transport system will be different from that which existed when the existing inter-

governmental agreement was formulated. Above average growth of traffic will occur mainly along the coastal strip of eastern Australia from south of Sydney to north of Cairns, which happens to be the part of Australia where geology, topography and climate combine to create difficult conditions for road and rail transport. Investment priorities will have to change to reflect the transport needs of this 2500 kilometre long east coast corridor. While it is true that some poor decisions were made in the past that have exacerbated the transport problems of the eastern states, that is no reason to ignore current pressures and/or argue for the maintenance of existing funding allocations.

The emphasis on improved appraisal of projects to be submitted for funding under the proposed AusLink program, the plan to evaluate projects in the context of the corridors that make up the national transport network, and the development of nationally consistent databases and evaluation techniques, will together lift the standard and performance of the transport planning profession and hopefully improve national decision-making for transport in Australia.

In summary, AusLink has the potential to move Australia forward by investing its transport dollars more effectively. The patterns of movement are well established and the maintenance of existing transport infrastructure will continue to be important; the scope for improvement lies in investing wisely to improve and extend the network.

Endnotes

1. $AUS 1.00 was $CAN 0.90 at the time of writing.
2. Local governments are legally instruments of the states; and the two territories (NT and ACT) are instruments of the Commonwealth, but govern internal matters as if they were states.

References

Anderson J (2002). Address to the Re-Engineering Forum. Sydney, 21 May.

Applied Economics (1995). Evaluation and Decision Making for Transport Infrastructure Investment. National Transport Planning Taskforce, Commissioned Work Volume 2. AGPS, Canberra.
Department of Transport & Regional Services (2002). AUSLINK – Towards the National Transport Plan. Green Paper. Canberra.
Department of Transport & Regional Services (2004). AUSLINK – Building our National Transport Future. White Paper. Canberra.
Dowcra G (1993). Roads Policy and Australian Federalism. BTCE Occasional Paper No 106. AGPS, Canberra.
Economic Planning Advisory Commission (1995). Private Infrastructure Task Force Report. AGPS, Canberra.
Industry Commission (1991). Rail Transport. Report No 13. AGPS, Canberra.
Industry Commission (1994). Urban Transport. Report No 37. AGPS, Canberra.
Maclean JD and Hickman JM (1976). A More Critical Look at some of the Assumptions used in Allocating Commonwealth Government Assistance for Land Transport Investment. Papers of the 2nd Australian Transport Research Forum. Adelaide.
National Road Transport Commission (1993). Cost Allocation and Charging. Working Paper 10. Melbourne.
National Transport Planning Taskforce (1995). Building for the Job – A Strategy for Australia's Transport Network. AGPS, Canberra.
Ockwell A, Tisato P, Harvey M, Evans C *et al* (2004). Development of a national appraisal methodology for transport proposals. Paper to 27th Australasian Transport Research Forum. Adelaide.
Pender H (1999). Taxing Cars – Fleecing the Fleet or Subsidising Smog. Australian Tax Research Foundation Research Study No 33. Sydney.
Productivity Commission (1999). Progress in Rail Reform. Report No 6. AusInfo, Canberra.
Rail, Tram & Bus Union (2002). AusLink – a Positive Federal Initiative, but too many missing links. RTBU Discussion Paper. Sydney.
Scrafton D (2003). The AusLink Green Paper – proposals, questions and consultation. Transport Engineering in Australia, Vol 8 no 2 pp 73-79.

Cargo Airships: Civilian Applications and National Defence

Dr. Barry Prentice, Professor, I.H. Asper School of Business, and Director, Transport Institute, University of Manitoba

Introduction

Civilian aviation and marine markets are essential for the military to be able to afford the volume of sealift and airlift required to meet the needs of an emergency. When needed, the military rely on stand-by contracts with civilian operators to provide crews and craft for military purposes. As the merchant marine demonstrated during World War II, military success can be highly dependent on the existence of a well-developed civilian transportation industry. Consequently, national defence always plays a role in transportation policy.

The development of a new generation of airships is creating a new tactical option for strategic emergency response. Airships could provide the military with a new asset to support the logistics of emergency situations that are not easily handled with existing means of transport. Airships could fill a gap between ocean shipping that has great capacity and low cost, but slow deployment, and airplanes that offer great speed, but limited capacity at a high price. Airships could complement other forms of transport and offer the military a low-cost, flexible logistics tool.

The military needs a commercial civilian market for this potential to be available[1]. The objective of this paper is to outline the potential commercial market for airships in Northern Canada, and the corresponding opportunity that airships present to the Canadian military. The paper begins with an overview of the northern transportation market and the state of the international airship industry. Subsequently, two economic case studies for using airships in the North are reviewed briefly. The last section of the paper considers the policy merits of airships for the Canadian military.

of the paper considers the policy merits of airships for the Canadian military.

Transport Options for the North

Canada faces significant transportation challenges and issues in the 70 percent of its northern landmass that has few all-weather roads. This region is characterized by vast distances, thin markets with low income populations and harsh, cold weather conditions. The sparse traffic density and few back-haul opportunities often lead to monopolies in transportation, and of course shipper complaints. At the same time, the North is a very fragile environment. Wildlife protection groups and the First Nations are concerned that the construction of all-weather roads over permafrost will harm the delicate flora and fauna of the tundra.

The four alternatives for transportation in the Canadian North are sealift, barges, and airlift and winter roads. Sealift is the least expensive form of transport available to communities with ocean access, and is capable of delivering 5,000 to 25,000 tonnes of fuel and other cargo, but only once per year[ii]. The size of the cargo also increases the potential environmental hazard. If a major oil spill occurs in the Arctic seas, Canada has no capability to respond rapidly (Prentice and Phillips, 2002).

Communities located on the coast of Hudson's Bay and the McKenzie River can take advantage of barge shipment. Barges are more convenient, but more expensive than ships because of the extra handling and smaller shipment size. The convenience advantages of barges can be measured against the costs of sealift where missing a sealift can delay a project for a full year. Although barge transport is interrupted by the long winter and the shoulder seasons, they permit several shipments per year.

Airplanes offer the greatest convenience with year round, fast service. The limitations of fixed-wing aircraft include the conditions of the landing strip, high freight rates and cargo weight and dimensions. Indivisible freight can only be shipped if it can fit through the airplane's cargo door. Oversized cargo or oddly shaped pieces must be shipped by other means.

Figure 1 presents typical aircraft that are used in the North. Generally speaking, these aircraft are old and near the end of their physical life. Parenthetically, it is worth noting that the industry is concerned that replacement aircraft are much more expensive and require longer runways (Aalders, 2003).

Figure 1 Airplanes used for Northern Transportation



Heavy cargo is shipped to remote communities in the interior via winter roads. Tractor-trailers operate on these temporary roads that are built over frozen lakes and cleared bush. Winter roads are open for one to three months annually depending on location and the weather in a given year. In addition to the annual cost of building the roads, truck shipments are 65 to 70 percent more expensive than the equivalent truck transport over all-weather roads[iii]. These trucks travel more slowly, suffer greater damage to suspension and tires, experience longer delays and face more risk associated with crossing frozen lakes.

In Manitoba, over 2,000 kilometres of winter roads are built every year. If stretched out and linked together, Manitoba's winter roads would reach from Winnipeg to Vancouver. Conversion of these temporary roads to all-weather gravel roads would cost approximately $1 billion plus the cost of annual maintenance. Major sections of all-weather roads are unlikely to be constructed any time soon in Manitoba where the province struggles to maintain its existing road network (Manitoba, 2003).

Climate change is a global problem, but has greater impact in the North (ACIA, 2004). In Manitoba, winter roads serve a population of 25,000 to 30,000 located in about two dozen remote communities. Winter roads usually open in January and close in March. Prior to the mid-1990s, winter roads east of Lake Winnipeg were able to operate 50 to 60 days every year. In half the years since 1996, barely 30 days of operation has been possible. In a severe year, like the El Nino of 1998, many winter roads can fail[iv].

First Nations populations are growing rapidly and Northern political leaders are demanding a better standard of living. Remote communities want economic development opportunities and better public services[v]. More resource development is happening in the north, but it is mainly the "low hanging fruit" variety - diamonds and gold mining. The advantage of mining gems and precious metals is that a month's worth of production can be brought out in an airplane. Base metals and other mineral products are much harder to develop because the cost of transporting these products from the mine to the market is too great in proportion to their value.

Many rich ore bodies have been identified in the North. Izok Lake, which is located near the border of Nunavut and the Northwest Territories, is one of the richest copper deposits on earth. Plans have been put forward to build a deep-water port at Bathurst Inlet on the Artic Ocean and to construct an all-weather road to connect Izok Lake. Added to the cost of building over permafrost, this road would run through the centre of the caribou calving lands. Even if the environmental concerns could be mitigated, copper prices would have to become very high to justify the cost of this road.

Although it may seem preposterous to envision airlifting base metal concentrates, this may be the most practical means of developing environmentally sustainable mining in the North. A new generation of airships is being developed that could offer freight rates that make such activities economic.

Figure 2 compares the economic and operating characteristics of the various means of northern transportation. Airships provide a very flexible option relative to other transportation modes commonly used in the north. Airships can be operated year round and are relatively independent of ground-based infrastructure. Furthermore, airships can be configured to carry a combination of passengers and cargo economically.

Figure 2 – Characteristics of Transportation Modes for the North

	Sea-lift	Barges	Airplane	Trucks	irships
Freight Cost	Lowest	Low	Highest	Medium	Medium
Payload Capability	5,000-25,000 T	200-500 T	1-10 T	20-42 T	1-300 T
Seasonality	Summer only	Summer only	Year round	Winter only	Year round
Delivery Speed	6-8 Weeks	1-2 Weeks	2-24 Hours	3-7 days	8-48 Hours
Infrastructure Requirements	Minimal	Moderate	High	High	Minimal
Infrastructure Costs	Minimal	Low	High	High	Minimal
Flexible access to delivery site	Limited to seas	Coasts and rivers	Runway limits	Limited by roads	No limits
Freight and passengers	No	No	Yes	No	Yes
Environmental Impact	Oil spills	Oil spills	Runways	Ice road routes	Minimal
GHG Emissions	Low	Low	High	Medium	Low

Airships can deliver large payloads and transport indivisible loads at speeds that are second only to fixed-wing aircraft. Moreover, they have a small environmental footprint. Airships require minimal infrastructure and have low greenhouse gas (GHG) emissions.

The Potential for Airships

Seventy-five years ago, giant Zeppelins crossed the Atlantic Ocean on a regularly scheduled basis, and could carry up to eighty tons. What makes airships different today is the opportunity created by new materials, better design tools and helium gas[vi]. When the Zeppelins were at the cutting edge of technology, nylon was a revolutionary material used for expensive ladies stockings and the Model A Ford was a popular car. The same technological advances that have been experienced by industrial textiles and airplanes since the 1930s can be applied to airships.

Figure 3 presents a blow up of the Tedlar/Mylar/Dacron composite fabric that TCOM uses to build envelopes for modern aerostats and airships. This "skin" material resists ultraviolet light, is non-flammable and loses less than 10 percent of the helium lifting gas per year. The breaking strength for a one-inch strip of this material is 500 pounds (Barlow, 2002). TCOM Aerostats built with these envelopes have been in continuous use for 10 to 15 years without problems[vii].

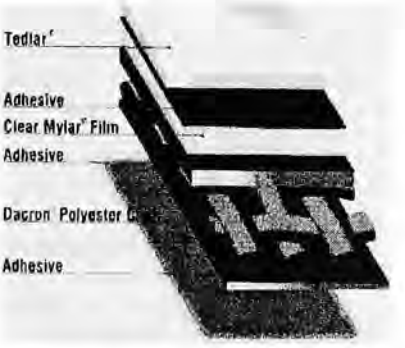


Figure 3
Envelope material used by TCOM for aerostats and the Skyship airships

Modern airships with envelopes made of this long lived, durable material can function without an elaborate internal structure. An internal catenary curtain that is suspended from patches sewn into the top of the envelope can support hard structures, like the gondola. In contrast to these durable envelope materials, the great Zeppelins were clad with painted canvass. This envelope added no structure, would become water soaked and is blamed for causing the fire on the Hindenburg (Van Treuren, 2001). New envelop materials make airships safer and more robust without adding weight. They also permit shapes and design concepts for airships that could not have been contemplated 70 years ago.

In the future, some airships will look very different than the traditional cigar shape. Researchers are testing hybrid lifting body airships and symmetrical airships, as well as conventional ellipsoid airships. Many airships in the future are going to be lightweight structures that are 30 to 50 percent heavier than air[viii]. Hybrid vehicles that take advantage of their hull shape and propulsion for aerodynamic lift could have a number of operational advantages. For example, these heavier-than-air vehicles could drop off cargo and return without taking on ballast, and may be able to offer faster ground-handling turn around times.

The competition to develop the next generation of airships is global in nature. Figure 4 lists 12 airship companies that are located in seven different countries that have products at various stages of development. In addition to these private companies, South Korea and Japan sponsor state research establishments that are testing remotely controlled airships for high-altitude applications. The range of activities of these companies includes aerostats, lighter-than-air (LTA) vehicles and hybrid vehicles.

An exciting part of the future for airships is the ability to use solar collectors for power. High altitude airships designed for missile defence and wireless communications require the development of solar collection systems. In the stratosphere,

where these airships will operate, the atmosphere has only one-fourteenth of its density at ground level. No airship could carry enough fuel to stay aloft and power its equipment for six months. The solution is thin membrane solar collection technology that can be woven into the envelope of the airship[ix]. What works at 70,000 feet should also work at 5,000 feet. Most commercial airships in the future are likely to be solar assisted in some fashion.

Figure 4 Leading Airship/Aerostat Developers

Location	Company	Aerostats	LTA Vehicles	Hybrid Vehicles
U.S.A.	Lockheed Martin		design	design
	TCOM	certified		
	American Blimp Co.		certified	
	Air Mgt. Services		certified	
	World Aeros	certified	prototype	design
	Ohio Airships			model
Germany	Zeppelin Luftschifftechnik		certified	
U.K.	Advanced Technologies Group		certified	model testing
Canada	21st Century Airships		prototype	
Russia	RosAeroSystems	certified	certified	
China	Vantage Airship Co.		certified	
Malaysia	ALOM		model	

The intended markets for next generation airships are high-altitude communications and defence surveillance, advertising, tourism, coast guard patrol and heavy lift. Airship designs will vary according to their particular application: long haul versus short haul, heavy lift versus high altitude and so on. Two cases for freight transportation in the north have been researched by Prentice and Thomson (2004, 2003). The first case examines the economics of using an 84-tonne lift airship to haul fuel to the Ekati diamond mine in the Northwest Territories. The second case examines the feasibility of using a smaller 30-tonne lift airship of a similar design to support remote communities in northern Manitoba.

The Ekati diamond mine is served by a 400 kilometre long winter road northwest from Yellowknife. The winter road also serves Diavik diamond and Lupin gold mine. Other mines are planning to open, but how much truck traffic can be put over these lakes is uncertain. The Ekati mine requires over 180 million litres of fuel a year to be moved over a winter road in tractor-trailers. The transportation cost is $40 million and these logistics have to be completed in 12 to 16 weeks.

The key assumptions of the model are based on an 84 tonne lift hybrid airship that can move 100,000 litres of fuel with an average speed of 161 kilometres per hour. The purchase price of the airship is US $60 million and the desired return on capital is ten percent. The airship is assumed to haul only fuel and is paid at the same rate as trucking over the winter road. No backhaul is considered. A detailed description of the analysis is presented in Prentice and Thomson (2003); the conclusion is that the airship reaches its profit target after 245 days of operation. Any revenues beyond the variable costs for the remainder of the year would be pure profit.

Of course, if this airship were available the first thing that any mine operator would say is, "hold the fuel; we have got more valuable things to move". Everything from mining equipment to spare parts and camp supplies need to be brought in over the winter road. With year round service, fuel would likely be used to top up the load on the airship as opposed to being the primary cargo. The cost reduction implications for the mine are obvious, but are not quantified in the study.

The second case (Prentice and Thomson, 2004) considers the needs of six remote First Nation villages in Manitoba. A 30-tonne lifting body airship is considered that has a purchase price of US $30 million and operating characteristics similar to the 84-tonne hybrid. The *2020 - Manitoba Transportation Vision* study done by the Province of Manitoba provides data on the freight and passenger demand of these remote communities. Figure 5 presents the amount of goods and the number of flights and passengers to these various communities. The analysis assumes that 75 percent of the goods are brought in by truck over the winter roads and the remainder is shipped by air.

Figure 5 – Passenger and Freight Demand Profiles: Manitoba Remote Communities

Community	Pop.	Annual Freight	Number of passengers	Number of flights
St. Theresa Point / Wasagamach	4,021	11,600	14,127	6,061
Garden Hill / Island Lake	3,204	13,000	37,933	11,245
Red Sucker Lake	773	3,200	6,073	1,923
Oxford House	1,998	8,150	9,649	3,102
God's Lake Narrows	1,393	6,260	9,741	3,396
God's River	466	2,100	5,957	2,399

Source: *2020 - Manitoba Transportation Vision*

The economics comparison uses a common origin at Winnipeg, Manitoba. The airship costs include trucking to Thompson where the airship completes the transport on direct flights to these six villages. Operationally, the analysis assumes that airship flies year-round, but in practice, utilization is limited to 285 days per year because of maintenance, crew rest and inclement weather. Annual utilization is based nine hours flight time per day for scheduled cargo/passenger operations. Point-to-point moves are not as efficient as a circular route, but for the purposes of modelling, a simpler route was used. No backhaul was considered in any of these cases. The airship is assumed to operate in a combination fashion (passengers and freight), which is common in the north.

The results of the analysis are presented in Figure 6. If the 30-tonne hybrid were only moving freight, it would not be competitive with existing transportation at this level of utilization. However, by combining passengers' fares with freight revenues the airship is competitive in four of the six communities. The truck costs do not include the cost of building the winter roads. Also, the indirect cost of inventory is ignored. If goods are only received once a year, everything has to be

bought and paid for at once. Then the inventory must be stored, which leads to high holding costs. These inventory holding costs are a major burden put on all remote communities.

Figure 6 – Comparative Costs for a 30-Tonne Hybrid to serve Remote Communities

Destination	Existing Costs ($/tonne)	Hybrid No Passengers ($/tonne)	Hybrid With Passengers ($/tonne)
St. TheresaPoint / Wasagamach	$435	$671	$304
Garden Hill / Island Lake	$450	$676	$366
Red Sucker Lake	$500	$696	$550
Oxford House	$530	$591	$456
God's Lake Narrows	$500	$648	$502
God's River	$525	$650	$552

Source: Prentice and Thomson, 2003

National Security and Economic Opportunity

The North is a logical place to introduce cargo airships. First, freight rates are high which would help a fledgling airship industry get established. A jug of milk that costs $3 in Winnipeg can cost $8-10 in these remote communities. High freight rates in the North account for most of this price difference. Second, competition is minimal, so once airships become profitable, their margins can be sustained. Northern shippers rely on older airplanes, winter roads, and not much else, especially for heavier lifts and awkward shapes. Third, the market in the North has enough potential to justify assembly line production of airships. The territory to be served extends from Labrador to Tuktoyaktuk. Finally, airships could help defend the sovereignty of the Canadian Arctic and provide significant social benefits to the citizens of these remote communities. Poverty lies at the root of most health and social issues. Effective transportation is necessary for Northern communities to attain economic viability.

Foreign submarines and incidents of unauthorized surface ships in the Arctic challenge Canada's claim to the North. Canada can protest but has negligible means of securing the borders of this area[x]. Sovereignty is becoming more important in the Arctic with evidence of a melting ice cap. Airships could give Canada an economic means to increase patrols in the North and defend Canadian sovereignty on all three coasts.

In terms of military deployment, airships provide a new ability to move bulky cargo. Helicopters that are fully assembled and ready to fly could be carried on an airship because these vehicles are limited only by cargo weight (Woodgerd, 2003). The military could take advantage of the lift capacity provided by a commercially driven airship industry that does not depend on DND for procurement, or place an additional burden on DND infrastructure or staffing. As with ocean shipping and fixed-wing air transport, standby contracts could guarantee that airships would be available when they are needed. The government would have to pay only the marginal costs for airship transport when logistical support was really needed. A commercial industry would keep the airship fleet at the cutting edge of technology and provide training of all crews and mechanics.

A civilian airship industry could also help meet Canada's international commitments in NATO and NORAD. With such a vast market for airships in the North, Canada could specialize as one of the main logistics suppliers to the NATO forces. Finally, the best means to enforce Canadian sovereignty in the arctic is to have a more permanent presence. Promoting commercial development in the North is consistent with and supports Canadian territorial claims.

Conclusion

Airships represent an enhancement option that will add to the existing northern transportation capabilities and provide an asset for military strategic airlift to the North and beyond. Canada has national interests in developing energy, resources, mining and mineral resources in the North, as well as serving remote populations better. Airships offer an opportunity to create high

technology jobs in aerospace, to stretch DND budgets, to demonstrate Canada's commitment to pro-Kyoto technologies and defend Canadian sovereignty.

The transportation challenges faced in the construction of a pipeline in the Mackenzie River valley or a mine in Nunavut are similar to a military campaign. Detailed logistical planning is involved in moving a huge volume of material into a difficult zone during a short window of opportunity. Mining companies and energy resource companies are natural partners for the military, as are the First Nations. It is not often that the military, business, environmentalists and First Nations can all support a common commercial development. The next generation of cargo airships presents a rare opportunity that touches multiple cords of National Interest.

References

Aalders, Les. in *Airships to the Arctic Symposium II*, "Moving Beyond the Roads". Barry E. Prentice, et al. (eds) Winnipeg, MB: Conference Proceedings, University of Manitoba, Transport Institute, October 2003: 101-113.

Althoff, William, F. *Arctic Mission by Airship & Submarine to the Far North.* (Second Edition) Auckland, New Zealand: Lighter-Than-Air Institute, 2000.

ACIA (Arctic Climate Impact Assessment). Impacts of a Warming Arctic. Cambridge University Press, 2004. (http://www.cambridge.org)

Barlow, David. in *Airships to the Arctic Symposium.* "Applications for Northern Transportation". Prentice, Barry E. and Shelley Turriff. (eds) Winnipeg, MB: Conference Proceedings, University of Manitoba, Transport Institute, October 2002:15-23.

Manitoba. *2020 - Manitoba Transport Vision.* 2003 http://www.gov.mb.ca/tgs/2020transport/index.html

Prentice, Barry E. and A.J. Phillips. "Airships to the Rescue: Re-supply and Emergency Response in Arctic Resource Development" *Canadian Transportation Research Forum.* Proceedings Issue: 37th Annual Meeting (2002): 371-382.

Prentice, Barry E. and Jim Thomson. "Economics of Airships for Northern Re-supply" *Proceedings of the 5th International Airship Convention and Exhibition*, Oxford, England, August 2004.

Prentice, Barry E. and Jim Thomson. "Airship Fuel Tankers for Northern Resource Development: A Requirements Analysis" *Canadian Transportation Research Forum*. Proceedings Issue: 38th Annual Meeting (2003): 592-606.

Woodgerd, Mike. "Ultra Large Airlifters (ULAs): A North American Window of Opportunity" *Canadian Transportation Research Forum*. Proceedings Issue: 38th Annual Meeting (2003): 587-592.

Woodgerd, Mike. "Mobilus: The Vision and Method to Create a New Component of the US Aerospace Industry Centered Upon Transport Airships". *Transformation Trends*. March, 2004. Also available at www.mtmc.army.mil

Van Treuren, Richard G. "Hindenburg: The Wrong Paint Hydrogen: The Right Fuel" Edgewater, FL: Atlantis Publications, 2001

Acknowledgement: This article is based on an earlier presentation made at the Canadian Institute of Strategic Studies.

---

[i] "The US military will utilize this commercial asset of ULAs [airships] and their support structure in a manner similar to how we now utilize airplanes through the Civil Reserve Air Fleet (CRAF). .... This civilian capability must be broad, deep and develop as rapidly as reasonably possible. Military considerations will be a key component of development from its inception and military will be a crucial first customer as well". (Woodgerd, 2004)

[ii] Goods destined for Iqaluit and other Nunavut and Nunavik communities are assembled in warehouses at Montreal during the spring, then shipped on the Umiavut, Anna Desgagnes, Arctic Viking, Lady Franklin and Aivik. The first ships arrive in Nunavut about mid-July.

[iii] Personal communication, Big Freight Systems, Inc. This freight rate differential applies to Manitoba. Further north, in the Territories, the costs are likely to be higher

[iv] In 1998, the Government of Canada had to spend an additional $12 million to fly in fuel and supplies to First Nations communities in Manitoba.

[v] For example, in Garden Hill, a community of 3,500 people in central Manitoba only a few houses have running water. The rest depend on outhouses and trek water in pails from four wells.

[vi] While non-flammable helium gas was available prior to the Hindenburg, and is used exclusively in modern airships, some design tools like strain gauges were not introduced until after this period. Today, advanced computer design and simulation software tools allow aircraft to be flown on paper.

[vii] More information can be obtained from the TCOM website www.tcomlp.com

[viii] Ohio Airships is building a hybrid vehicle called a Dynalifter that looks like an oversized airplane. It uses a central frame to support stubby wings and body that create dynamic lift that adds to the buoyant lift of the helium gas. The test flight of the first two person Dynalifter is scheduled for spring 2005. More information can be obtained from the Ohio Airships website www.ohioairships.com For other hybrid airships see: www.atg-airships.com , www.aeros-airships.com and http://www.hacinc.us/

[ix] ITN Energy Systems and Iowa Thin Film Technologies are developing the thin-film photovoltaic (PV) arrays for the Lockheed-Martin airship.

[x] As former Prime Minister Brian Mulroney lamented: "not so much as a canoe" (cited in Althof, 2001).

Barry Prentice is a Professor of Supply Chain Management at the I.H. Asper School of Business, and the Director of the Transport Institute at the University of Manitoba. His research and teaching interests include logistics, transportation economies, trade and development. In 1999, Dr. Prentice was named Manitoba Transportation Person of the Year by National Transportation Week.

Some Second Thoughts on the Governance of Canadian Airports

Michel Boucher
École nationale d'administration publique
Université du Québec

1. Introduction

The transfer by Transport Canada to independent not-for-profit entities of the 26 largest airports, initiated at the beginning of the nineties, is now completed. Not only did these new providers offer the full range of services previously supplied by Transport Canada, a public monopoly, but they also developed new commercial opportunities unimaginable in the old regime. Transport Canada managers have high expectations for future competitive services among airports because these new entities are not restrained by the political context and bureaucratic rules that prevail in a public department. In short, they assume that the overall competitive framework built around the delegation will promote efficiency, integrity, transparency and accountability in airport management. Over the years, Transport Canada, the owner of and landlord of airport infrastructures, realized that the rules of governance put in place were not binding and that some airport managers acquired by their location market power whose use led them to go against Transport Canada's interests.

In 2001, the *Canada Transportation Act* Review Panel mentioned some weaknesses in governance and monitoring arrangements that could lead to unsatisfactory future performance in Canadian airports. The Panel also recommended substituting the not-for-profit entity for the more traditional profit maximizing private firm. This institutional

change was seen as solving many problems, among the most controversial being airport rent. This issue has been a Gordian knot since the beginning of the commercialization and it continues to generate animosity between airport managers and Transport Canada. In March 2003, the Minister of Transportation tabled before Parliament Bill C-27 entitled the *Canada Airports Act*. This proposed legislation intended to regulate the manner in which all airport authorities are governed, set fees and charges and manage their facilities. Bill C-27 received first reading in the House of Commons, but died on the order paper when Parliament adjourned in the fall of 2003. A new Bill was expected to be tabled in the first months of that year. The content of the new Bill sought to promote efficiency in airport management and to make management more accountable and transparent than in the past to users and stakeholders.

The purpose of this article is to throw some light on the recent performance of Canadian airport service providers since 2001. Section 2 explains shortly how and why Transport Canada's expectations did not come true. Too many imbroglios and quid pro quos, not enough information on both sides to anticipate the consequences of the delegation process and an involuntary kind of leniency by Transport Canada are at the origin of the problem. Section 3 presents the content of Bill C-27 as a possible answer to the *Canada Transportation Act* Review Panel. Section 4 describes the overall performance of the main airports and evaluates their efficiency, mainly financial ones. Section 5 contains some remarks related to the general behaviour of airports and Transport Canada.

2. The delegation process in the airport industry: An evaluation

The before delegation Transport Canada had a branch in charge of and managing all internal and operational decisions taken to operate the network of Canadian airports. Its public monopoly status allowed it to determine the level and the quality of services offered. Quality is fairly similar across the country, except for a few variations. When commercializing airport providers, Transport Canada managers decided that the not-for-profit private non-share corporation is the

best substitute replacing the various bureaucratic arrangements under the Transport Canada umbrella. The situation changes from a coast-to-coast public monopoly, whose activities and decisions are centralized in Ottawa, to many local natural monopolies, disseminated among metropolitan cities.

Transport Canada prefers the not-for-profit corporation over other corporate forms, for its main characteristic, namely that members who control it have no rights to claim net cash flow[1]. Although accumulating net cash flow is allowed, a not-for-profit organization cannot distribute it among its members. All surplus funds must be reinvested in the organization. Nobody has a right on it[2]. It also means implicitly for Transport Canada that airport managers cannot have incentives to appropriate the net cash flow in some indirect ways. In this new context, each new airport provider, under local control, is expected to exercise self-regulation and to operate in a commercial and cost-effective manner. On the other hand, Transport Canada does not intend to impose any regulations governing some conducts that a new airport provider can adopt, being a potential natural monopoly. In short, the results of the delegation process imply that each local airport offering services and allowing stakeholders opportunities for inputs on its decisions can be more efficient than those managed when Transport Canada was the monopoly operator for all Canadian airports. Evidence confirms how superior is the new regime based on the autonomy of each airport infrastructure provider relatively to the large centralized bureaucratic approach.

The promoter of the not-for-profit approach, Transport Canada, never conceived at all that new airport authorities could have divergent objectives from those expected by the department. Incentives were not aligned between airport authorities and Transport Canada, leading to inefficiencies. An ingredient of market efficiency is that the rules of game be known to every player. In some situations, breaking a rule requires a strong penalty and in other cases, only a minor reprimand. In the real world, enforcements must apply, even in an emerging market as that of the new airport providers in Canada.

We neglect voluntarily miscalculations and mistakes emerging from a trial and error process because variability is reduced over time and leads gradually to an efficient converging path. Our interest is concentrated on inefficiencies caused by Transport Canada's own behaviours and inactions. The most important fact underlying Transport Canada's general conduct is that it does not set up incentives inducing the new airport operators to promote Transport Canada's interests in day-to-day operations. Corporate governance rules whose intentions are to constrain airports managers' behaviour, do not have strong effects, being perceived by managers as minor irritations that Transport Canada will hesitate to punish.

Agency theory describes two types of principal-agent problems that a principal faces when delegating activities to an agent[3]. The first source of inefficiency is moral hazard or hidden action and the second is known as adverse selection or hidden information. These two problems of opportunism originate from the difficulties that the principal encounters in observing and measuring the anticipated conduct of agents. In the typical hidden action case, which is ex ante opportunism, the principal certainly cannot observe in detail whether the agent is making appropriate decisions. The principal cannot check to see if the actions of the agent are as diligent as they could be. This form of asymmetric information is related to the superior knowledge of the agent. The solution when dealing with moral hazard is to include specific clauses in contracts, reducing its impact.

In the typical hidden information scenario, which is ex post opportunism, the agent has made some observation that the principal has not made. The agent uses this observation in making decisions. However, the principal cannot check whether the agent used his information in the way that best serves the principal's interest. The principal must therefore develop better selection criteria when hiring agents to avoid it.

What follows is a short list containing six types of information asymmetries, the first being four moral hazard cases and the second

group being two adverse selection cases found in the delegation process done by Transport Canada.

First of all, Transport Canada does not implement monitoring expenditures to ensure that airport managers will align their interests on those promoted by Transport Canada. Only on rare occasions does Transport Canada dare to use them. Absence of monitoring induces airport managers to use their natural monopoly power. The introduction of airport improvement fees (AIFs) is a good example[4]. Secondly, the remuneration for senior executives is not clearly determined as there is no provision in the law for such remuneration. By adopting a not-for-profit entity to manage commercial activities, it is plausible that managers, being unable to obtain compensation corresponding to their real productivity, attempt to increase it by various means. A third type of moral hazard lies in the weak influence of important users (air carriers) who do not have a voice on issues, being not represented on the board of directors. Fourthly, many stakeholders are concerned that the not-for-profit status leads to over-investments by managers relatively to the real demand. A fifth form, an adverse selection type, involves the development of "off-airport" activities and ancillary activities. These actions diverge fundamentally from Transport Canada's interests that are to guarantee the delivery of airport services to all Canadian users at a least cost. Sixthly, many financial practices by airports managers show little transparency. Those defaults include a lack of information on the determination of user fees; avoiding to issue calls for tenders; no release of information revealing the name of sole-source suppliers; restrictions on making public information related to the compensation paid to senior executives and no promotion for an independent performance review every five years. It may be argued that those results reflect Transport Canada's lack of experience in developing not-for-profit governance arrangements.

But some important problems still persist between Transport Canada and the members of the National Airports System. The major one on the agenda is the airport rent required by Transport Canada as the caretaker of airport infrastructures. The underlying cause of the

dispute is the current formula used to compute the net revenue for fixing airport leases. The amount of ground lease for an airport has two components, a combination of fixed and variable costs. The relation linking airport rent to net revenue reduces incentives for cost minimization. Gillen et al. explain how the formula limits "the revenues an airport can retain through efficient pricing and entrepreneurial drive"[5]. The consequence for the airport manager is simple: his cost minimization strategies are not optimal and create inefficiency and welfare losses. Gillen et al's solution is that the ground lease "should be collected via a per-passenger fee, reflecting both incentives compatibility and external market conditions"[6]. The issue is more controversial than previously because future forecasts predict that the overall annual rent paid by airport providers will increase at an exponential rate. Moreover, Transport Canada has an additional problem with the Department of Finance: The latter maintains that it owns the ground lease.

The second is the not-for-profit organisation's status for airport providers. A traditional not-for-profit receives revenue primarily from its donors whereas an airport authority produces commercial services that are subsequently sold to the users. Assimilating an airport infrastructure to a not-for-profit entity is nonsense. Not only they do not share the same objective, but their internal structure is totally different. The federal act describes the not-for-profit non-share organization as "a body corporate and politic, without share capital, for the purposes of carrying on, without pecuniary gain to its members, objects,...of a national patriotic, religious, philanthropic, charitable, scientific, artistic, social, professional or sporting character, or like objects"[7].

The *Canada Transportation Act* Review Panel

In its 2001 report, The *Canada Transportation Act* Review Panel advanced suggestions to improve further the performance of airport infrastructure providers. We present the most important suggestions with the motives invoked by the Panel[8]. However, an overview of the main problems generated by Transport Canada governance model is

relevant, although some of them were previously mentioned. The Panel found many weaknesses in governance and monitoring arrangements that will threaten the future efficiency of airports. Some airports use their market power to obtain more revenue than necessary and others use AIFs for any purpose. The checks and balances of the system do not operate well, impeding a convergence towards an operational equilibrium for all the parties involved. Transport Canada believed initially that financial institutions would discipline directly airports by managing their debts. But as airport managers show how efficient they are in covering partly their loans with the AIFs, financial institutions lend them money, being consider a low risk in their portfolio. The governance regime introduced by Transport Canada for Local Airport Authorities (LAAs) and Canadian Airport Authorities (CAAs) that focuses on consultations with users and interested parties and on the accountability of directors to nominating bodies is not very effective as disciplinary mechanisms. Some airports that set up for-profit subsidiaries, some being related to non-airport and others being ancillary activities, can be involved in cross-subsidization, when occurring difficulties. Finally, ground lease, already a problem in itself, limits actual arrangements and affects some important accountability requirements.

Gaps and weaknesses in current institutional arrangements induce the Panel to call into question "whether the not-for-profit model is the optimal approach". After showing that the classical for-profit model has superior advantages over public alternatives[9], the latter being a common property type of resource or signifying a complete absence of ownership rights, the Panel expresses that the differences between airport providers and all those providers of other commercial services are not sufficient to rule out the private firm model for airport infrastructures.

A typical for-profit regime requires that significant public policy objectives can continue to be considered. Safety is an important government responsibility, whatever the type of organization chosen. The potential market power for private infrastructure providers must

be addressed as well. A cost-benefit analysis may reveal whether the costs of government intervention are higher or not than the costs generated by unregulated for-profits airports. The economic literature on public utilities, telecommunications, natural gas and oil pipelines, and the airline industry has shown that regulatory process per se are costly, distort the firms' incentives and cause loss of economic efficiency. Thirdly, adopting the for-profit model will solve a once-and-for-all determination of appropriate compensation for the true value of airports providers investments made over the years. There will be no more contractual disputes between Transport Canada and the 26 airports as seen since the commercialization of airport infrastructures. The Panel's conclusion is to develop "a long term strategy to realize the full benefits of market-oriented service provision"[10].

Very few travellers will dare to challenge that the efficiency of each of the airport providers has improved, when comparing them with the public monopoly managed by a branch of Transport Canada. Each airport infrastructure provider has adopted its own user-pay principles to get the best number of possible passengers. In the meantime, Transport Canada has learned much in its discussions with independent airport authorities.

Some can surely make a case that the learning process of Transport Canada was very slow because nobody in Transport Canada had the necessary know-how to figure out how heavy are the transition costs required to move from a public monopoly entity to a decentralized market composed of individual airports located around Canada. The learning-by-doing process observed in Transport Canada implies that its transaction costs fall with knowledge acquired because its managers grow in skill as their experiences multiply. It is easier now, than previously, for Transport Canada to evaluate how airport entities manage federal assets, get clear responses when points of view are divergent and in the long run, be able to analyze their respective performance. The dialogue between Transport Canada and each airport providers and their Association is more constructive and less acrimonious than before. Both sides understand the content and

the intentions of the delegation process, namely to produce an efficient airport system allowing passengers to move from one place to another at a minimal cost.

3. Bill C-27 *Canada Airports Act*

The proposed *Canada Airports Act*, introduced as Bill C-27 in the House of Commons on March 20, 2003, had not reach the Committee stage when Parliament adjourned in December. The content of Bill C-27 can be considered as a response to the questions and recommendations raised by the Report of the *Canada Transportation Act* Review Panel. A brief presentation of the major changes in the nature of infrastructure helps us to understand the fundamental intentions of Transport Canada policy towards the long term tendencies of airports.

In Part 2 of the Act, Transport Canada retains its non-share capital, not-for-profit corporation status (subsection. 35(1)). Activities of airport authorities are better defined than before (subsection. 50 (1) and 50(3)) as are their investments in other corporations such as subsidiaries (art. 56 and art. 57) that must be incorporated under the *Canada Business Corporations Act*. The unconditional exemption from the *Income Tax Act* is also eliminated in the new Act.

Transport Canada develops in Part 3 a complete framework for the corporate governance of airports. Here are some examples: An airport must still reveal fully any contracts involving expenditures in excess of $100,000 not awarded under a public bid solicitation process (art. 96); Every airport must invite the air carriers operating in the airport to meet with the transport authority at least once a year (art 97); Article 98 requires that the airport authority must provide for mechanisms to consult with their community. Part 5 presents how each airport authority must disclose and be accountable every year for audited financial statements, annual reports and annual meetings open to the public and so on. Part 5 also explains how each airport must establish performance indicators and be ready to accede to a performance review every five years. Part 6 describes a general

structure for the existing, new and revised fees imposed by airport operators, the necessity of transparency when determining the methodology for new fees, obligations for consultation with users and stakeholders and possible appeals to the Canadian Transportation Agency in certain cases. The Auditor General of Canada agrees in her February 2005 Report that the main features of this governance framework are of high quality and must be put in place as soon as possible[11].

Bill C-27 shows that Transport Canada has learned much over years, when discussing, arguing and trading with airport authorities. Corporate governance in general is more inclusive as users, stakeholders, air carriers and community interests are involved. The board, its roles, its rules and duties are more articulated and defined than those actually in place. Consultation with users, air carriers and stakeholders are part of the 60 year contract linking them together. Our main comments on the non adopted Bill C-27 are the following.

Transport Canada does not accept to adopt the private firm model as recommended by The *Canada Transportation Act* Review Panel. Transport Canada does not take the opportunity, as suggested by the Review Panel, to enjoy the benefits of the complete privatization of infrastructure providers as other countries (United Kingdom, Australia and New Zeeland) have done since 1990's.

Transport Canada does not intend to change the complex formula for the ground lease that many persons, academics as well as stakeholders and users, describe as inefficient and inflexible. In the actual regime, airport providers do not have incentives to be cost-minimizing agents. They are not rewarded for their efficiency when increasing their net cash flow. All increases in the net cash flow coming from their economic ingenuity, based on energy, efforts, investments in money and in time, augment at the ground lease paid to Transport Canada.

Their marginal activities to increase the net cash flow are counter productive as they go against their airport's interest. A simple formula

based on passenger traffic as suggested by Gillen et al. is more coherent than the actual one, if the governance regime's objective aims at advancing efficiency in airport operations. Gillen's proposition seems to provide as much revenue as the existing formula and at much less risk than the actual regime generates.

Evidence from the tragic events of September 11th and the outbreak of Severe Acute Respiratory Syndrome (SARS) in Toronto reveals how a lack of flexibility can change rapidly the economic situation faced by airports. A formula of which some of its components are counter-cycling guarantees at least shorter variations in the ground lease than the automatic rent payments actually required by Transport Canada to each airport. This flexible rule is fiscally sound step and a little bit fairer than the actual arrangement.

The style of corporate governance proposed by the Bill C-27 is in general compatible with that advanced by the Review Panel. This approach reinforces the accountability and control that were missing before the Report of the *Canada Transportation Act* Review Panel. Compliance and enforcement are really now part of the contract between both parties Transport Canada and the airports authorities. For instance, members of a board and officers that commit an offence against the Department of Transport is liable to a fine, the amount varying from less than $50,000 for an individual and less than $250,000 for a corporation or an airport operator.

In the meantime, Transport Canada has terminated the review of the National Airports Rent Policy. Transport Canada has not yet given answers to the impact of ground lease on airport system and air carriers, on the fair value received by government for the transfer to airport authorities and on a possible sharing of the ground lease among users, taxpayers and airport authorities if the rent must be reduced.

The Auditor General of Canada writes that "the Department has separated calculating the value of the airports from determining how much rent to charge"[12]. To do so, that Transport Canada must

evaluate the airport's value and other possible land uses of the airport. Then, the landowner computes the discounted cash flow of future revenues generated by the airport activities and those revenues of available land uses[13]. The ground lease paid by the airport must be based on the value generated by its airport activities and not at all by others possible land uses. For instance, the Vancouver airport is so beautifully located that it could generate more revenue to the landowner if he decides to change his use by building outstanding condos, mansions and villas around golf courses.

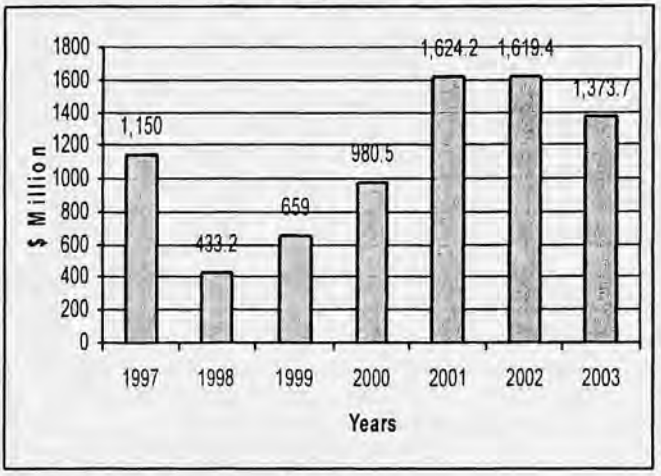
4. The performance of infrastructure providers

We evaluate the performance of the first eight most important airports in Canada from 1997 to 2003. The main elements considered are the level of investments realized to satisfy their consumers, their main sources of revenue, their operational efficiency as cost-minimizing operators, the equity generated by their overall activities, their annual net benefits and their dispute about the rent payment paid to Transport Canada.

The acquisition of capital assets for that particular period is $7.890 billion or around $1.120 billion a year. The largest investments occurred between 2000 and 2003 and they represent 71.4 % of the overall volume of the period. The Greater Toronto Airports Authority (GTAA) is the main source of the boom, having invested $5.204 billion or in percentages, 66.4% of the capital assets acquired by airport providers. The seven remaining airports invested $2.686 billion or 33.6%. Figure 1 presents the evolution of investments for the period by our eight chosen airports. The Auditor General of Canada writes in her recent 2005 February Report that "Transport Canada invested around $1.1 billion during the 1983 to 1992 period when all airport providers invested $6.6 billion from 1999 to 2003"[14]. Although the commercialization of airport providers is a clear success overall, it does not imply that all investments are perceived cost-conscious by users, air carriers and stakeholders. The GTAA is actually considered to be out of step by many national and international associations. The International Air Transport

Association (IATA) wrote in November 2004 about the GTAA situation that "We need an efficient terminal, but we don't need a Taj Mahal, or Versailles, or whatever you want to call it"[15].

Figure 1
Capital Assets Invested by Eight Airports, 1997-2003, in $ million

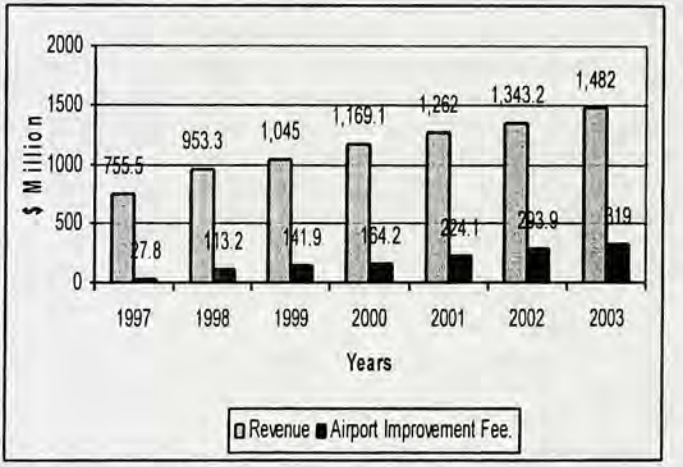


Source: Data come from annual reports

Canadian airports have three types of revenue. Aeronautical revenues are composed of landing fees charged to airline companies and airport fees imposed on passengers. Airport improvement fees are paid by users for boarding an aircraft. Non-aeronautical revenues hail from concessions, property leases and parking fees. Figure 2 shows the growth of total revenues collected by our eight airports and also the importance of the AIFs over the years. Presently airport improvements fees represent 30% of total revenues collected by airport providers because nearly every Canadian airport obliges its travellers to contribute directly to its investments. The AIF depends on the relative market power of each airport. For example, the GTAA has introduced in 2001 a conventional airport improvement fee as all

airports have done since 1993 and a transit fee based on the reasoning that more than 50% of travellers go through Toronto airport when going to their final destination.

Figure 2
The Growth of Total Revenue of Airports and Airport Improvement Fees generated from Eight Airports, 1997-2003, in $ million



Source: Data come from various annual reports.

Airport providers search in general to be cost-minimized when offering services to their customers. But some of them, due to their location, can enjoy relative market power. In so doing, they can charge higher fees for travellers whose demand is inelastic and lower fees for those whose the demand is elastic. Let us use the case of the Greater Toronto Airports Authority as a prominent example of market power, being the largest hub in Canada. Many airports among the 26 airports offer direct services to this airport and vice versa. This relationship between the Pearson hub and the various spokes is cost-conscious if airlines flying to Toronto and back obtain the best services at the lowest costs. The International Air Transport

Association (IATA) that represents 277 airlines, accounting for more than 95% of global scheduled air traffic, writes in November 2004 that "the Pearson Airport has increased its landing fees by 208% already since 1998"[16]. Moreover, an Air Canada representative argues recently that new fee increase of 18% for 2005 simply contribute to the uncompetitiveness of Toronto as a destination. In an intensive survey analyzing aircraft landing fees for a large number of world airports in 2003, Tae Oum[17] reveals that Pearson ranked as the most expensive in North American airport in landing fees in three specific aircraft categories. Globally, Pearson airport ranks second behind Tokyo's Narita Airport as the most expensive airport at which to land. Finally, As Tae Oum's survey also measures "operating efficiency ", he shows that Vancouver Airport is the number 2 performer in North America after Atlanta and that Toronto places 20th in that ranking of more than 40 North American airports.

If air carriers modify their schedule and air fares accordingly and if competing airports reduce their different fees too, they will attract in the short run marginal travellers, and inframarginal customers in the long run who used to transit by the Toronto hub. They will offer value for money in contradistinction to Pearson airport. This substitution will exert pressure on all those dealing with the Toronto airport to change the existing situation and to contract their cost of doing business. So, the GTAA airport actual situation comes from its relative market power, its financial internal policies and its high cost operation. In a word, the relative market power of GTAA generates inefficient behaviours and the resultant costs are passed on to passengers.

Table 1 presents a brief financial comparison of the Vancouver airport and the Pearson airport from 1997 to 2003. The main elements considered are capital assets or investments, collected revenues, the airport's equity as activities go up and the ground lease to Transport Canada.

Table 1
Financial Comparisons of the Vancouver Airport and the Toronto Airport, 1997-2003, in $million

	Vancouver				GTAA			
Years	Capital	Revenue	Equity	Net benefits	Capital	Revenue	Equity	Net benefits
1997	51.0	213.6	239.1	44.1	881.5	274.5	12.1	19.3
1998	73.4	236.7	303.7	64.6	193.8	372.2	42.7	30.5
1999	92.8	245.1	353.1	49.5	367.3	417.2	81.6	40.0
2000	66.6	265.0	410.2	57.2	766.1	449.7	102.1	20.0
2001	98.0	266.9	472.6	62.7	1,186.8	517.3	88.0	-13.7
2002	54.8	248.5	516.9	44.4	1,025.5	577.9	64.0	-20.4
2003	60.8	274.5	540.4	23.5	887.8	646.9	1.2	-66.8

Source: Data come various annual reports

The first and fifth column contain the volume of capital assets that each airport has acquired over the period to fulfill its obligations to its clients and the landlord. The Vancouver airport has invested $497.2 million, a mean of $71.0 million a year and the GTAA airport $5,203.8 million, a mean of 743.1 million per year. The amount of money going to the capital assets by the Vancouver Airport corresponds to 9.6% of that invested by Toronto managers.

The second and sixth columns represent the overall revenues generated by both airports. The Vancouver airport has received $1,750.3 million in revenue relatively to $3,255.7 million for the GTAA during the seven year period. The Vancouver airport revenue is 53.8% of those obtained by the Pearson Airport. The Vancouver airport initiated in 1993 the AIF's while the GTAA introduced its two types of fee only in 2001.

The third and seventh columns show the evolution of equity as both airports increase their revenue to offer better services to their consumers. The Vancouver management steadily increased its equity by a factor of 2.3 over the period, namely from $239.1 million in 1997 to $540.4 million in 2003. In the meantime, the Toronto airport's equity value did not follow the same increasing path. The equity in 1997 was $12.2 million, grew to $102.1 in 2000 and dropped dramatically to $1.2 million in 2003. In a word, many

inefficiencies mentioned in the previous paragraphs are partly confirmed when looking at financial elements contained in Table 1. The erosion of GTAA equity is caused by precipitous and massive investments that those capital assets are now a real menace for many years to come, if for instance, there are sudden increases in the Canadian interest rate caused by external macroeconomic conditions[18].

The annual net benefits, located in the fourth and eighth columns, indicate the balance of revenues over expenses for a year and a positive number shows an increase of equity for the next year. The net benefits generated by the Vancouver airport are always growing as time goes on while the net benefits coming from the GTAA activities have declined steadily since 2000, and were negative for the last three years. The measures of the volume of equity and net benefits are good predictors of a firm's health, whatever its status as a private firm or a non-for-profit entity. Toronto airport must change how it does business by reducing its high costs in order to satisfy its clients. In the meantime, its potential competitors, namely airports being considered as its satellites or spokes, must increase their share of the market by reducing their costs. They have a golden opportunity to attract travellers who have a price-elastic demand.

The last considered analyzed element is the ground lease that each Canadian airport must give back annually Transport Canada for using its land. Table 2 presents, for each column, the respective rent payment by each airport per year, the total amount paid by each airport during a seven year period and the overall annual ground lease by eight airport providers. For example, the Vancouver airport paid $56.4 million in 1997 and $66.3 million in 2003. Its overall contribution for seven years was $424.5 million. Moreover, the annual contribution for each airport for the year 1997 was $181.8 million and $241.0 million in 2003. The total rent payment for the 1997-2003 was $1,512.5 million. Vancouver, Toronto and Calgary airports have paid in 93.8% of the overall ground lease to Transport Canada.

Canadian airport providers argue that they have invested a lot of money to improve and update their respective airports. They are also profitable organizations that generate more equity or net cash flow than before, when airports were a Canadian public monopoly. In so doing, their financial success increases further the value of the land relatively to the previous public regime.

Table 2
The Evolution of Ground Lease for Eight Airports and the Total 1997-2003, in $million

Years	Vancouver	Toronto	Calgary	Edmonton	Halifax	Montreal	Ottawa	Winnipeg	Total
1997	56.4	107.1	18.3	0	0	0	4	1	181.8
1998	43.5	115.8	19.0	0	0	0	5.3	1.4	190.2
1999	60.3	117.9	19.9	0	0	0	5.9	1.9	211.2
2000	63.9	122.3	20.7	0	0	0	6.1	3.4	216.4
2001	67.8	126.4	21.5	0	0	0	8.7	3.6	228.0
2002	66.3	134.5	22.5	0	3.1	6.4	11.0	3.6	244.3
2003	66.3	125.2	23.4	0	3.1	4.8	11.3	3.8	241.0
Total	424.5	849.2	145.2	0	6.2	11.2	52.3	18.7	1,512.5

Source: Data come from various annual reports

Airports want their share of the net cash flow because they are the only temporary owners of this additional value added. Private managers in the post-reform period perform better than public servants did in the former regime because incentives for performance are stronger now that before. The new tenants have financed new investments. They have taken risks and have completely changed the airport's operations. They want to be the only ones to could claim the returns to such efforts.

The Auditor General of Canada[19] points out that if the Review of the National Airports Rent Policy by Transport Canada recommends a reduction in the rent payment, the Airports authorities intend to share the savings with air carriers. The real problem facing Transport Canada is who will be the recipients. Until now, airports, stakeholders and users are expected to receive some price decrease while taxpayers are rarely mentioned.

It can be deduced that airport providers and interest groups related to the airport's activities may enjoy some diminution in their costs. It is plausible that some savings may be passed on to users. Taxpayers, being an ineffective pressure group, may either receive some money back by a tax cut or nothing at all. I do believe that taxpayers will not be among those who will split the ground lease surplus, if there is a real one.

Finally, the Department of Finance does not want at all to change the airport rent scheme. As the Department believes in owning the ground lease, Transport Canada must therefore find funds in its own department to reduce the ground lease. The most important debate about how the rent payment must be distributed is not among users and taxpayers and Transport Canada, but between Transport Canada and the Department of Finance. Forecasts indicate that in 12 years from now, rent payments will be around $1 billion.

5. Concluding Comments

Although the commercialization of airport providers is a big step towards efficiency, our brief survey, mainly financial, demonstrates that there are places where the search for efficiency is inadequate. The GTAA hub has market power that insulates it from operating efficiently, but such power does not mean other airports, many of them being spokes to the Toronto hub, will not pose a threat. Toronto's airport costs have escalated steadily in recent years and if the trend continues, airports at the spokes, who are now being mindful of costs will increase their respective share of the market and they will send a signal to travellers that efficient airport providers outside Toronto offer good services at low costs.

It is true that Transport Canada has put its house in order, being now more involved and effective than before, proactive and playing generally well itsrole of landowner and caretaker when trading with airport providers. Its learning curve is far better than it was at the end of the last century, but there are still problems that must be attenuated over the following years.

Notes and References

[1] The main source of the following paragraphs is Michel Boucher, *Les fournisseurs d'infrastructures de transport nouvellement commercialisés; analyse des principes de gouvernance, d'imputabilité et de performance*, Working Paper prepared for the *Canada Transportation Act* Review, March 2001.

[2] Transport Canada neglected that a not-for profit organization cannot raise equity capital and must rely on fees and debt to finance its activities.

[3] M. C. Jensen and W.H. Meckling (1976), *Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure*, Journal of Financial Economics, vol.3, October , p.305-360.

[4] A plausible explanation is that each airport has a potential market power and that each of them is ready to raise airport improvement fees (AIFs) to guarantee its loans. This action can be also perceived by the banking industry as a bonding expenditure.

[5] David Gillen et al. (2001), *Airport Financing, Costing, Pricing and Performance*, Working Paper prepared for the *Canada Transportation Act* Review, March 2001.

[6] David Gillen et al. (2001), Airport Financing, Costing, Pricing and Performance, op. cit. p.4.

[7] Michel Boucher (2001), *Les fournisseurs d'infrastructures de transport nouvellement commercialisés; analyse des principes de gouvernance, d'imputabilité et de performance*, op. cit. p. 8.

[8] Vision and Balance, Report of the *Canada Transportation Act Review*, Ottawa, June 2001, Ch.9. p. 149-173.

[9] Vision and Balance, Report of the *Canada Transportation Act Review*, op, cit. p. 159.

[10] Vision and Balance, Report of the *Canada Transportation Act Review*, op, cit. p. 161.

[11] Report of the Auditor General of Canada, February 2005, *Transport Canada- Overseeing the National Airports System*, Ch. 2, p. 6-7.

[12] Report of the Auditor General of Canada, February 2005, *Transport Canada- Overseeing the National Airports System*, op. cit., p. 5.

[13] The land used for airport purposes, although it has no other short-term use, has an opportunity cost corresponding to the sum obtained by the owner when selling it. In the long run, all possible changes are possible, if so desired.

[14] Report of the Auditor General of Canada, February 2005, *Transport Canada- Overseeing the National Airports System*, op. cit., p. 5.

[15] *The Globe and Mail*, November, 17, 2004, p. B9.

[16] *The Globe and Mail*, November, 17, 2004, p. B9.

[17] Tae Oum, *The Globe and Mail*, November, 17, 2004, p. B9.

[18] The Auditor General of Canada reveals in its last Report that Transport Canada does not have a strategy for responding in the event that an airport authority encounters serious financial difficulty. Op. cit.p.9 .

[19] Report of the Auditor General of Canada, February 2005, *Transport Canada- Overseeing the National Airports System*, op. cit., p. 5.

Airport Congestion and Market Structure: The Case of Variable Passenger Time Costs

Anming Zhang and Andrew Yuen
Sauder School of Business and Centre for Transportation Studies, University of British Columbia

November 2004

Abstract: Brueckner (2002) develops an analytical framework to show how market structure in aviation markets affects optimal airport congestion pricing. In this paper we modify the Brueckner model such that for a given number of peak flights, time costs resulting from travel in the congested peak period are higher for some passengers than others owing, for example, to their higher time valuation. We find that for each market structure, our modified model tends to predict greater congestion, relative to the social optimum, than Brueckner. As a result, the modified congestion tolls are higher than Brueckner's tolls. Moreover, while Brueckner's prediction – that carriers' internalization of congestion rises as airport concentration rises – still holds in the modified model, airport congestion appears to become less responsive to the change in market structures under the modified model. This may provide an explanation for the inconsistency between Brueckner's prediction and some of the empirical studies – the latter seem to indicate a more limited power of market structure in explaining observed variations of airport delays.

Acknowledgement: Financial support from the Social Science and Humanities Research Council of Canada to Anming Zhang (SSHRC Grant No. 410-2003-1613) is gratefully acknowledged.

Correspondence: Anming Zhang, Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, B.C., V6T 1Z2, Canada. Phone: (604) 822-8420; Fax: (604) 822-9574; E-mail: anming.zhang@sauder.ubc.ca

1. Introduction

Congestion and air traffic delays at busy U.S. airports have become a major public policy issue in the last few years. In 2000, the Aviation Investment and Reform Act for the Twenty-first Century (AIR-21), which contains provisions aimed at stimulating competition and economic development and easing congestion at busy airports, was signed into law. The Airport Congestion Relief Act of 2001 further amended AIR-21 by temporarily limiting the number of airline flights at La Guardia Airport, and the Secretary of Transportation, in consultation with the Port Authorities of New York and New Jersey, continued a comprehensive study on possible remedies to alleviate the congestion.[1] Congestion is perhaps even more acute at some of the major European and Japanese airports.

Since the early work of Levine (1969), Carlin and Park (1970) and Borins (1978), economists have approached airport congestion by calling for the use of the price mechanism, under which landing fees are based on an aircraft's contribution to congestion. The approach involves charging different landing fees at different times: during peak hours, flights are charged higher rates than during off-peak hours.[2] Congestion pricing is desirable from an efficiency point of view, because social marginal cost varies depending on the time the runway and slots are used and congestion pricing better reflects this varied marginal cost than a flat fee structure.[3] Three empirical models of airport congestion pricing have

---

[1] http://www.ncsl.org/statefed/AIR21cb.htm

[2] Landing fee is the main component of airport pricing. At Hong Kong International Airport, for example, airport charges consist of three components: landing, aircraft parking, and terminal building charges (TBC). Landing fees are based on aircraft movements, while TBCs are passenger based. In a "typical" case (a B747 with 418 seats, 66.8% load factor and 4-hour turnover) it has been calculated that the landing fee accounts for 68.7% of the total charge, parking 6.6% and TBC 24.7%.

[3] A less mentioned benefit of congestion pricing is that it may facilitate airlines' entry onto routes from congested airports. Dresner, Windle and Yao (2002) find that among the airport barriers they

been developed, namely, the standard peak-load pricing model (Morrison 1983; Morrison and Winston 1989), the deterministic bottleneck model (Vickrey 1969; Arnott, De Palma and Lindsey 1993), and the stochastic bottleneck model (Daniel 1995, 2001). Daniel and Pahwa (2000) compare the three models by using the data from Minneapolis-St. Paul (MSP) airport. In their simulation results, the traffic rates, delay levels and (positive) congestion tolls are on average remarkably similar. They argue that the stochastic bottleneck model could predict the behavioral response of air carriers to airport congestion pricing; consequently it could properly estimate congestion fees. Raffarin (2004) argues that introducing congestion costs in the European air traffic control (ATC) pricing rule is an efficient means to improve the social surplus and tackle congestion in European airspace. The congestion pricing proposed by economists has not really been implemented, and the existing landing fees are based primarily on aircraft weight.[4]

As pointed out by Brueckner (2002), most studies in the above airport congestion literature extend the road-pricing literature without taking account of important differences between the two markets. Road users are atomistic and so have no market power but air carriers may have market power at a given airport. Brueckner demonstrates how congestion tolls are modified when the economic agents causing congestion have market power. Specifically, when an airport is dominated by a monopolist, congestion is fully internalized,

---

analyzed, gate utilization during peak periods is the most significant deterrent to entry in the U.S. airline industry.

[4] Apart from the price mechanism and direct administrative allocations mentioned above, other remedies for airport congestion focus mainly on expanding physical capacity, including increasing the size of airports, building new runways, and improving the air traffic control technology, all of which can be time-consuming, expensive undertakings. On the other hand, Daniel (1995) shows that congestion pricing, while virtually free to implement, may achieve the same goal through an improvement in capacity utilization. Based on an analysis of MSP airport, he concludes that existing capacity could serve 30% more traffic under congestion pricing.

yielding no role for congestion pricing. Under a Cournot oligopoly, carriers internalize the congestion they impose on themselves; as a result, the congestion externality cost borne by other carriers will need to be priced. This means a positive congestion fee, but the fee is still lower than the fees calculated in, e.g., Daniel and Pahwa (2000). Brueckner's analysis therefore suggests that earlier research may have overestimated the appropriate congestion fees, owing to the latter's assumption of atomistic competition.[5]

In Brueckner's analysis, he makes an assumption that for a given number of peak flights, time costs resulting from travel in the congested peak period are the same for all (peak) travelers. As pointed out by Brueckner, this assumption may be unrealistic. In particular, the higher time valuation of business passengers would suggest that, rather than depending on the number of peak flights alone, the time cost should be an increasing function of passengers' time valuation. This observation is supported by extensive literature on the value of travel time savings, which indicates that air travelers value their travel time in proportion to their wage rate (see Section 2 below for individual studies). In this paper we formally incorporate Brueckner's idea of variable passenger time costs, and examine how this modification affects the extent of internalization and the level of congestion tolls under various market structures.

Our second objective is to shed some light on the significance of market structure (airport concentration, or degree of competition) in explaining airport congestion. Brueckner (2002) predicts that carriers' internalization of congestion rises as airport concentration rises. He further tests this hypothesis using cross-

[5] One limitation of Brueckner (2002) is that it assumes that a congested airport is connected to another uncongested endpoint. To overcome this limitation, Brueckner (2003) analyzes internalization and congestion pricing in a network setting, which includes multiple congested airports serving as network hubs. It is found that the earlier results remain valid in the network setting. In a somewhat simpler model, Pels and Verhoef (2004) obtain similar conclusions that airlines may internalize some delay and so that portion of the externality cost should not be accounted for in congestion tolls.

section data from the 25 most delayed U.S. airports in 1999, and finds some rudimentary supporting evidence. On the other hand, Morrison and Winston (2003) do not find that air carriers internalize congestion as airport concentration rises, when they analyzed a larger sample of 78 U.S. airports accounting for 67% of all domestic airline operations in 2000. Daniels (1995), in his study of MSP airport dominated by the hub airline Northwest, also rejects internalization in favor of atomistic behavior. One very interesting and insightful explanation for the inconsistency of the analytical and empirical results is provided by Mayer and Sinai (2003). While Brueckner has abstracted the potential network benefits of hubbing in his analysis, Mayer and Sinai argue that delays at airports dominated by one large carrier are the equilibrium outcome of the hub airline equating marginal benefits from hubbing with the marginal cost of delays. Hubbing leads to narrower, more congested traffic banks. Their data on U.S. domestic flights by major U.S. carriers from 1988-2000 show that delays were increasing in hubbing activity at an airport and decreasing in airport concentration.[6] Of the two factors, however, the empirical impact of airport concentration, which they use as a proxy for the extent to which delay costs are internalized by the carriers at the airport, is much more modest than for hubbing. They thus argue that an optimal congestion toll would have a relatively small impact on air traffic delays since hub carriers already internalize most of the costs of hubbing, and a toll that does not take the network benefits of hubbing into account could reduce social welfare.

Our framework follows Brueckner (2002); in addition, we conduct a simulation study of the analytical model. We find that for each market structure, our modified model tends to predict greater airport congestion, relative to the social optimum, than Brueckner. As a result, the modified congestion tolls are higher than Brueckner's tolls. Furthermore, while Brueckner's prediction – that carriers' internalization of congestion rises as airport concentration rises –

[6] The data, obtained from the U.S. Department of Transportation, include flights from 1988-2000 by all air carriers with more than 1% U.S. market share, with a total of over 66 million flights covering more than 250 airports.

continues to hold in the modified model, our analysis reveals that congestion would become less responsive to the change in market structures under the modified model than in the Brueckner model. This "softened" effect of the market-structure variable may provide an alternative explanation for the inconsistency between Brueckner's prediction and some of the empirical results, which seem to indicate a more limited power of market structure in explaining observed variations of airport delays.

While focusing on carriers' internalization incentive and its implications for optimal congestion tolls, Brueckner (2002) also suggests that an airport might be underused when strong market power exists, an effect he refers to as the "residual market power." Pels and Verhoef (2004) explicitly consider such market-power distortions and their impacts on the optimal congestion toll. They consider situations where airports need not cooperate optimally to maximize joint welfare, but instead may engage in a form of tax competition in the pursuit of maximizing local rather than global welfare. Their analysis shows that the second-best optimal congestion tolls are typically lower than what would be suggested by congestion costs alone and may be even negative. We shall also discuss the implication of the modified model for market-power distortions. Our analysis shows that the modification might alleviate the residual market-power effect.

2. Basic Model and Modification

2.1. The Brueckner Model

This section describes the basic model of Brueckner (2002). The model distinguishes between two travel periods, peak and off-peak, at a given airport. The peak period consists of a relatively short time window containing the day's most desirable travel times (e.g. early morning or late afternoon) and is always congested. On the other hand, the off-peak congestion never occurs. Passengers are represented by a continuum with index θ, which is uniformly distributed between 0 and 1 with unit density.

Travel benefits depend on θ and the travel periods. The benefits from off-peak travel are given by the function $b_o(\theta)$, which are independent of n_o – the number of flights in the off-peak period – because of the absence of any congestion in the off-peak period. However, the benefits from peak travel depend on both θ and n_p, which denotes the number of flights in the peak period. So the peak-travel benefits are given by $B_p(\theta, n_p)$. This benefit function is further assumed to be additively separable in its arguments:

$$B_p(\theta, n_p) = b_p(\theta) - t(n_p), \qquad (1)$$

where $b_p(\theta)$ represents the gross benefit of peak travel and $t(n_p)$ denotes the additional passenger time cost due to congestion, which is increasing in n_p.

Potential passengers are sorted such that $b_p^{'}, b_o^{'} > 0$, i.e., the (gross) travel benefits are higher for a high-θ passenger than a low-θ passenger. As such, θ may be interpreted as an index of the passenger's tendency to travel on business (as opposed to leisure travel), for business travel is a crucial job requirement. Moreover, since business travel must occur during the early and late peak hours to avoid disruption of the work day, peak-travel benefits should increase relative to off-peak benefits as θ increases, yielding $b_p^{'} > b_o^{'}$ for all $\theta \in [0,1]$. This is referred to as the single-crossing assumption. To avoid a degenerate equilibrium, it is further assumed that the levels of the benefit functions must be such that they intersect at an intermediate value of θ. Consequently, $b_p(\theta) > (<) b_o(\theta)$ hold for high (low, respectively) θ, indicating that peak benefits are higher (lower, respectively) than off-peak benefits for business (leisure, respectively) passengers.

On the production cost side, c is the operating cost per flight, and $g(n_p)$ is the additional operating cost due to peak-period congestion, with $g(\cdot)$ increasing in n_p. Letting s be the number of seats per flight and assuming, for simplicity, a 100% load factor, the welfare function at the social optimum can be written as:

$$W = \int_{\underline{\theta}}^{\theta^*} b_o(\theta)d\theta + \int_{\theta^*}^{1}[b_p(\theta) - t((1-\theta^*)/s)]d\theta - c(1-\underline{\theta})/s \quad (2)$$
$$-[(1-\theta^*)/s][g((1-\theta^*)/s)]$$

where those travelers with θ higher than θ^* choose to travel in the peak period, those with θ between $\underline{\theta}$ and θ^* travel off-peak, and those with θ below $\underline{\theta}$ do not travel. Assuming an interior solution for θ^*, then the solution, denoted θ_s^*, is characterized by the first-order condition (subscript s for social optimum), (3)

$$[b_p(\theta_s^*) - t(n_p) - b_o(\theta_s^*)] - (1/s)(1-\theta_s^*)t'(n_p) - (1/s)$$
$$[g(n_p) + n_p g'(n_p)] = 0$$

where $n_p = (1-\theta_s^*)/s$, and by the second-order condition $W_{\theta^*\theta^*} < 0$ (here subscripts denote partial derivatives). Since the left-hand side of (3) equals $-W_{\theta^*}$, this expression is, by the second-order condition, increasing in θ^*. To ensure the second-order condition for optimization, it is assumed that both $t(\cdot)$ and $g(\cdot)$ are convex functions.

2.2. Modification: Variable vs. Constant Passenger Time Costs

A key implication of the separability condition (1) is that passenger time costs do not depend on θ. Specifically, for a given n_p, time costs resulting from peak travel are the same for all peak travelers. Brueckner (2002, p.1361) points out: "It should be noted that, under

the business-leisure interpretation of the passenger continuum, the separability assumption for peak benefits may be unrealistic. In particular, the higher time valuation of business passengers would suggest that, rather than depending on n_p alone, time cost t should be an increasing function of θ." Indeed, von Wartburg and Waters (2004) have a detailed discussion on the relationship between the value of travel time savings (VTTS) and income. Their comprehensive survey of the relevant studies (e.g. MVA Consultancy et al. 1987; AHGC 1999; Gunn et al. 1999; Wardman 2001; and Mackie et al. 2001) indicates that the income elasticity of VTTS is positive. For example, Mackie et al. (2001) report this income elasticity is in the range of 0.72-0.82. Furthermore, positive estimates of VTTS for air travel time have been incorporated in many studies (e.g. Carlton, Landes and Posner 1980).

In this paper, we formally incorporate Brueckner's idea of variable time costs into his analysis by providing a modification to the peak benefit function, $B_p(\theta, n_p)$. We examine how the original results change under the more realistic assumption. To distinguish between the original model and this paper, we shall use "hat" to represent variables of this paper; so, for example, the time cost is now denoted as $\hat{t}(\theta, n_P)$. We shall assume

$$\hat{t}(\theta, n_p) = H(\theta) \cdot t(n_p).$$

Hence, the peak benefit function is now given by:

$$\hat{B}(\theta, n_p) = b_p(\theta) - H(\theta) \cdot t(n_p), \qquad (4)$$

where $H(\cdot)$ is an increasing function shown in Figure 1. For a given number of peak flights, therefore, travelers with high θ will suffer more due to congestion than travelers with low θ. Notice that Brueckner's analysis is a special case with $H(\theta) = 1$ for all $\theta \in [0,1]$. For the purpose of comparison between the original and modified analyses, we shall assume that $H(\hat{\theta}_s^*) = 1$, where $\hat{\theta}_s^*$ is

the socially optimal θ^* under the modified model, and that $\int_0^1 H(\theta)d\theta = 1$.

**** Figure 1 Here ***

Similar to Brueckner (2002), assumptions are made so as to avoid the degeneration of equilibrium. In particular, to ensure that there are passengers in the peak and off-peak periods, $b_p(\theta) > b_o(\theta)$ must hold for high θ and $\hat{B}(\theta, n_p) = b_p(\theta) - H(\theta)t(n_p) \le b_o(\theta)$ must hold for low θ, respectively. It is also assumed that $b_p'(\theta) - H'(\theta)t(n_p) - H(\theta)t'(n_p) > b_o'(\theta)$.

3. Analysis

3.1. The Social Optimum

With the modification, the welfare function at the social optimum can be written as: (5)

$$W = \int_{\underline{\hat{\theta}}}^{\hat{\theta}^*} b_o(\theta)d\theta + \int_{\hat{\theta}^*}^{1} [b_p(\theta) - H(\theta)t((1-\hat{\theta}^*)/s)]d\theta - c(1-\underline{\hat{\theta}})/s$$

$$-[(1-\hat{\theta}^*)/s][g((1-\hat{\theta}^*)/s)].$$

Differentiating (5) with respect to $\underline{\hat{\theta}}$, the first-order condition reduces to:

$$b_o(\underline{\hat{\theta}}_s) \ge c/s, \quad (6)$$

with equality holding if $\underline{\hat{\theta}} > 0$. This result is the same as the one obtained by Brueckner, indicating that the modification does not alter

$\hat{\underline{\theta}}$. Intuitively, in the allocation between travelers and non-travelers we only need to consider the operating cost and off-peak travel benefits, which do not involve the peak benefit function upon which our modification was imposed.

More interest, therefore, lies in the impact of the modification on the choice of θ^*, which divides passengers between peak and off-peak travel. Using the assumption that $H(\hat{\theta}_s^*) = 1$, we derive the following first-order condition for θ^*:

$$[b_p(\hat{\theta}_s^*) - t(\hat{n}_p) - b_o(\hat{\theta}_s^*)] - (1/s)\int_{\hat{\theta}_s^*}^{1} H(\theta)d\theta \cdot t'(\hat{n}_p) - (1/s)[g(\hat{n}_p) + \hat{n}_p g'(\hat{n}_p)] = 0 \quad (7)$$

where $\hat{n}_p = (1 - \hat{\theta}_s^*)/s$. The items in the first bracket on the left-hand side of (7) give the change in travel benefits for the marginal traveler, namely, $\hat{\theta}_s^*$, shifting from the off-peak to peak periods. The second term in (7) captures the congestion effect of the marginal traveler: to serve this passenger, the number of peak flights should be increased by $1/s$. This term represents total additional time costs to the peak passengers. The last term in (7) is the additional operating cost of moving the traveler from the off-peak to peak periods. Thus, equation (7) shows that the marginal travel benefit is set equal to the incremental congestion and operating costs of switching a passenger from the off-peak to peak periods.

Given that $H(\hat{\theta}_s^*) = 1$ and $H(\cdot)$ is an increasing function, it follows that $H(\theta) > 1$ for all $\theta > \hat{\theta}_s^*$, thus yielding $\int_{\hat{\theta}_s^*}^{1} H(\theta)d\theta > \int_{\hat{\theta}_s^*}^{1} 1 d\theta = 1 - \hat{\theta}_s^*$. Using this inequality, $H(\hat{\theta}_s^*) = 1$, and (7), we then obtain:

$$[b_p(\hat{\theta}_s^*) - t(\hat{n}_p) - b_o(\hat{\theta}_s^*)] - (1/s)(1-\hat{\theta}_s^*)t'(\hat{n}_p)$$
$$-(1/s)[g(\hat{n}_p) + \hat{n}_p g'(\hat{n}_p)] > 0$$

or equivalently, $-W_{\theta^*}(\hat{\theta}_s^*) > 0$. Since equation (3) says $-W_{\theta^*}(\theta_s^*) = 0$ and the expression $-W_{\theta^*}$ is increasing in θ^* by the second-order condition, it follows immediately that $\hat{\theta}_s^* > \theta_s^*$, leading to:

Proposition 1: *At the social optimum,* $\hat{\theta}_s^* > \theta_s^*$, i.e., *there should be fewer flights in the peak period under the modified model than under the Brueckner model.*

Thus, the modified model changes the social-welfare benchmark: it reduces the congested travel segment from $[\theta_s^*, 1]$ to $[\hat{\theta}_s^*, 1]$. Essentially, the additional time costs due to congestion for peak passengers, who have higher θ's, rise under the modified model. The first-order condition for the choice of θ_s^* balances the benefits and costs of switching the marginal traveler from the off-peak to peak periods. To avoid the higher passenger time costs, fewer flights in the peak period should now be arranged.

3.2. The Competitive Case

In the competitive case, the operating costs determine fares as well as allocation of passengers:

$$f_P = [c + g(n_P)]/s, \quad (8)$$

$$f_o = c/s, \quad (9)$$

where f_P and f_o are the fares in the peak and off-peak periods, respectively. As in the social optimum case, the modification does not alter the allocation between travelers and non-travelers as compared to the original analysis. Here, $b_o(\hat{\underline{\theta}}_c) - f_o \geq 0$, where subscript c stands for the competitive case.

For the allocation between peak and off-peak passengers, the traveler with $\hat{\theta}_c^*$ will be indifferent between traveling in the peak or off-peak periods. In particular, $b_p(\hat{\theta}_c^*) - H(\hat{\theta}_c^*)t(n_p) - f_p = b_o(\hat{\theta}_c^*) - f_o$. By substituting equations (8) and (9), the condition determining $\hat{\theta}_c^*$ reduces to:

$$[b_p(\hat{\theta}_c^*) - H(\hat{\theta}_c^*)t(\hat{n}_p) - b_o(\hat{\theta}_c^*)] - (1/s)g(\hat{n}_p) = 0 \qquad (10)$$

where $\hat{n}_p = (1 - \hat{\theta}_c^*)/s$. Equation (10) implies that $\hat{\theta}_c^* < \hat{\theta}_s^*$: Consistent with the result in Brueckner (2002), competitive carriers fail to internalize the congestion effect, leading to overuse of the peak period. Nevertheless, as $H(\cdot)$ is an increasing function and $H(\hat{\theta}_s^*) = 1$, that $\hat{\theta}_c^* < \hat{\theta}_s^*$ implies $H(\hat{\theta}_c^*) < 1$. Compared with the corresponding first-order condition in Brueckner, namely, $[b_p(\theta_c^*) - t(n_p) - b_o(\theta_c^*)] - (1/s)g(n_p) = 0$, we obtain

$$\hat{\theta}_c^* < \theta_c^*. \qquad (11)$$

That is, more flights will be arranged in the peak period under the modified model than in Brueckner (2002). Intuitively, in the competitive case, each airline only equates its marginal benefit with its marginal cost, without taking into account the congestion costs it imposes on other airlines. Under the modified model, the time cost of the marginal traveler will be less than that in the original model. As a consequence, airlines are more willing to provide flights in the congested peak period of the airport.

This overuse of the peak period may be avoided by imposing a congestion toll. While no toll is charged for off-peak flights, the peak-period congestion toll equals the congestion cost imposed by an additional flight and it is given by:

$$\hat{R}_c(\hat{n}_p) = \int_{\hat{\theta}_c^*}^{1} H(\theta)d\theta \cdot t'(\hat{n}_p) + \hat{n}_p g'(\hat{n}_p) \qquad (12)$$

where $\hat{n}_p = (1-\hat{\theta}_c^*)/s$. When the carriers are charged $\hat{R}_c(\hat{n}_p)$ per peak flight, the peak fare, f_P in (8), is augmented to include the additional expression $\hat{R}_c(\hat{n}_p)/s$. Equation (10) is then modified to include the missing congestion terms involving $t'(\cdot)$ and $g'(\cdot)$, and the new equilibrium reduces to the social optimum.[7]

Proposition 2: *In the competitive case,*
(i) the congestion problem is worsened,
(ii) the congestion toll is greater,
under the modified model than under the Brueckner model.

As can be seen from the proof, the congestion is worsened owing to two sources. First, the modified model implies less use of the peak period from the social point of view than the original model (Proposition 1). Second, as indicated in (11), the modified model predicts greater use of the peak period by competitive carriers than the original model. The situation is illustrated in Figure 2. Note that the worsened congestion can also be seen from the optimal congestion tolls in the two cases.
**** Figure 2 Here ****

3.3. Monopoly

As a perfectly discriminating monopolist solves the same optimization problem as the social planner, the modification would not change the result, shown in Brueckner (2002), that the social optimum would be replicated under conditions of a perfectly discriminating monopolist. In the following, we will examine the situation where the monopolist cannot discriminate and so must charge uniform peak and off-peak fares. Hence, the monopolist solves the following profit-maximization problem: (13)

[7] The proofs of the propositions have been omitted due to space constraints and are available upon request from the authors.

$$\underset{\hat{\theta}^*,\hat{\underline{\theta}}}{Max}[b_p(\hat{\theta}^*) - H(\hat{\theta}^*)t(n_p) - b_o(\hat{\theta}^*)](1-\hat{\theta}^*) + b_o(\hat{\underline{\theta}})(1-\hat{\underline{\theta}})$$

$$- c(n_p + n_o) - n_p g(n_p).$$

As in the cases of social optimum and competitive firms, the modification does not alter the allocation between travelers and non-travelers. The first-order condition for choice of $\hat{\underline{\theta}}$ is (subscript *m* for "monopoly"):

$$b_o(\hat{\underline{\theta}}_m) - (1-\hat{\underline{\theta}}_m)b_o'(\hat{\underline{\theta}}_m) \geq c/s,$$

which shares the same expression as the one in Brueckner (2002).

With regards to the allocation between peak and off-peak travelers, differentiating expression (13) with respect to θ_m^* yields:

$$[b_p(\hat{\theta}_m^*) - H(\hat{\theta}_m^*)t(\hat{n}_p) - b_o(\hat{\theta}_m^*)] - \hat{n}_p H(\hat{\theta}_m^*)t'(\hat{n}_p)$$

$$+ (1-\hat{\theta}_m^*)H'(\hat{\theta}_m^*)t(\hat{n}_p) - (1/s)[g(\hat{n}_p) + \hat{n}_p g'(\hat{n}_p)] \qquad (14)$$

$$- (1-\hat{\theta}_m^*)[b_p'(\hat{\theta}_m^*) - b_o'(\hat{\theta}_m^*)] = 0,$$

where $\hat{n}_p = (1-\hat{\theta}_m^*)/s$. As compared to equation (7), the term $(1/s)\int_{\hat{\theta}_s^*}^{1} H(\theta)d\theta \cdot t'(n_p)$ is now replaced in (14) by $n_p H(\hat{\theta}_m^*)t'(n_p) - (1-\hat{\theta}_m^*)H'(\hat{\theta}_m^*)t(n_p)$. As $H(\cdot)$ is an increasing function, the former term is greater than the latter. Hence, ignoring the last term in (14), which captures the "residual market-power effect," we obtain

$$\hat{\theta}_m^* < \hat{\theta}_s^*. \qquad (15)$$

In other words, the monopolist does not fully internalize the congestion cost. This is in contrast to Brueckner's Proposition 1, which states that a non-discriminating monopolist fully internalizes the congestion cost. Intuitively, as the non-discriminating monopolist

can only charge a single price for peak-hour passengers, lower-θ passengers in the peak period implicitly should subsidize higher-θ passengers; as a consequence, the monopolist does not fully internalize the congestion cost. This overuse problem may be solved by imposing a congestion toll. The congestion toll on each additional peak flight should be set equal to the additional cost it imposes on other peak flights, giving rise to:

$$\hat{R}_m(\hat{n}_p) = \int_{\hat{\theta}_m^*}^{1} H(\theta)d\theta \cdot t'(\hat{n}_p) - s[\hat{n}_p H(\hat{\theta}_s^*)t'(\hat{n}_p) - s\hat{n}_p H'(\hat{\theta}_s^*)t(\hat{n}_p)] \quad (16)$$

Notice that the last term in (14), namely, $-(1-\theta_m^*)[b_p'(\theta_m^*) - b_o'(\theta_m^*)]$, is negative given the single-crossing assumption $b_p' > b_o'$. This implies that the monopolist allocates too few passengers to the peak period, relative to the social optimum, and is referred to as the "residual market power effect". As shown in the following result, this residual market-power effect might become smaller under the modified model than in the original model.

Proposition 3: *(i) Absent of the residual market-power effect, a non-discriminating monopolist will only partially internalize the congestion cost; (ii) If $H'(\theta)$ is sufficiently small at $\hat{\theta}_m^*$ and assuming $(1-\theta)[b_p'(\theta) - b_o'(\theta)]$ is decreasing in θ, then underuse of the peak period due to the residual market-power effect is less serious under the modified model than in the Brueckner model.*

In the above result the assumption, that $(1-\theta)[b_p'(\theta) - b_o'(\theta)]$ is decreasing in θ, holds if benefit functions $b_p(\theta)$ and $b_o(\theta)$ are linear, which will for example be assumed in the simulation study of Section 4). Also note that combining the two results in Proposition 3 would yield the result that

the tendency here is towards overuse of the peak period, rather than underuse as predicted by the Brueckner model.[8]

4. Simulation Results

This section conducts a simulation analysis for the allocation of passengers as well as congestion pricing under various market structures. We consider symmetric airlines, and use specific functional forms and related parameters described as follows.

Passengers: Following Pels and Verhoef (2004) we consider linear benefit functions:[9]

$$b_o(\theta) = \alpha_o + \gamma_o\theta, \qquad b_p(\theta) = \alpha_p + \gamma_p\theta$$

where $\alpha_o, \gamma_o, \alpha_p, \gamma_p > 0$, $\alpha_p > \alpha_o$ and $\gamma_p > \gamma_o$. Furthermore, both $H(\theta)$ and $t(n_p)$ are linear:

$$H(\theta) = \lambda\theta, \qquad t(n_p) = sn_p t = (1-\theta^*)t$$

where $\lambda > 0$. Hence, $B_p(\theta, n_p)$ is given as:

$$B_p(\theta) = b_p(\theta) - H(\theta)t(n_p) = \alpha_p + \gamma_p\theta - \lambda\theta(1-\theta^*)t.$$

Airlines: We consider linear operating cost functions and linear $g(n_p)$, which implies constant marginal congestion cost. The cost functions are therefore given as follows:

$$g(n_p) = \varphi s n_p = \varphi(1-\theta^*)$$

$$c(\theta, n_p) = c_0(1-\underline{\theta})/s + [(1-\theta^*)/s]g(n_p) = c_0(1-\underline{\theta})/s + [(1-\theta^*)/s]\varphi(1-\theta^*)$$

where $\varphi > 0$.

---

[8] The analysis for Cournot Oligopoly has been omitted due to space constraints and is available upon request from the authors.

[9] It is noted that Pels and Verhoef used linear demand functions. To be consistent with our analytical model, benefit functions are used here.

Choice of Parameters: Table 1 shows the parameterization for our simulation exercise. Some of the parameters are derived from Pels and Verhoef (2004). According to their demand specification, $\theta \in [0,5000]$, suggesting that the air travel market consists of 5,000 potential travelers. As Pels and Verhoef do not consider the benefit function in the off-peak period, we assume that $\alpha_o = 0$, and $\gamma_o = 0.8$. Moreover, λ is determined by the assumption that $H(\hat{\theta}_s^*) = 1$.

**** Table 1 Here ****

Tables 2 and 3 present the simulation results. The results are consistent with our analytical results. Consider Table 2. First, $\underline{\theta}$ (the allocation between travelers and non-travelers) remains, as expected, the same for each market structure between this paper and Brueckner (2002). Second, as predicted by Proposition 1, the socially optimal θ^* (the allocation between peak and off-peak travelers) is greater under the modified model than in Brueckner. Third, θ^* falls as airport concentration falls, suggesting that airport delays rise as concentration falls, a result that is consistent with Brueckner. Third, whilst the θ^*'s are smaller under oligopoly (and perfect competition) than the social optimum, indicating overuse of the peak period, the monopolist would under use the peak period. These reveal that under the given simulation specifications, the residual market-power effect dominates the congestion externality effect under monopoly, while the reverse holds for oligopoly. Fourth, for each market structure, the modified model tends to predict greater congestion, relative to the social optimum, than Brueckner. In particular, as predicted by Proposition 2, the congestion problem is worsened in the competitive case.

The last point may also be seen from Table 3, where congestion tolls are computed for each market structure. Table 3 shows that the modified congestion tolls are higher than Brueckner's tolls for each and every market structure. The percentage increase

ranges from ∞ under monopoly conditions, 61% under duopoly, 37% under trigopoly, to 12% under perfect competition. These results are consistent with our analytical results.
**** Table 2 Here ****
**** Table 3 Here ****

Figure 3 plots the congestion toll against the market structure variable k, the number of airlines at the airport. The figure supports the argument made by Brueckner (and this paper) that congestion is negatively correlated to airport concentration. Furthermore, the marginal increase in optimal congestion toll falls as k rises. That is, although airport congestion is worsened owing to more serious externalities as the number of carriers rises, the marginal external effect falls. In particular, Figure 3 shows that the level of congestion toll remains almost unchanged for $k > 5$ suggesting, therefore, that for the practical purpose of congestion pricing we might just treat $k > 5$ as atomistic airlines and apply the results of the road-pricing literature. Finally, the figure shows that the optimal congestion toll increases in k but at a lower rate under the modified model; as a consequence, our congestion-toll curve is flatter than Brueckner's curve. Thus, airport congestion would be less responsive to the change in market structures under the modified model than in the Brueckner model. This may provide an explanation for the discrepancy between Brueckner's prediction and some of the empirical results, which seem to indicate a more limited power for market structure as an explanatory variable of actual variations of airport delays.
**** Figure 3 Here ****

5. Concluding Remarks

Our main objectives in writing this article are to better understand the implications of incorporating variable passenger time costs into Brueckner's analysis for airport congestion and congestion pricing, and to shed some light on the significance of market structure in explaining airport congestion. In our analytical and simulation analysis, we found that for each market structure our modified model

tends to predict greater congestion, relative to the social optimum, than Brueckner. In particular, the congestion problem is worsened in the competitive and monopoly cases. The modified congestion tolls are higher than Brueckner's tolls for each and every market structure. While Brueckner's prediction that carriers' internalization of congestion rises as airport concentration rises still holds in the modified model, our analysis further reveals that airport congestion would become less responsive to the change in market structures. This may provide an explanation for the discrepancy between Brueckner's prediction and some of the empirical results, which seem to indicate a more limited power of market structure in explaining observed variations of airport delays. Our analysis also shows that the modification might alleviate the residual market-power effect.

Figure 1. Variable Passenger Time Costs: The H(θ) Function

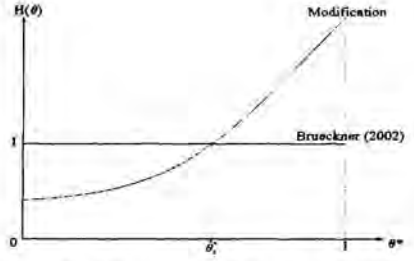


Figure 2. Allocation of Peak and Off-peak Travelers: The Social Optimum and Competitive Market

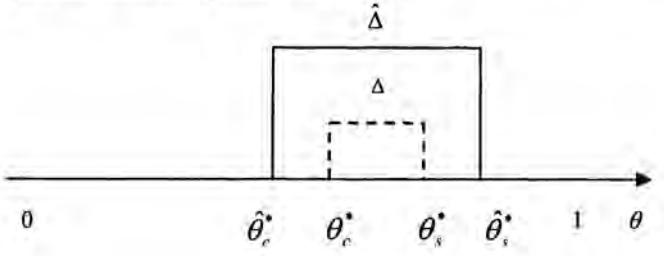


Figure 3. Market Structure and Congestion Tolls

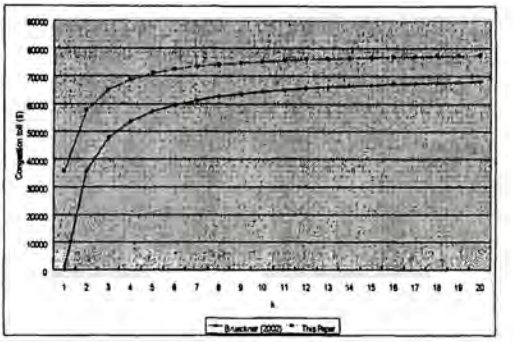


Table 1. Parameter Values for the Numerical Model

Benefit functions		Airline cost functions	
α_o	0	c_0	100,000
γ_o	0.8	s	200
α_p	0	φ	38.08073115
γ_p	1		
t	0.25		
λ	0.000279697		

Table 2. Simulation Results of the Allocation of Passengers

	Social Planner		Monopoly		Duopoly		Trigopoly		Perfect Competition	
			k=1		k=2		k=3		k=∞	
	$\underline{\theta}$	θ^*	$\underline{\theta}$	θ^*	$\underline{\theta}$	θ^*	$\underline{\theta}$	θ^*	$\underline{\theta}$	θ^*
Brueckner (2002)	625	3,575.3	2,812.5	3,920.1	2,083.3	3,517.2	1,718.8	3,320.2	625	2,782.5
With Modification	625	3,690.8	2,812.5	3,831.9	2,083.3	3,299.4	1,718.8	3,007.6	625	2,157.7

Table 3. Simulation Results of Congestion Tolls

	Monopoly	Duopoly	Trigopoly	Perfect Competition
	k=1	k=2	k=3	k=∞
Brueckner (2002)	0	35,753.2	47,670.8	71,506.4
With Modification	35,954.6	57,632.4	65,107.6	79,808.6
% increase	∞	61.2%	36.6%	11.6%

References

AHCG Accent Marketing & Research and Hague Consulting Group, "The Value of Travel Time on UK Roads," The Hague, Netherlands, 1999.

Arnott, R., A. De Palma and R. Lindsey, "Economics of a Bottleneck," Journal of Urban Economics 27 (1993) 111-130.

Borins, S.F., "Pricing and Investment in a Transportation Network: The Case of Toronto Airport," Canadian Journal of Economics 11 (1978) 680-700.

Brueckner, J.K., "Airport Congestion When Carriers Have Market Power," American Economic Review 92 (2002) 1357-1375.

Brueckner, J.K., "Internalization of Airport Congestion: A Network Analysis," Working Paper, Institute of Government and Public Affairs, University of Illinois at Urbana-Champaign, 2003.

Carlin, A. and R.E. Park, "Marginal Cost Pricing of Airport Runway Capacity," American Economic Review 60 (1970) 310-319.

Carlton, D.W., W.M. Landes and R.A. Posner, "Benefits and Costs of Airline Mergers: A Case Study," Bell Journal of Economics 11 (1980) 65-83.

Daniel, J.I., "Congestion Pricing and Capacity of Large Hub Airports: A Bottleneck Model with Stochastic Queues," Econometrica 63 (1995) 327-370.

Daniel, J.I., "Distributional Consequences of Airport Congestion Pricing," Journal of Urban Economics 50 (2001) 230-258.

Daniel, J.I. and M. Pahwa, "Comparison of Three Empirical Models of Airport Congestion Pricing," Journal of Urban Economics 47 (2000) 1-38.

Dresner, M., R. Windle and Y. Yao, "Airport Barriers to Entry in the U.S.," Journal of Transport Economics and Policy 36 (2002) 389-405.

Gunn, H., J.G. Tuinenga, Y.H.F. Cheung and H.J. Kleijn, "Value of Dutch Travel Time Savings in 1997," in H. Meersman, E. Van de Voorde and W. Winkelmans (eds.), Proceedings of the 8th World Conference on Transport Research, Vol. 3, Pergamon: Amsterdam, 1999.

Levine, M.E., "Landing Fees and the Airport Congestion Problem," Journal of Law and Economics 12 (1969) 79-108.

Mackie, P.J., M. Wardman, A.S. Fowkes, G. Whelan, J. Nellthorp and J. Bates, "Values of Travel Time Savings in the UK: A Report on the Evidence," ITS Working Paper No. 567, Institute of Transport Studies, University of Leeds, 2001.

Mayer, C. and T. Sinai, "Network Effects, Congestion Externalities, and Air Traffic Delays: Or Why Not All Delays Are Evil," American Economic Review 93 (2003) 1194-1215.

Morrison, S.A., "Estimation of Long-run Prices and Investment Levels for Airport Runways," Research in Transportation Economics 1 (1983) 103-130.

Morrison, S.A. and C. Winston, "Enhancing the Performance of the Deregulated Air Transportation System," Brookings Papers on Economic Activity, Microeconomics (1989) Special Issue, 61-123.

Morrison, S.A. and C. Winston, "Air Travel Delays: Another Look at Appropriate Policy," Unpublished Manuscript, Northeastern University, 2003.

MVA Consultancy, Institute of Transport Studies, University of Leeds, and Transport Studies Unit, University of Oxford, "The Value of Travel Time Savings," Policy Journals, Newbury, Berks, U.K., 1987.

Pels, E. and E.T. Verhoef, "The Economics of Airport Congestion Pricing," Journal of Urban Economics 55 (2004) 257-277.

Raffarin, M., "Congestion in European Airspace – A Pricing Solution?", Journal of Transport Economics and Policy 38 (2004) 109-126.

Vickrey, W., "Congestion Theory and Transport Investment," American Economic Review 59 (1969) 251-260.

von Wartburg, M. and W.G. Waters II, "Congestion Externalities and the Value of Travel Time Savings," Chapter 2, Towards Estimating the Social and Environmental Costs of Transportation in Canada, A Research Report prepared for Transport Canada, pp. 54-140, 2004.

Wardman, M., "A Review of British Evidence on Time and Service Quality Valuations," Transportation Research Part E 37 (2001) 107-128.

National Road Network (NRN)

(Canada's digital road centerline framework data)



www.GeoBase.ca

Marcel Sabourin
msabouri@nrcan.gc.ca
Natural Resources Canada
Centre for Topographic Information
2144 King Street West
Sherbrooke, QC J1J 1E8

Abstract

Natural Resources Canada (NRCan) plays a lead national role in the acquisition, management and dissemination of GIS data associated with ground transportation networks. Road, railway and pipeline networks are key geospatial layers that serve as an essential source of information to support decision-making in a wide array of applications.

Where agreements are in place, NRCan along with its network of stakeholders have initiated and are now maintaining a good portion of Canada's 1.1 million kilometres of digital road centerline data referred to as the National Road Network (NRN). This initiative has succeeded in bringing multilevel governments (federal, provincial, territorial and municipal) to a common understanding (model) of a Canadian road network framework dataset.

NRCan has invested a significant amount of effort over the last five years in modeling acquiring and updating the National Road Network. The NRN provides a consistent geometric description and a set of 'core' attributes for all publicly accessible roads. The modeling work for the NRN was primarily based on Linear Reference System (LRS) concepts and complies, to the greatest degree possible, with internationally recognized Geographic Data Files Committee (GDF / ISO TC-204) vocabulary and definitions.

Together NRCan and provincial Stakeholders participate in the updates of the road network and maintain it in a sustainable manner.

As an End User it is important for you to know that NRN centerline data are currently available at no cost and without restriction through the GeoBase portal (www.geobase.ca).

The paper and presentation will provide a brief historical overview of the NRN acquisition project as well as a high level review of the model. Information regarding the current status and strategic maintenance of the NRN will also be presented.

Introduction

In the past, organisations involved in the acquisition, management and dissemination of road centerline data have done so primarily in response for their specific needs and with little regards for a homogeneous standardized National acquisition method. Emerging technology, increase use of International (ISO) and open (OGC) standards now allow these same departments to review their current practices in an attempt to redefine past habits and convert to a National Standard.

Across all sectors, it is becoming increasingly recognized that basic geographic information serves as a direct input to several aspects of sustainable development, logistics planning, emergency preparedness, public policy, health studies, statistical studies, wildlife management, etc...

The rapid development and widespread proliferation of distributed databases and the Internet have only increased the demand for access to a variety of geographic data, including standardized data produced by government. User applications are becoming more sophisticated and users themselves more familiar with geospatial data therefore more demanding in regards to the types of data that serves as input.

Many data clients point to the lack of a more integrated homogeneous and standardized framework set of data. The Canadian Council on Geomatics (CCOG)[1] decided to tackle this issue. In 1999 the CCOG mandated the Ministry of Natural Resources Ontario (OMNR) and Natural Resources Canada (NRCan) to work together in researching and defining a National Standard for Canada's road centerline framework now referred to as the National Road Network.

As a result, users with an interest in the field of geomatics now have access to quality geospatial information via the GeoBase portal (www.geobase.ca). One of the layers within the portal is the NRN centerline data. GeoBase is a federal, provincial and territorial

governments initiative that is overseen by the Canadian Council on Geomatics (CCOG). The GeoBase vision is based on the premises of ensuring the provision of, and access to, a common, up-to-date and maintained base of quality geospatial data for all of Canada. Through the GeoBase portal, users have access to quality geospatial data at no cost and with unrestricted use.

NRN model development

Between 1999 and 2001 many meetings between OMNR and NRCan as well as several National Workshop sessions (which involved provincial, territorial and federal representatives) were held to achieve National consensus on defining and adopting the NRN V1.0 data model and data dictionary by all stakeholders involved.

During the development stage, NRCan reviewed and studied many Road Network models. The Geographic Data File (GDF) [2] Standard Version 4.0 of the ISO TC 204 group was considered the best choice and was adopted as a guide in the creation of the NRN. GDF is an application independent data model. The GDF Standard was designed in Europe for traffic and transport related extensions. The model is proven to be flexible, and extendible.

Whenever possible the NRN data dictionary terms and definitions have been derived from the GDF Feature and Attribute Catalogues.

The resulting NRN V1.0 model is also based on NCHRP 20-27 [3] Linear Referencing System (LRS) concepts. This approach allows for the management of geometric representation separately from the attributive information (referred to as "event" in LRS). However it is important to note that in order to satisfy the greatest number of road network data users, NRN data has also been modeled and will be distributed as a conventional (segmented at attributes) road network.

Version 1.0 of the NRN model [4] is currently available via the GeoBase portal. NRCan is currently leading the work for V2.0 of the model.

NRN Data

The primary data source of NRN road centerline data was produced with field driven Differential Global Positioning System (DGPS) technology. Additional sources, such as existing accurate photogrammetric provincial and municipal data were also integrated and updated. During the initial acquisition of the NRN data, efforts were made to utilize and update as much existing authoritative 'closest to source' centerline road data as was possible.

Natural Resources Canada (NRCan) managed and produced, in partnership with several provinces, the first version of the NRN. Initial data collection of the NRN began in the summer of 1999 and was recently completed in March 2005.

NRN Features

The NRN is composed of three features:

Road Element: A road is a linear Section of the earth that is designed for or the result of vehicular movement. A Road Element is the representation of a road between Junctions. Two Junctions always bound a Road Element.

Ferry Connection: The average route a ferryboat takes when transporting vehicles between two fixed locations on the Road Network. Two Junctions always bound a Ferry Connection.

Junction: A point feature that is always connected to one or more Road Element or Ferry Connection. A Junction is defined at the intersection of three or more roads, at the intersection of a Road Element and a Ferry Connection, at the end of a dead end road and at the intersection of a Road Element or Ferry Connection with a National, Provincial, or Territorial Boundary.

NRN Core attributes (or events)

The following list contains 'core' (mandatory) attributes that have been defined and adopted by national consensus (federal, provincial, territorial, municipal and institutional representatives).

Core Attributes for Road Element:
Route Name
Route Number
Road Class
Number of Lanes
Structure
Toll Point
Blocked Passage
Road Surface
Exit Number
Core Attributes for Ferry Connection:
Route Name
Route Number
Road Class
Core Attributes for Junctions;
Junction Type
Exit Number

Agreements reached between NRCan and provincial producers are to ensure that the following 'core' NRN geometry (features) and attributes are maintained. The GeoBase portal distributes these 'core' data at no cost and without restrictions.

The NRN model extends beyond this set of 'core' attributes. NRCan is involved with federal and provincial stakeholders for the population of Addressing data within the NRN. This information however is not available to the public.

For a complete list of all nationally defined attributes refer to the following site:

ftp://ftp.cits.rncan.gc.ca/pub/information/NRN_V2_Draft/Draft_C_August_25_2004/

NRN Geometry

Two Geometric Types are used to represent NRN entities.

- The Road Element and Ferry Connection are represented by linear geometry referred to as GB_Curve as stipulated in the ISO 19107 [5] standard.

- A point, GB_Point that also respects the same standard, represents the entity Junction.

Metadata

Metadata are information that describe data quality and origin. NRN is produced in cooperation with several partners. Consequently, metadata are important and must be associated with every single part of the data that comes from different partners.

Two types of metadata are provided, one for the dataset and one for the object. **Dataset metadata** contains the information that is common to the entire dataset. **Object metadata** qualifies specifics about each object.

For additional information regarding Metadata for the NRN refer to the following site:

http://www.geobase.ca/doc/specs/pdf/GeoBase_metadata_NVD.pdf

NRN data views

NRN data is distributed following two data views:

- Linear Referenced System (LRS) view (Spring 2005)
- Segmented view

The Network Linear Elements forms the geometric structure of the Road Network. In accordance with the LRS model these elements are segmented for two reasons only:

- Intersection at the same level with another Network Linear Element
- Intersection at Provincial or Territorial boundaries.

In addition to these previous segmentation rules, the Segmented view of the Network Linear Elements segments the geometry at each attribute value change.

Spatial Constraints

While the NRN has no explicit topology, it does ensure that the data are free from any spatial inconsistencies such as overshoots and undershoots. The NRN data is 2-Dimension (X,Y) data.

The NRN data is seamless within datasets and forms a continuous network over the Canadian landmass. Segmentation only occurs at Provincial and Territorial limits (called Dataset Boundaries).

Unique Identifiers (UUID)

Identifiers (IDs) must be assigned in a permanent fashion. In order to ensure their stability, the assigned IDs must be insignificant (inconsequential) in their expression. In other words, IDs must not contain any information relative to the data. Previous experiences have demonstrated that encapsulating information within the ID can

force a modification to IDs without any real change having occurred to the data.

Several Standards have dealt with the Road Network. Most of them invoke the importance of the use of Identifiers without ever specifying in any particular manner the format nor the method of application. The following documents, GDF, GIS-T (GIS in transportation data standards) [6] and CEN TC 278 have no specifications related to Identifiers. NSDI Framework Transportation Identification Standard was the only document that clearly defined and described an Identifier code [7]. Within the ISO TC 211/SC: Geographic information standard - Encoding [8] we found a UUID definition that matched our requirements.

An application domain defines a universe and an identification scheme called universal unique identifiers (UUIDs). A UUID is assigned to an object when it is created and is stable over the entire life span of the objects. A UUID of a deleted object cannot be used again. UUIDs are required for long term distributed data management and for realizing update mechanisms. These identifiers are also called persistent identifiers. A special name server may be used to resolve persistent identifiers. The identifiers are unique within a well-defined limited universe defined by an application domain.

One of the most important features of the NRN data is that each *Road Element, Junction and Ferry Connection as well as* each *Event (Attribute)* have a unique National Identifier (NID) which data type is a Universally Unique identifier (UUID). The UUID consists of a record of 16 bytes and must not contain padding between fields. The total size is 128 bits. For use in human readable text, a UUID string representation (32 characters) is specified as a sequence of fields. The following string is a UUID example:

378a3917e824422cb25f268b8295da51

For more information:
http://www.opengroup.org/onlinepubs/9629399/apdxa.htm#tagcjh_20

Planimetric precision

The accuracy of geometric representation data is given as the difference between the position of the geometric representation associated with an object and the real ground position of the corresponding topographic feature, as measured with respect to the geodetic network. Accuracy, which can vary from one occurrence of an object to the next, is provided as an attribute to each object.

Many techniques can be used for acquisition or maintenance of the NRN data. However, the NRN aims at attaining, for all occurrences in all areas, a planimetric (X,Y) data accuracy of 10 meters or better.

Planimetric data accuracy is expressed as the circular map accuracy standard (CMAS) of a topographic feature derived from the equation below:

Standard circular error: $\sigma_c = 0.7071\ (\sigma_x^2 + \sigma_y^2)^{1/2}$

σ_x : standard deviation in the X-axis

σ_y : standard deviation in the Y-axis

Circular map accuracy standard: $CMAS = 2.1460\ \sigma_c$

Coordinate reference system

Horizontal reference system:

NRN spatial data are expressed in reference to the North American Datum 1983 Canadian Spatial Reference System (NAD83CSRS).

Horizontal coordinate system:

NRN data are stored in latitude (ϕ) and longitude (λ) geographic coordinates.

Unit of measure (coordinate system axis units):

The measuring unit for storing horizontal spatial data is the degree, given to 7 significant decimals. Coordinates are expressed as real values. The

number of significant decimals corresponds to the decimetre when projected in UTM coordinates.

Vertical reference system:

The initial NRN data are 2D (X,Y). There is no elevation (Z) associated with NRN data.

Current Coverage and updates

By visiting the GeoBase portal (http://www.geobase.ca/) on a regular basis, users will be kept informed of GeoBase-NRN updates availability. The maintenance of the National Road Network data is funded through negotiated joint federal/provincial/territorial agreements. The data maintenance process has already begun within certain regions. Partnership agreements are securing 'closest to the source' maintenance principles of GeoBase.

The following table contains the basic statistics regarding the validity dates and the number of km per region

Province/Territory	Release Date	Validity Date	Number of kilometres
AB*	2004-06	2003-09	226,030
BC*	2005-03	n/a	n/a
MA	2004-07	2003-09	87,818
NB	2005-03	2003-06	31,830
NL*	2003-01	2001-10	20,563
NS*	2003-01	2001-11	28,304
ON*	2003-11	2002-08	191,371
PEI*	2004-09	2003-11	6,381
QC	2003-11	2002-11	148,108
SK*	2004-07	2004-03	250,062
NU	2003-01	1999-11	308
YUK	2004-09	2003-08	5,849
NWT	2004-09	2003-08	5,068

*Negotiated and signed maintenance agreements for the NRN between NRCan and its provincial counterparts.

Address information

The major advancements in the NRN V2.0 model are the profound modifications that were introduced for the addressing component of the model.

The model was compartmentalized into eight (8) components:
Road Network Features, Linear Events, Point Events, Address Range Events, Street and Place Names Tables, Point Location Events, Linear referencing System Method and Geometry ISO 19107 Profile.

Update management

NRN data constitutes the best representation of the real world phenomena of interest, unless proven otherwise. The geometric data in the NRN must be subject to a measured amount of change. Changes occur when a new information source offers a better representation than the preceding one. A buffered tolerance is used to compare and consider if sufficient geometric change has occurred.

Updating makes it possible to establish a parallel between existing data and the new data from an update. The latter results in certain effects on the data. These effects can be classified as:

Addition
When a new object has no geometric counterpart in the NRN, a new object is added, with a new unique National Identifier (NID).

Retired
When an object no longer represents a feature, the object is retired.

Modification
An object is said to be modified if one of its descriptive attributes or its geometric representation is different. Two types of modification are possible.

Descriptive Modification
A descriptive difference occurs when a pair of objects from the same class is geometrically identical but has different attribute values. For example, the surface type of a specific road may have changed over time from unpaved to paved.

Geometric Modification
A geometric difference occurs when a pair of objects from the same class has different geometries but depicts the same phenomenon.

Confirmation
Alongside change, there is confirmation of objects when their geometric and descriptive attributes have not been modified.

Conclusion

Now that the initial build (acquisition) phase is complete across the country the next challenge will be for NRCan and its partner's to maintain the NRN on a regular basis. NRN project members are meeting with provincial, territorial and federal representatives on an ongoing basis. Discussions regarding capacity, understanding and, of course, resource commitments are the major topics.

The stakeholders involved in this new challenge must also strive to ensure that NRN road centerline data collection is done in a homogeneous fashion and that the database holdings for these are kept synchronized.

The long-term goal for the NRN project will of course be the establishment a network of distributed databases allowing users easy access to standardized NRN data via the GeoBase portal.

Acknowledgements

The author wishes to acknowledge the contribution of NRCan's Centre for Topographic Information (CTI) Sherbrooke management for allowing Denis Boutin (co-author of the NRN) and myself the time needed to properly reflect on the development of the NRN Model & Standards. Also as important are the NRN team members, Denis Bergeron, Guy Bergeron, Pierre Desjardins and René Pellerin in their efforts to work together in achieving national success in this project.

This acknowledgement is also extended to all provincial, federal, municipal, territorial and institutional contributors in their efforts at making this project a national success.

Additional NRN documentation

NRN Product Specifications in PDF format:
http://www.geobase.ca/doc/specs/pdf/GeoBase_product_specs_NRNC1.pdf
NRN Data Model in PDF format:
http://www.geobase.ca/doc/specs/pdf/GeoBase_Data_Models_NRNC1.pdf
NRN Segmented Conceptual data model in PDF format:
http://www.geobase.ca/doc/specs/pdf/GeoBase_Conceptual_Model_SEG_NRNC1.pdf
NRN Linear referenced Conceptual data model in PDF format:
http://www.geobase.ca/doc/specs/pdf/GeoBase_Conceptual_Model_LRS_NRNC1.pdf
NRN Segmented data Feature catalogue in PDF format:
http://www.geobase.ca/doc/specs/pdf/Geobase_Feature_Catalogue_SEG_NRNC1.pdf
NRN LRS data Feature catalogue in PDF format:
http://www.geobase.ca/doc/specs/pdf/Geobase_Feature_Catalogue_LRS_NRNC1.pdf
NRN dataset downloads:
http://www.geobase.ca/geobase/en/list.jsp?produit=nrnc1

References

[1] Canadian Council on Geomatics: http://www.geobase.ca/geobase/en/about/ccog.html

[2] ISO 14825:2004(E), Intelligent transport systems, Geographic Data Files (GDF), Overall data specification, 620 p..

[3] National Cooperative Highway Research Program, "*A Generic Data Model for Linear Referencing Systems*", Research Results Digest, Number 218, September 1997, 28 p.

[4] National Road Network, Linear Referenced Conceptual data model; http://www.geobase.ca/doc/specs/pdf/GeoBase_Conceptual_Model_LRS_NRNC1.pdf

[5] ISO 19107, '*Geographic information spatial schema*', 2003-05.

[6] Dueker, Kenneth J. and Butler, J. Allison, "*GIS-T Enterprise Data Model* with Suggested Implementation Choices",Center for Urban Studies School of Urban and Public Affairs Portland State University, October 1, 1997.

[7] National Spatial Data Infrastructure, "*NSDI FRAMEWORK TRANSPORTATION IDENTIFICATION STANDARD -- Public Review Draft*", FGDC-STD-999.1-2000, Ground Transportation Subcommittee Federal Geographic Data Committee, December, 2000.

[8] ISO Technical Committee 211, Working group 4, "*Geographic Information – Encoding*", ISO/CD 19118.3, June 15, 2001.

Reviving the Activity Space: Can "Old" Behavioural Geography Provide "New" Insights to the Urban Transport Policy Debate?

Ron N. Buliung
Joint Program in Transportation, University of Toronto, 35 St. George Street, Toronto, ON, CAN, M5S1A4, E-mail: ron.buliung@utoronto.ca

Abstract. The intrinsic complexity of cities demands a multifaceted approach to addressing urban issues. Useful concepts and methods for urban analysis have emerged from decades of theoretical and empirical research in urban and behavioural geography. As an example, consider the *action space* and *activity space* concepts, introduced to the literature during the mid-1960s and early 1970s. The *action space* contains the totality of environmental information possessed by an individual, while the *activity space* contains a subset of locations visited during participation in daily activities. In this paper, I initially argue for a refinement of the activity space concept to include additional descriptive properties that have potential use for the evaluation of transport and land use policy effects. I then demonstrate the policy assessment potential of the activity space by exploring the relationship between transit provision and the dispersion of household activity locations. The findings lend preliminary support to "New Urbanist" claims regarding the provision of transit as an approach to reducing auto-dependence in North American cities.

1 Introduction

Urban population growth, the geographical distribution of opportunities, and household demand for goods and services combine to place mounting pressures on cities. Credible evidence suggests that North American cities are faced with several challenges including urban sprawl, land conversion, auto-dependence, congestion, and related environmental and human health effects. An important aspect

of addressing these challenges involves understanding the manner in which households use cities when conducting daily activities and travel. Advancing our understanding of household activity-travel behaviour can inform the development of urban policy that encourages the use of local opportunities, potentially leading to reductions in motorized travel.

In this paper I revisit the *activity space* concept, originally conceived by behavioural geographers during the 1970s. The early literature is reviewed, followed by an examination of recent innovations stimulated by the availability of detailed activity-travel survey data and commercial GIS software. I then argue for the refinement of the activity space to include several properties that may hold value for evaluating urban transport and land use policy. Following the conceptual discussion, I develop an applied example to explore the relationship between the dispersion of household activity locations and proximity to transit. The example demonstrates the potential of the activity space as a concept for evaluating the ability of transit-oriented development (TOD) to achieve desirable outcomes. The paper concludes with a general summary of research findings.

2 Background

Beginning in the mid-1960s, a group of geographers became interested in the conceptual and empirical study of individual and household spatial behaviour. Processes of interest included, migration and residential location choice (Brown and Moore, 1970; Johnston, 1972; Wolpert, 1965), and relationships between urban spatial structure and individual behaviour (Horton and Reynolds, 1971).

Several interesting conceptualizations of urban spatial behaviour emerged from this early research. For example, partially influenced by the Lewinian concept of *life space*[1], Wolpert (1965) devised the *action space* as part of a *field theoretic* approach to describing migration search behaviour. Wolpert (1965) defined the *action space*

[1] Lewin conceived of the "life space" as the individual and his/her psychological environment, often using topological constructs to aid with conceptual illustration (Lewin, 1951).

as the subjective space containing locations about which individuals have sufficient information to elicit move/stay responses. Horton and Reynolds (1971) would extend Wolpert's work in two ways. First, they contemplated the discrete and continuous properties of action spaces. Specifically, they suggested that locations about which an individual has some familiarity could be described in terms of their spatial extent. Using Wolpert's characterization of *place utility* within the action space, Horton and Reynolds then suggested that the action space could have a continuous representation as a utility surface. Second, they introduced the *activity space* to describe the subset of locations within the action space that an individual contacts during daily activity participation. Similar to the action space, the activity space possesses discrete and continuous characteristics.

These concepts would be applied and adapted in other research during the early 1970s. For example, Brown and Moore's (1970) paper on intra-urban migration introduced the *awareness space* as a concept to describe locations that comprise an individual's knowledge of spatial opportunities within the urban environment. They also devised the *indirect contact space* to describe the set of locations that individuals become familiar with through secondary contacts (e.g., acquaintances, mass media, etc.). Similar to the features differentiating action and activity spaces, a distinction would continue to be made in the literature between *knowledge* of the urban environment and *actual* daily contact with spatial opportunities.

Seemingly unrelated initiatives also emerged during the late 1970s and early 1980s to describe the spatial distribution of locations physically contacted by households and individuals during activity participation. Using data from a 1968 travel survey for Washington DC, Zahavi (1979) and Beckman et al. (1983) explored relationships between urban form and *travel probability fields*. Operationally, the travel probability field is a *standard deviation ellipse*[2] estimated using traffic zone centroids that are meant to characterize household and activity locations.

---

[2]Computational formulae can be found elsewhere (Ebdon, 1988; Lefever, 1926; Levine, 2002; Yuill, 1971)

There appears to be renewed interest in concepts that emerged from Behavioural Geography prior to the 1990s. Partially influenced by the arrival of GIS software, the period since the early 1990s has been one of considerable geocomputational innovation. The availability of activity-travel micro-data, coupled with the data management and analytic capabilities of GIS, has facilitated a shift in the scale of analysis. Instead of conducting research at the zonal scale, researchers have achieved success with the conceptual and empirical description of individual and household level activity-travel behaviour.

Two parallel themes have emerged in recent literature. The first concerns implementation, extension, and refinement of *time geographic* concepts, such as the *potential path area*. Using complex algorithms implemented using GIS, this literature typically attempts to describe *potential* individual activity-travel behaviour in space and time (e.g., Miller, 1991; Kim and Kwan, 2003; Kwan, 1999; Scott, 2003). The second theme involves the use of spatial statistics and computational geometry to describe primarily the spatial characteristics of individual and household *activity* and *action spaces*. A few examples associated with this second theme are examined more closely because they relate well with the objectives of this paper.

Dijst (1999) developed linear, circular, and elliptical *actual action spaces* for a sample of individuals of two Dutch municipalities. Newsome et al. (1998) discuss the *observed activity space* as a spatial entity that can describe the area within which individuals are likely to engage in regular activities. Schönfelder and Axhausen (2003) develop three spatial abstractions for the *activity space* using confidence ellipses, kernel density estimation, and shortest path networks. The last approach attempts to capture the subset of network segments that likely comprise an individual's daily path through physically contacted activity locations. Buliung and Kanaroglou (2005a) have conceptualized the household activity space as the convex hull of the set of activities physically contacted by a household's individual members during a single day. What these studies share is the common goal of developing geometric spatial

abstractions to describe potential and observed activity-travel behaviour.

While most the research discussed in this section has been interesting and innovative from behavioural and computational perspectives, widespread adoption of related concepts for urban policy assessment has been slow to develop. Recent examples include an evaluation of the effect of two-earner households on sustainable transport (Dijst, 1999), and a study of the relationship between the spatial extent of the household activity space and a multi-centred metropolitan development strategy (Buliung and Kanaroglou, 2005b). These and other recent examples, (e.g., Newsome et al., 1998; Schönfelder and Axhausen, 2003), have demonstrated that the activity space and related concepts hold potential value for evaluating urban policy effects.

3 Refining the Activity Space

Horton and Reynolds (1971) initial conceptualization of the activity space includes four properties that can be used for descriptive purposes. These are, (1) person - i.e., the activity space belongs to an individual, (2) activity locations, (3) spatial extent, and (4) intensity of activity participation across the broader action space. Despite recent experimentation with procedures for estimating the spatial extent of the activity space (e.g., Buliung and Kanaroglou, 2005a; Newsome et al., 1998; Schönfelder and Axhausen, 2003), there has been no formal attempt to *refine* the activity space concept.

Partially informed by recent research, I propose a *refinement* of the activity space originally devised by Horton and Reynolds (1971). Within this context, *refinement* refers to the process of developing a more detailed specification of something that has been previously specified at a more general level (Rumbaugh et al., 1999). The approach to conceptual modelling that I have adopted has been influenced by the literature on object-oriented analysis and design (e.g., Booch, 1994). To this end, I have adopted the Unified Modelling Language (UML) to aid with the visual expression of the conceptual modelling exercise (e.g., Booch, 1994; Buliung and Kanaroglou, 2004).

The primary objective of the refinement process is to broaden the list of characteristics that can potentially be used to describe activity spaces. This process has been documented in figure 1, where *ActivitySpace_71* refers to the original concept, and *ActivitySpace_05* identifies the new specification. Original properties have been supplemented with several new ones including, (1) Household, (2) Dispersion, (3) TravelDistance, (4) TravelTime, and (5) DateTime. These properties are essentially state variables that can be used to describe the spatial structure and travel demand associated with an activity space over time. Implementation within a broader activity-travel database (e.g., Buliung and Kanaroglou, 2004) could also be used to link activity space objects with individual and household sociodemographics. The remainder of this section elaborates on the definition of these additional properties.

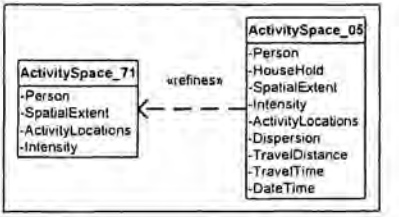


Figure 1. Refining the activity space.

The *Household* property facilitates a logical connection between the activity spaces of individuals within the same household. The property also suggests that activity spaces can be estimated at the household level. For example, the union of the sets of activity locations, physically contacted by members of the same household, can be used to operationally define the *household activity space* (e.g., Buliung and Kanaroglou, 2005a).

The *Dispersion* property is meant to hold a value to describe the spatial distribution of activity locations within individual or household activity spaces. Estimation can be achieved using *centrographic statistics* such as standard distance and the standard

deviation ellipse (e.g., Bachi, 1963; Buliung and Kanaroglou, 2005a; Ebdon, 1988; Lefever, 1926; Levine, 2002).

Total travel costs through the set of locations within an activity space can be stored in the *TravelDistance* and *TravelTime* properties. These can be determined at the individual level, or for an entire household by summing across the travel costs associated with the activity spaces of individual household members. Numerous approaches are available for estimating travel costs. These range in complexity from the construction of activity tours using Euclidean segments to connect successive activity locations (e.g., Buliung and Kanaroglou, 2005b), to the use of shortest path algorithms implemented using various link impedance measures (e.g., distance, free-flow cost, congested travel times).

A unique aspect of the proposed refinement involves the use of a *DateTime* property to embed the activity space concept within a discrete time domain. Conceivably, the spatial and non-spatial properties of the activity space will change over time in response to individual and household lifecycle, and adjustments in the urban environment (e.g., transport capacity, metropolitan development). Furthermore, an individual's knowledge of the urban environment likely evolves on a daily basis, combining with lifecycle and urban environmental factors to produce *daily activity spaces*. Through a process of temporal aggregation – i.e., the creation of monthly or yearly activity spaces using locations contacted over these longer time periods, we can extract activity spaces that represent activity-travel behaviour over a particular time interval. As recent studies have shown, activity spaces estimated using daily or weekly observations hold potential value for policy assessment (e.g., Buliung and Kanaroglou, 2005b; Schönfelder and Axhausen, 2003). Clearly, one of the challenges to operationalizing these concepts within a single study involves the collection of detailed longitudinal activity-travel survey data.

4 Policy Application

In this section, I examine the relationship between the weekday household activity space and proximity to transit stops. This represents an initial step toward using the activity space for

evaluating the success of transit-oriented development (TOD) in achieving desirable outcomes. For example, there is an expectation that the provision of transit, complemented by compact mixed-use development, will elicit a consumer response that includes selection of activity locations closer to home (Nelson and Niles, 1999; Porter, 1997; Tunlin and Millard-Ball, 2003). A behavioural response of this sort could give rise to smaller activity spaces containing geographically concentrated activity locations.

Study Area and Behavioural Observations

Households used in this research have their primary residence in the Portland, Oregon metropolitan area. With a resident population of approximately 529,121, the city of Portland lies at the centre of the study area. Portland has maintained a vital commercial role in Oregon since the middle of the 19th century and continues to be an important regional commercial and cultural hub. The broader metropolitan region is managed by Metro, the only elected regional government in the US. Metro's jurisdiction includes approximately 1.3 million residents, the urbanized portions of three counties (Washington, Clackamas, and Multnomah), and 24 cities. Metro has long been recognized as an early adopter of an integrated approach to the development of transportation and land use policy.

Behavioural observations have been drawn from Metro's 1994-1995 "Household Activity and Travel Behavior Survey". The survey was conducted in the Portland Six-County region during the spring and autumn of 1994 and the winter of 1995. Activity-travel observations for two-consecutive days of the week were recorded for 9,471 persons from 4,451 households. Overall, respondents recorded 122,348 activities and 67,891 trips (Cambridge Systematics, 1996). Activities include those occurring inside and outside of the home. Activity locations were geocoded using intersection or street addresses with the approach influenced by the quality of respondent recorded location descriptions.

The Household Activity Space and Proximity to Public Transit

Following the proposed conceptual refinement of the activity space, I operationally define the *household activity space* to be a spatial entity comprised of the set of locations physically contacted by the

individual members of a single household over time[3]. The remainder of this section investigates the sensitivity of the dispersion of activities within the household activity space to the placement of transit stops within the study area. The sample used in this study includes the unique weekday observations of 3,649 households from the Portland survey. The reduction in sample size from the original survey occurred because certain households were assigned to weekend days, or had incomplete activity descriptions. Exclusion of weekend observations is based on evidence suggesting a weekday/weekend differential in activity-travel behaviour (e.g., Bhat and Srinivasan, 2004).

The dispersion property of the household activity space has been operationalized using the *standard deviation ellipse*. The level of dispersion associated with a set of activity locations increases with the length of the major axis and the area of the standard deviation ellipse. Computational formulae for the standard deviation ellipse are well known to the spatial statistical literature and not reported here (e.g., Ebdon, 1988; Lefever, 1926; Levine, 2002). Levine's modifications to Ebdon's specification have been used in this analysis. Elliptical parameters, (e.g., semi-major, semi-minor axes), are estimated using the location coordinates of the daily activities reported by individual members of each household.

At the design level, the literature suggests that transit demand is sensitive to the placement of stops relative to household and business locations. Specifically, research indicates that demand decreases rapidly at distances exceeding .40km (1/4 mile) (Beimborn et al., 1991). Informed by the literature, the sample has been subdivided into two *transit service categories* using a residence to transit threshold distance of .40km (1/4 mile). Transit supply has been modelled as the location of bus or light rail stops within the study area at the time of the Portland survey. Proximity to transit has been

[3] Opportunities for activity participation in web-space raise the possibility of extending the activity space concept to include *virtual locations*. This perhaps leads one to distinguish the *physical* from the *virtual* activity space and consider the potential for inter-space activity substitution.

specified as the Euclidean distance from a household's place of residence geocode to the closest transit stop (LTR or bus). Within the sample, 2,551 households are living in the optimal distance band (≤ .40km), leaving 1,098 in the category with poorer spatial access to transit. There is an initial expectation that households living beyond the threshold distance will conduct daily activities using dispersed activity locations. The implication is that a dispersed activity set will give rise to greater daily travel requirements. With this in mind, the relationship between activity dispersion and transit stop placement potentially holds value for evaluating the performance of TOD.

The analysis begins with a visual case study exploring the dispersion of activities associated with the activity spaces of a sub-sample of 12 households (Figure 2). A GIS-based tool for the visualization and exploration of activity-travel data has been used to support this analysis. In a separate paper, Buliung and Kanaroglou (2005a) describe the operational characteristics of the exploratory software system. In this study, the tool supports interactive geovisualization of activity dispersion in a manner facilitating the generation of hypotheses concerning activity-travel demand and transport infrastructure. In a computational setting, analysts can add or subtract additional structural layers (e.g., transit stops, LTR facilities, etc.). Evidence from the case study suggests that households located closer to transit appear to have ellipses that are smaller in area, with shorter major-axis lengths, than households located outside the specified threshold distance. Beyond the initial

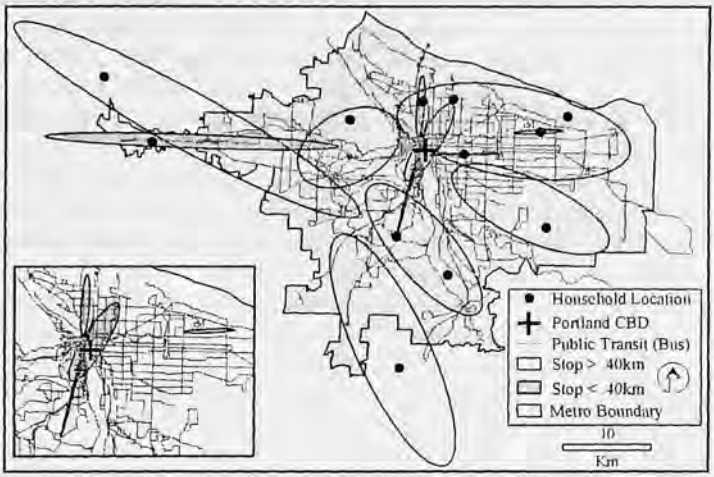


Figure 2. Dispersion of activities and proximity to transit.

behavioural observations, the case study demonstrates the potential of interactive geovisualization as an approach to exploring relationships between household activity-travel behaviour and transport infrastructure.

To further substantiate these initial observations, the standard deviation ellipse has been estimated for the entire sub-sample of 3,649 households. Parametric difference of means *t-tests* were used to compare the major axis lengths and elliptical areas of households allocated to the transit service categories[4]. The evidence suggests that households living beyond the specified threshold distance have longer major axis lengths than those living closer to transit, $t(3,649) = 12.80$, $p < .001$. A similar result was found for the area of the ellipse. On average, households living closer to transit have ellipses with smaller areas than those living beyond the threshold distance, $t(3,649) = 8.09$,

[4] Log transformation for normality was applied to the major axis length and elliptical area variables. Reported t-values are based on an assumption of unequal variance.

$p < .001$. Taken together, these results suggest that households living beyond the distance threshold typically conduct their daily activities using a more dispersed set of activity locations. This preliminary evidence suggests that there is an association between the stop placement dimension of TOD and the spatial outcomes of household activity-travel decisions.

While the results of the statistical comparison across transit service categories are not entirely surprising, a closer examination suggests that these findings may have practical value from a travel perspective. Separate ellipses generated from the mean elliptical properties estimated for each transit service category can be used for visual comparison (Figure 3). The result both reinforces findings concerning the spatial dispersion of activities across transit service categories while suggesting that the spatial potential exists for greater daily travel when households are located further from transit stops.

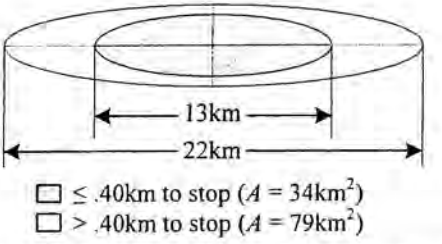


Figure 3. Standard deviation ellipses for transit service categories.

The findings reported in this section have been presented with two caveats. First, the metropolitan scale of analysis potentially obscures marked differences in dispersion within regional sub-areas where TOD initiatives have been aggressively pursued. An assessment of local effects has been left as a future research objective. Second, this experiment has been conducted with the

understanding that successful TOD depends on more than proximity to transit stops. Land use mix, density of development, auto-ownership, and the characteristics of households and individuals living and working close to transit combine to influence demand as well (e.g., Filion, 2001; Porter, 1997; Tunlin and Millard-Ball, 2003).

5 Conclusions

In this paper, I have argued for a refinement of the activity space concept to include several properties beyond those originally conceived by Horton and Reynolds (1971). With my proposed refinements in place, I have demonstrated an approach to evaluating the dispersion of activities within a household's activity space. The relationship has been explored between the dispersion property of the activity space and access to transit in the Portland, Oregon metropolitan area. Application of the activity space as a device for interactive geovisualization signifies an interesting innovation that would not have been possible during the 1970s. Evidence has also been presented suggesting an association between the dispersion property of the household activity space and proximity to transit. The findings suggest that households with poor spatial accessibility to transit tend to have spatially dispersed activity locations within their activity spaces. While it has been recognized that successful TOD depends on more than network structure and distance to stops, this research demonstrates that the activity space is a potentially useful concept for evaluating the success of transit-oriented development.

Acknowledgements. The author would like to thank Bill Stein and the Metro organization for providing data and technical assistance. The author would also like to thank the National Science and Engineering Research Council for postdoctoral funding support.

References

Bachi, R. 1963: Standard distance measures and related methods for spatial analysis. *Papers of the Regional Science Association* 10, 83-132.

Beckmann, M.J., Golob, T.F., and Zahavi, Y. 1983: Travel

probability fields and urban spatial structure: 2. Empirical tests. *Environment and Planning A* 15, 727-738.

Beimborn, E., Rabinowitz, H., Gugliotta, P., Mrotek, C., Yan, S. 1991: *Guidelines for transit sensitive suburban land use design*. US Department of Transportation, Washington DC.

Bhat, C.R. and Srinivasan, S. 2004: A multidimensional mixed ordered-response model for analyzing weekend activity participation. *Transportation Research B* 39: 255-278.

Booch, G. 1994: *Object-oriented analysis and design with applications 2nd edition*. Massachusetts: Addison-Wesley.

Buliung, R.N. and Kanaroglou, P.S. 2004: On design and implementation of an object-relational spatial database for activity/travel behavior research. *Journal of Geographical Systems* 6, 1-27.

Buliung, R.N. and Kanaroglou, P.S. 2005a: A GIS toolkit for exploring geographies of household activity/travel behavior. *Journal of Transport Geography*, forthcoming.

Buliung, R.N. and Kanaroglou, P.S. 2005b: Metropolitan polycentric urban form and household activity-travel behaviour. *Growth and Change*, currently under review.

Cambridge Systematics, 1996: *Data Collection in the Portland, Oregon Metropolitan Area: Case Study*. Report DOT-T-97-09. US Department of Transportation (USDOT), Washington DC.

Dijst, M. 1999: Two-earner families and their action spaces: A case study of two Dutch communities. *Geojournal* 48, 195-206.

Ebdon, D. 1988: *Statistics in geography 2nd Edition*. Oxford: Blackwell.

Filion, P. 2001: Suburban mixed-use centres and urban dispersion: What difference do they make? *Environment and Planning A* 33, 141-160.

Horton, F.E. and Reynolds, D.R. 1971: Effects of urban spatial structure on individual behavior. *Economic Geography* 47, 36-48.

Kim, H-M. and Kwan, M-P. 2003: Space–time accessibility measures: A geocomputational algorithm with a focus on the feasible opportunity set and possible activity duration. *Journal of Geographical Systems* 5, 71-91.

Kwan, M-P. 1999: Gender and individual access to urban opportunities: A study using space-time measures. *Professional Geographer* 51, 210-227 .

Lefever, D.W. 1926: Measuring geographic concentration by means of the standard deviational ellipse. *The American Journal of Sociology* 32, 88-94.

Levine, N. 2002: CrimeStat II: A spatial statistics program for the analysis of crime incident locations (version 2.0). Ned Levine & Associates: Houston, TX/ National Institute of Justice: Washington DC.

Lewin, K. 1951: *Field theory in social science; selected theoretical papers*. D. Cartwright (ed.). New York: Harper and Row.

Miller H.J. 1991: Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information Systems* 5, 287-301.

Nelson, D. and Niles J. 1999: Essentials for transit-oriented development planning: Analysis of non-work activity patterns as a method for predicting success. (http://www.globaltelematics.com/dn-jn-pt.htm)

Newsome, T.H., Walcott, W. and Smith, P.D. 1998: Urban activity spaces: Illustrations and application of a conceptual model for integrating the time and space dimensions. *Transportation* 25, 357-377.

Porter, D.R. 1997: Transit-focused development: A synthesis of research and experience. *TRCP Report* No. 20. National Academy Press, Washington DC.

Rumbaugh J., Jacobson I. and Booch G. 1999: *The Unified Modeling Language Reference Manual*. Massachusetts: Addison-Wesley.

Schönfelder, S. and Axhausen, K.W. 2003: Activity spaces: Measures of social exclusion? *Transport Policy* 10, 273-286.

Scott,D.M. 2003: Comparison of two GIS-based algorithms for generating potential path areas. *Paper presented at the 50th Annual North American Meeting of the Regional Science Association International, March.*

Tunlin, J. and Millard-Ball, A. 2003: How to make transit-

oriented development work. *Planning* 69, 14-21.
Yuill, R. S. 1971: The standard deviational ellipse - An updated tool for spatial description. *Geografiska Annaler. Series B, Human Geography* 53, 28-39.
Zahavi, Y. 1979: *The 'UMOT' Project Report.* Report DOT-RSPA-DPB-20-79-3, US Department of Transportation, Washington DC.

Developing a GIS-based detailed traffic simulation model for the Montreal region: Opportunities and challenges

Timothy Spurr and Murtaza Haider, Ph.D.
McGill University, Montreal Canada (http://www.milute.mcgill.ca)

1. INTRODUCTION

Deciding on an appropriate transportation planning policy can be difficult. One approach is to try and model the travel behaviour of individuals and then use these models to test the impacts of various policy options. But accurate modeling of individual behaviour is not straightforward. Any analysis of travel behaviour requires an acknowledgement of a vast array of externalities and unpredictable outcomes. While this difficulty will never be completely resolved, the advent of low-cost computing power now affords researchers the capability to process enormous amounts of data quickly. As a result, complex models accounting for thousands or even millions of variables and observations can be constructed and evaluated. Given sufficient data, it is now possible to build - at relatively low-cost - detailed models of urban regions which can be useful in informing public policy.

The road networks in many cities throughout the world, including Montreal, have been modeled according to the deterministic user-equilibrium framework. The present research is not breaking new ground in this respect but one or two departures from the norm are worth noting. First of all, the majority of large-scale traffic models use the EMME/2 software. The model described in this paper uses TransCAD. Secondly, while most models use greatly simplified networks containing only highways and major streets, the present model contains *all* streets in the region. Such detail is possible due to the integration of Geographic Information Systems (GIS) with conventional transportation modeling tools in the TransCAD software.

GIS is a powerful transportation planning tool for several reasons. First and foremost, it can be used to efficiently link travel behaviour patterns and socio-economic data within a spatial framework. Secondly, it incorporates geographic details which are essential in the construction of realistic links and nodes as well as in the discretization of space. Thirdly, it requires little formal programming although such capabilities are available for complex analyses. Finally, GIS generates visual output of the analysis results which are invaluable in both the application and validation of the model.

The retention of local streets and detailed network geometry may bestow only small benefits on a static model of automobile traffic, especially if the assignment method remains aggregated to a system of zones. But these details become essential in accurate models of transit where the precise location of bus stops and their corresponding intersections is necessary to measure walking distances, transfer times, accessibility and overall performance of the transit system. The importance of geographic detail is even greater if an integrated transportation-land use model is the ultimate goal. Land use data exists at the parcel level and transportation data must therefore be at least as disaggregate and as detailed.

2. REVIEW OF CURRENT PRACTICE

Most of the recent research in traffic modeling describes theoretical formulations of dynamic modeling that are ever more adept at capturing the temporal distribution of choices and activities. Static models, meanwhile, are rightly criticized for poorly representing utility-maximization theory and being unable to handle situations where demand exceeds capacity or where congestion is not continuous (Verhoef, 1997). While dynamic models are undoubtedly better at applying the theory, there are usually no direct comparisons of results with static models when used in planning practice (Wu et al., 2001; Nagel et al., 2000; Boyce and Bar Gera, 2003). Researchers are encouraged to pursue investigations into dynamic, micro-level modeling due to the ever-increasing computing power made available.

Modelers likely face the law of diminishing returns whereby enormous additional investment in spatial-temporal complexity will generate only marginal improvements in model quality. This is especially true when the model is used to forecast several years into the future. The fact that most transportation planners working both inside and outside academia continue to use static model frameworks such as EMME/2 is evidence of their persistent relevance (Florian, 1999; JPINT, 1998; Tremblay, 2004; Krajczar, 1998).

If static models continue to represent the norm in large-scale transportation modeling, methods of improving the static framework are of natural interest. One of the significant events in this regard is the advent of Geographic Information Systems (GIS). Since the early 1990s, enormous progress has been made in the development of GIS which offer solutions to problems inherent in the transport modeling process, including the modifiable areal unit problem (MAUP), boundary problems and spatial sampling, spatial dependency and spatial heterogeneity (Miller, 1999). Spiekermann and Wegener (2000) provide a good example of how GIS can be used to mitigate the problems inherent in the use of zone systems for transportation analysis. Furthermore, GIS allows for greater flexibility in managing data sources, can be used to construct highly detailed transportation networks and generates easily-decipherable visual output (Arampatzis et al., 2004). Also, the general user-friendliness of a GIS interface can greatly reduce the costs, in terms of both time and money, incurred in the construction of a transportation planning module (Souleyrette and Anderson, 1998; Johnston and de la Barra, 2000). Finally, as the modeling practice becomes more and more concerned with spatial and temporal disaggregation, GIS plays an increasingly important role (Frihida et al., 2002).

3. DATA

The study area for the traffic assignment model was defined as the three census metropolitan areas (CMAs) of Montreal, St-Jean-sur-Richelieu and Salaberry-de-Valleyfield. The census tracts and enumeration areas of these regions formed the basis of the system of traffic analysis zones. The travel demand data was based upon the

1998 O-D Survey undertaken by the Agence Métropolitaine de Transport (AMT, 2001). Detailed street network data with link travel times was acquired through DMTI Spatial Ltd. This network contained approximately 135,000 bi-directional links.

4. METHODOLOGY

4.1. Zone system and O-D matrices

The three CMAs comprising the study area were segmented into 947 zones. The centroid of each zone was the point from which traffic was generated. In addition, 34 pre-existing network nodes located around the perimeter of the study area were used to generate external trips. The result was a 981x981 O-D matrix.

Three one-hour O-D matrices for auto-drive trips were constructed based on the trip data provided in the O-D survey. The first matrix contained all trips that began between 6 am and 7 am, the second matrix contained trips beginning between 7 am and 8 am, and the final matrix contained trips beginning between 8 am and 9 am. Total travel demand for the three hours was roughly 878 000 trips.

4.2. Construction of the network

Traffic assignment models in TransCAD require the construction of network files. A network file is a digital representation of a line layer as a series of nodes and directional links. When a network file is built, the user is asked to provide a set of links in a line layer that will form the basis of the network file, as well as link attributes to be included in the network. Examples of link attributes are segment length, travel time and capacity.

In order to model congestion, additional line layer attributes had to be incorporated into the network file. The BPR volume- delay function has two parameters, alpha and beta, which vary by segment according to the road's functional class and free flow speed. Each network link therefore had to be assigned values for alpha and beta. The present model is based upon parameters estimated by Horowitz (1991) – see Table 1. Horowitz did not provide an estimate for global

parameters on urban roads operating under interrupted flow conditions because factors other than the volume-to-capacity ratio – such as signals and intersection characteristics – affect the link performance. Nevertheless, global values were required and were estimated ad hoc: alpha was taken to 0.6 and beta was taken to be 2 based on the trends in parameter variation by road functional class (see Table 1). Delays incurred at intersections had to be estimated in a similarly coarse fashion and were applied globally. A left turn was assigned a cost of 0.5 minutes, a right turn 0.2 minutes and a through movement 0.05 minutes.

The next step in the network construction process is the connection of centroids to the actual network using centroid connectors. These dummy links are supposed to substitute for local roads. Therefore, they should not connect directly to highways or highway access ramps but rather to collector streets or local roads. TransCAD's GIS capabilities were used to select those nodes that were not connected to highways or ramps. Only nodes in this set would be considered as candidates for connections to centroids.

Once constructed, the centroid connectors were assigned a travel speed of 40 km/h – the same as for local streets. The travel time over the link was simply the length divided by the speed. Finally, each centroid connector was assigned a capacity of 99999 vehicles per hour in order to prevent travel time increasing with volume.

4.3. Validation and adjustment of the model

Overall model performance is evaluated based upon comparisons between predicted and observed flows on those links for which there are observation data. The observed flow data provided by the MTQ contained 197 observation posts, each one corresponding to a single directional link.

Several statistics could be computed to quantify the model's performance. The first is the percentage error which is simply the difference between total observed and predicted counts divided by the total observed counts. In general, the predicted flows were expected to be lower than the observed flows, especially when aggregated over all the observation posts. While it is possible for the O-D survey to

under-represent the population, it is unlikely that it is over-representative in the aggregate. The detailed nature of the street network was expected to provide many alternate paths and thereby reduce flows over some routes.

The next statistic that can be computed to measure performance is the percent root mean squared error (% RMSE). This is an aggregate statistic which measures the performance of the model as a whole, comparing forecast flows to observed counts. The U.S. Department of Transportation's Travel Model Improvement program recommends a %RMSE of less than 30 (TMIP, 2001).

Finally, a linear regression model was employed to measure the correlation between forecasts and counts. Ideally, the intercept of the linear function should be zero and the slope of the line should approach 1. The USDOT recommends a region-wide R-squared of at least 88% (TMIP, 2001.).

After the level of performance had been assessed using the above statistics, an error investigation began. One major potential source of error lies with the O-D matrix and its validity was checked immediately following a trial run using visual inspection and the text output of the assignment procedure. The matrix should display an expected level of travel demand and this demand should be equal to the total demand expressed by the appropriate records in the tabular survey data.

The first level of assignment error-checking was accomplished using the flow map. The flow map is a combination of two thematic map types: a colour theme, and a scaled symbol theme. Each link is assigned a colour based upon its volume-capacity ratio and a width based upon its hourly volume (see Figure 1). Serious network problems were apprehended by looking at this map. For example, a link on a major highway may have had no colour and only a hairline width, indicating that there was no traffic on that link. Such occurrences suggested either there is no access to this link due to an error in digital representation, or the link had faulty attributes that rendered alternative paths less costly.

The flow map also displayed illegal traffic movements. These were especially common on bridges and freeways where access was limited. If traffic was jumping onto overpasses instead of using

ramps, it would usually be apparent in the flow map. These problems could be corrected by adding turn prohibitions and TransCAD provides a tool for doing this.

Another method of error detection involved looking at a scatter plot of forecasted vs. observed traffic volumes. Data points that were especially far from a line traced from the origin with a slope of 45 degrees represented locations where the model was performing poorly and these were investigated. Determining the root cause of these inaccuracies, however, was very difficult, especially if the link in question is a controlled urban road.

The Montreal region affords a natural benefit to modelers because of its geography. Both the city of Montreal and the city of Laval occupy islands in the St. Lawrence River. Both of these islands are accessible by 22 bridges, 21 of which hosted MTQ observation posts. Obviously, there is no other way for a vehicle to enter these two islands, so this setup allows for a very efficient screen-line and cordon analysis (see Figure 2). It was hoped that, even if predicted flows on individual bridge links did not match the observed flows, the forecasts for each region taken as a whole would be a good fit with observations. If this was in fact the case then the characteristics of individual bridges could be modified so that the flow was properly distributed.

A similar approach was taken with freeway links. In an interrupted flow regime, it is usually the signaling and signage that determines the road capacity. The performance of arterial, collector and local roads was not expected to be very good due to the complete absence of intersection data in the model. The capacity of freeways, on the other hand, depends mostly upon the free-flow speed and the number of lanes, both of which had been accurately represented on most freeway segments. It was therefore expected that the performance of freeway links would be good indicators of the validity of the network and the model in general. Since most of the region's large bridges carry freeways, freeway links were of interest in the validation process because their performance has a direct effect on the performance of bridges.

5. RESULTS

Model performance results appear in Table 2. This table shows that the r-squared for all 3 hours of 0.8759, and the %RMSE is 30.3. The two columns on the right-hand side of the table show the number of bridge and freeway links whose forecast flows were within 5% of observed values. Overall, the model is still approaching the minimum TMIP performance standard but the accuracy of forecast flows on individual links varies considerably.

In addition to the %RMSE and the r-squared, the %error can also be used to evaluate model performance. As can be seen in Table 2, the first hour of the AM peak generates traffic flows 15% below observed levels. In the second hour, forecast flows are nearly 6% above observations. In the third hour, flows are roughly 2.5% too low. It is difficult to explain the "missing" traffic in the first hour based solely upon the aggregate performance measures. However, the "excess" traffic forecasted between 0700 and 0800 may be due to the inability of the static model to deal with queues. During this hour, many Montreal arteries experience demand well above capacity resulting in queue formation upstream of the link and forced flow on the link itself. The static model cannot represent either of these phenomena and so flow continues to appear on links even though absolute capacity has been reached. The fact that the static model performs well when all three hours are taken together is evidence of this phenomenon. By 9 am, most of the queues on the region's roads have cleared and so, over the three hour AM peak, total demand has been met and the total percent error is around -3%.

Three additional aggregate measures were used to test the model's validity. Average trip length, average trip time and average trip speed were calculated based upon output generated by TransCAD in the traffic assignment procedure. Average trip length was computed by dividing the number of vehicle-kilometres traveled by the interzonal travel demand. Average trip time was obtained by dividing total vehicle-minutes traveled by interzonal travel demand. Finally, average speed can be computed by dividing average trip

length by average trip time. The results are enlightening (see Table 3).

The longest trips, both in terms of distance and time, occur in the first hour of the AM peak. This hour also contains the smallest percentage of intra-zonal trips (which are not assigned to the network), and the smallest travel demand. These findings are intuitively correct. People who must make longer trips must depart earlier if they wish to arrive on time. Also, since most stores and schools are still closed between 6 and 7 am, there are very few local trips being made. Furthermore, this hour displays the highest average vehicle speed which can be attributed to the relatively low level travel demand and therefore low congestion. Moreover, people making longer trips are more likely to use high-speed facilities like freeways.

During the second hour, between 7 am and 8 am, average trip length is considerably shorter while average trip duration is only slightly shorter than in the previous hour. Travel demand has nearly doubled and the percentage of intra-zonal trips has increased. Once again, this situation seems to correspond well with reality. This is the hour during which most people leave for work or school and local trips become more prevalent as stores start to open. Travel demand reaches a peak. Therefore, roads are highly congested and average speed drops. Trips are shorter on average, but most of the time savings are cancelled out by the increased congestion.

During the final hour of the AM peak, demand drops off only slightly, but trips are noticeably shorter and the intra-zonal percentage is much higher. Average speed increases due to easing of congesting indicating that fewer trips are being made along high density corridors. This, in turn, suggests that trip purposes are more local and more "random". Many people are commuting to work and school, others are running errands, going to meetings or dropping off family members or colleagues.

The progressive diversification of trip purposes with each successive hour can be verified in the analysis of the O-D survey presented in Table 4. In the first hour, work trips comprise nearly 80% of the total. By the third hour, that share has decreased to 57% while the share of "other" trip purposes has risen to 36%. Also note the increase in the percentage of shopping trips from 0.4% in the first

hour to 5.2% in the third. It seems, therefore, that the model is accurately representing these changes in travel patterns.

6. CONCLUSIONS

The goal of this research was to build a detailed model of travel behaviour in the greater Montreal area. It was hoped that current GIS technology would allow this goal to be accomplished fairly easily without an enormous allocation of labour and monetary resources. Indeed the authors, working over a 2-year period, have been able to construct a 240,000 link network that can reliably predict current traffic flows. In fact, the present network performs sufficiently well that the next stages of the regional planning research can be contemplated, while the model itself is improved over time.

Such a detailed transportation model lends itself to integration with land use models where the areal unit of interest is the individual parcel. Indeed, this integration is essential in effective forecasting of land use, urban development and the interaction of built form with travel patterns. The transportation model described in this paper serves as a strong foundation for achieving this objective.

7. REFERENCES

Agence métropolitaine de transport. (2001). *Enquête orgine-destination 1998 de la région de Montréal (V.3).*

Arampatzis, G., Kiranoudis, C.T., Scaloubacas, P., Assimacopoulos, D. (2004). A GIS-based decision support system for planning urban transportation policies. *European Journal of Operational Research,* 152, 465-475.

Boyce, D., Bar-Gera, H. (2003). Validation of multiclass urban travel forecasting models combining origin-destination, mode, and route choices. *Journal of Regional Science,* 43, 517-540.

Florian, M. Understanding Traffic Congestion. (1999). *ORMS Today* April 1999. Retrieved January 6, 2005 from http://www.lionhrtpub.com/orms/ORMS-archive.html

Frihida, A., Marceau, D.J., Theriault, M. (2002). Spatio-Temporal Object-oriented data model for disaggregate travel behaviour. *Transactions in GIS*, 6, 277-294.

Horowitz, A.J. (1991). *Delay-Volume Relationships for Travel Forecasting: Based on the 1985 Highway Capacity Manual*. Washington, D.C.: Federal Highway Administration U.S. Department of Transportation.

Joint Program in Transportation (JPINT) Data Management Group. (1998). *G.T.A. A.M. Peak Hour Network Coding Standard Part 1 – Notation*. Toronto: University of Toronto.

Krajczar, K. (1998). *Evolution of the Greater Vancouver Transportation Model*. Presented at the 13th Annual EMME/2 User's Conference.

Miller, H. (1999). Potential Contributions of Spatial Analysis to Geographical Information Systems for Transportation (GIS-T). *Geographical Analysis*, 31, 373-409.

Nagel, K., Rickert, M., Simon, P.M. & Pieck, M. (2000). *The Dynamics of Iterated Transportation Simulations*. Working Papers 00-02-012, Santa Fe Institute.

Souleyrette, R.R., Anderson, M.D. (1998). Developing small area network planning models using desktop GIS. *Journal of Urban Planning and Development*, June 1998, 55-71.

Spiekermann, K., Wegener, M. (2000). Freedom from the Tyranny of Zones: Towards New GIS-Based Spatial Models. In Fotheringham, A.S. and Wegener, M (Eds.). *Spatial Models in GIS: new potential and new models*. Philadelphia: Taylor and Francis.

Transportation Model Improvement Program (TMIP). "Assignment Procedures" in *Model Validation and Reasonableness Checking*

Manual. U.S. Department of Transportation. 2001. Retrieved September 30, 2003 from http://tmip.fhwa.dot.gov/clearinghouse/docs/mvrcm/ch7.stm

Tremblay, Pierre. (2004, Dec. 18). Urban Transportation Planning Tools at Transports-Quebec. Presentation at McGill University.

Verhoef, E. (1997). *Times, speeds, flows and densities in static models of road traffic congestion and congestion pricing*. Discussion paper retrieved January 6, 2005 from http://www.tinbergen.nl/discussionpapers/97070.pdf

Wu, Miller, Hung. (2001). A GIS-based decision support system for analysis of route choice in congested urban networks. *Journal of Geographical Systems*, 3, 3-24.

8. TABLES

Table 1 – Estimated BPR function parameters (Horowitz, 1991)

Coefficient	Freeways			Multilane		
	70 mph	60 mph	50 mph	70 mph	60 mph	50 mph
alpha	0.88	0.83	0.58	1.00	0.83	0.71
beta	9.80	5.50	3.60	5.40	2.70	2.10

Table 2 – Performance of AM peak hour models

Hour	R-squared	% RMSE	% Error	Bridge links within target range (out of 42)	Autoroute links within target range (out of 44)
0600-0700	0.8692	35.89391	-15.1613	2	6
0700-0800	0.8305	36.49927	6.222203	6	8
0800-0900	0.8201	35.31554	-2.54309	7	5
AM Peak (all 3 hours)	0.8759	30.34158	-3.27609	8	7

Table 3 – Performance of AM peak hour models; trip data

			Trip Lengths		
Hour	Total Demand	% Intra-zonal	Minutes	Kilometres	Avg Speed
0600-0700	193559.47	7.75	22.44	21.07	56.33
0700-0800	378600.05	8.80	20.06	16.60	49.67
0800-0900	352772.96	11.36	14.17	13.52	57.21
AM Peak (all 3 hours)	924932.48	9.56	18.37	16.40	53.58

Table 4 – AM Peak auto-drive trips classified by purpose

	Work Trips	School Trips	Shopping Trips	Other
0600-0659	79.66633	1.633656	0.3913191	18.3882659
0700-0759	67.96934	4.385146	1.2195222	26.8132624
0800-0859	56.57752	4.329476	5.2275161	36.1289279

9. FIGURES

Figure 1 – Traffic assignment model output

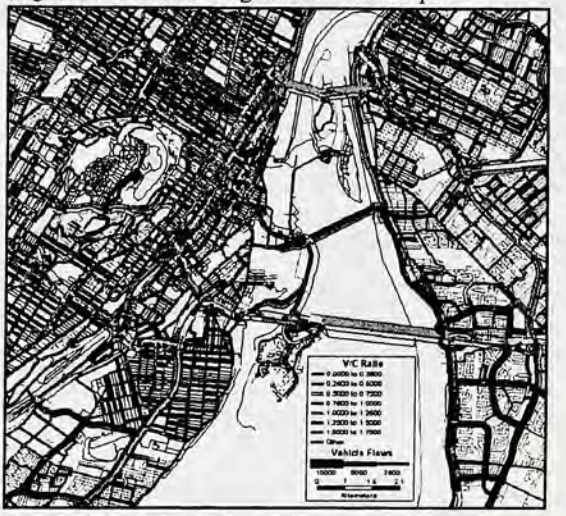
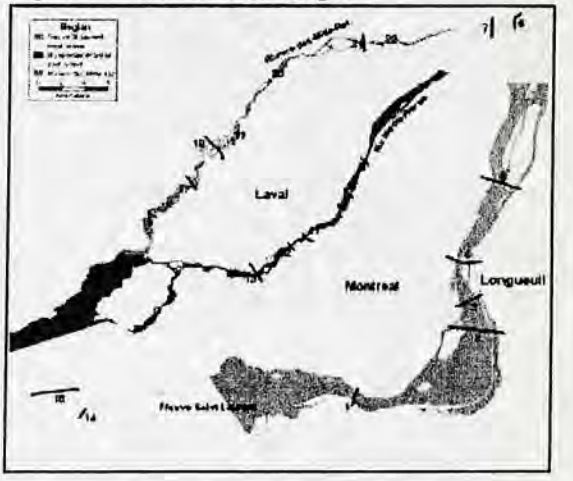


Figure 2 - Montreal Area bridges



THE IMPORTANCE OF SECURITY IN THE SERVICE OFFERINGS OF CANADA'S CUSTOMS BROKERS

by

Michael A. Haughton
School of Business and Economics
Wilfrid Laurier University

Abstract

Concerns about potential security vulnerabilities that arise when the movement of products transcends national borders have led to calls for more highly evolved supply chain management (SCM) practices. Given the national security implications of such vulnerabilities, a key element in these SCM practices is compliance with the pre-emptive security initiatives that governments have developed; e.g., Canada's Partners in Protection (PIP) program. For assistance in achieving compliance, importers and exporters have been urged to include third parties among potential sources of expertise to be drawn on. This paper is based on a study of the extent to which one of these third party groups –Canadian customs brokers– is prepared to provide the required help. In particular, the research offers insights concerning where Canadian customs brokers are in relation to their US-based counterparts.

Introduction

Dominated by writings since the horror of 9/11, the literature on supply chain management (SCM) discipline has emphasized that the current security conscious era calls for more highly evolved supply chain practices; see, e.g., Sheffi (2001), Helferich and Cook 2002), Russell and Saldanha (2003). A central goal of the required practices is securing supply chains from acts that can disrupt the efficiency and

continuity of trans-border product flows; e.g., an act of terrorism that destroys supply chain assets and infrastructure. But curiously, the government policies aimed at minimizing vulnerability to those acts that can harm national security can also impede smooth supply chain operations; e.g., time-consuming physical inspection of imported goods; see, e.g., Flynn (2000). Thus, effective management of supply chains that transcend national borders requires approaches that not only reflect awareness of the potential security vulnerabilities but also comply with security-related government policies.

In Canada, an illustration of the logic that a corporation's compliance with such policies is consistent with good SCM is documented in the Customs Action Plan 2000-2004. In that document, the then Canada Customs and Revenue Agency (now the Canada Border Services Agency (CBSA), the new home department of Canada Customs) presents a fictional case study to explain the trans-border logistical efficiencies for companies that operate under one of its policies – the Customs Self Assessment (CSA). Empirical affirmation of this logic can be found in statements from companies such as JCPenney, Nike, and Liz Clayborne. Specifically, they mention that the lower-than-industry customs inspection rates of their incoming shipments from overseas result from having supply chain security policies and compliance records that meet or exceed government requirements. See, for example, Douglas (2002) and Shanahan (2004).

Still, expressed concerns by Canada's international trade sector about the operational and financial burdens of making the adjustments required to comply with this and a slue of allied policies are well documented in the practitioner literature; e.g., Smyrlis and Toews (2003), Foster (2003), Harps (2003a). Customs brokers, as acclaimed experts on customs regulations, are touted among the third parties that corporations should consider in their search for assistance in complying with government policies.

But this does not guarantee adequate business for all customs brokers because in order to win a corporation as a client, they have to prove themselves capable. Janet Reuter, Rockwell Automation's director of

import/export compliance, affirms this in the following statement concerning the use of customs brokers:

> *"Select forwarders and brokers that will work on their own process improvement, and challenge your organization to do the same... then measure their performance, such as evaluating how long it takes the freight forwarder to turn documents over to the customs broker, and how fast the broker clears shipments."*
>
> Excerpted from Harps (2003b)

This raises the question of to what extent are Canada's customs brokers *actually* improving their own processes and developing the capabilities to facilitate objectives such as fast shipment clearance? The research to help in answering this question is based on a study involving member organizations of the Canadian Society of Customs Brokers (CSCB), the Canadian International Freight Forwarders Association (CIFFA), and the National Customs Brokers and Forwarders Association of America (NCBFAA). The remainder of this paper reports on the research in 4 sections as follows. The next section is an overview of some of the key policies in the Canadian government's approach to balancing the requirements of security and efficiency in trans-border trade, and some specific roles that customs brokers can assume in helping their clients to comply with these policies. Following that overview is description of the research methodology, which precedes the section on the paper's findings and related insights. The paper then concludes with a summary and an outline of some follow-up research.

Government Security Policies and Trans-border Trade

To date, the Canadian government has put forth a broad range of security-conscious policies to direct trans-border trade. Underpinning these policies is the goal of maximizing the number of shipments falling into the category of being legitimate and presenting low risk to a nation's security. That way, the conduct of physical border/customs inspections can focus on the few shipments suspected of being outside of that category. Prompted in part by heightened security

fears following the bombings of 9/11, the Free and Secure Trade (FAST) and Partners in Protection (PIP) are among the latest in the evolving set of Canada's security policies. Because FAST and PIP are representative of the aforementioned goal in the set of initiatives, an outline of these two suffices to depict the essence of government's security philosophy and to illustrate some ways in which customs brokers can be of value to their clients. According to the FAST fact-sheet at the CBSA's website, the initiative allows companies to prove, through government tests of their import/export procedures, that the procedures result in US/Canada trans-border shipments that warrant government certification as being legitimate and representing low security risks.

For as long as a company and those with custody of its shipments during the shipments' trans-border journeys (carriers, drivers) remain certified, it will enjoy faster average border/customs clearance than others that are not FAST-certified. The Partners in Protection (PIP), a post-9/11 policy, also reflects the spirit of encouraging the trade community to support government security goals by making a commitment to developing security-enhancing procedures internally. However, "partner" status does not automatically earn FAST benefits since it has a lower threshold than FAST certification. In particular, the focus of PIP accreditation is that firms voluntarily sign on to avow their *commitment* to security. So partner status is really a step towards becoming FAST-certified. The Customs-Trade Partnership Against Terrorism (C-TPAT) in the US is equivalent to Canada's PIP.

Becoming a PIP signatory would certainly be one way for a broker to move towards meeting the sort of client expectations expressed in the earlier quote from Rockwell Automation's director of import/export compliance. For example, partner status provides an opportunity for deep understanding of government's requirements concerning the flow of both goods and supporting documentation in an import/export transaction. Thus, it seems likely that customs brokers availing themselves of such opportunities will be better able to guide their clients towards security-enhancing procedures. Consequently, as one gauge of Canada's customs brokers' readiness to assist their clients,

this study will focus on the rate at which CSCB members have signed on to the PIP (and, to provide a reference point, the rate at which US-based customs brokers have signed on for C-TPAT).

A noteworthy point is that the motives for PIP-certification tend to be more evident for customs brokers who take some custody of goods (e.g., storage, transportation, packaging, partial processing) during the import/export journey. That is because they face the obvious risk that, under their watch, biohazard, contraband, etc., can be introduced to contaminate an otherwise compliant shipment. Still, custody of the import/export information/documents and the underlying information systems also has security risks. An example would be a system that fails to accurately screen out customers to which exports are deemed by the exporting country's government as security vulnerabilities. Similarly, unscrupulous customs brokerage employees who handle sensitive information on imports/exports can be a source of security risks. In fact, at the 2002 U.S. Customs Symposium, a presentation by the owner of a customs brokerage firm highlighted job applicant screening among several security measures developed as part of his firm's C-TPAT program.[1] Therefore, while not always immediately obvious, the motives for PIP certification can be just as strong for customs brokers without shipment custody as it is for those with custody.

Also worthy of note is that customs brokers without carrier operations are not eligible to move beyond PIP to seek FAST-certification since FAST is only for carriers and importers. To clients, this move by an eligible broker would probably be a stronger signal of a carrier-based broker's capabilities than PIP. However, the metric of FAST-certification among customs brokers, which is being analyzed as part of a larger study in progress, is not covered in the present paper.

A second gauge of customs brokers' readiness that this paper uses concerns certification in the ISO 9000 series. Despite criticisms of

[1]As reported in *COMMUNICATOR*, Vol. 10, No. 1, Winter 2003 (the newsletter of Barthco, a US-based customs broker).

ISO 9000 certification as a metric of quality (e.g., by Philip Crosby as quoted in Heizer and Render, 2001), it is usable because (at least being in the process of) becoming certified evinces attainment of significant milestones in the pursuit of a goal that is a relevant client consideration: continuous process improvements. The third gauge, which focuses on CSCB members only, is the level of involvement of brokers in pre-PIP/FAST government initiatives geared to expedited customs clearance; i.e., issuing customs clearance decisions prior to the arrival of goods at border/customs checkpoints. These initiatives include Accelerated Customs Release Operating System (ACROSS) and Pre-Arrival Review System (PARS). The logic here is that, compared to customs brokers that do not, those using the earlier expedited clearance systems, which lay the basis for later government initiatives such as FAST, are better placed to help their clients benefit from those subsequent initiatives. The next section discusses the methodology around obtaining and analyzing data on these gauges in order to answer this paper's core question.

Research Methodology

The methodology for determining the number of Canadian customs brokers (i.e., regular members of the CSCB) signing up to PIP was to combine searches of the CSCB members' websites, the websites of the CBSA, and other websites to fill in information gaps. Analogous sites were searched for information on US-based brokers; e.g., the U.S. Customs and Border Protection (CBP), which is the equivalent agency to the CBSA. The CBSA displays the list of PIP members[2] (last update used was February 11, 2005) and this was supplemented by searching the websites of the CSCB members up to March 4, 2005). The website search method was automated (using Google and Yahoo) to determine if the broker indicated participation in PIP. All customs broker participants on the CBSA list were found to indicate their participation at their individual site so this indicated adequate reliability in the approach of searching individual customs broker websites to determine the more recently admitted partners. This

[2] http://www.cbsa.gc.ca/general/enforcement/partners/list-e.html

bolstered the basis for that approach; i.e., customs brokers are likely to broadcast that they are partners instead of not doing so.

For comparison with the U.S., where, unlike Canada, customs brokers and freight forwarders are members of a single organization (the NCBFAA), the list of 152 regular CSCB member firms and the list of 167 regular member firms of CIFFA were combined to yield a net total of 282 firms (after accounting for dual memberships). So although the intended research focus was on customs brokers, CIFFA members were also included to ensure a valid reference point from the U.S. The U.S. CBP does not make its list of C-TPAT signatories publicly available and the most recent available statistics found on the Internet was from a presentation at the U.S. Customs Symposium in 2002. Despite the date difference between NCBFAA sign-ups for C-TPAT and CSCB/CIFFA sign-ups for PIP, the comparison, as will be shown, is still quite instructive.

For the ISO certifications, searches of the brokers'/forwarders' sites were used. Although the methodology was automated (using Google and Yahoo), visual website inspections were also necessary to overcome a weakness in the automated approach. In particular, when the only indication that a customs broker is ISO-certified is an object, e.g., the ISO certification logo at the broker's home page, a text-based automated search of "ISO" failed to detect it. From among the 709 NCBFAA members, 80 were a randomly selected for website perusal. A website search was not feasible for every single CSCB/CIFFA member (some seemed yet to have a website, some members' sites were under construction, etc.) so only 170 CSCB/CIFFA sites could be searched. Of note is that some NCBFAA members are also members of CSCB and/or CIFFA (i.e., they have operations in both the U.S. and Canada. From the perspective of getting insights into the broker/forwarder expertise and capability that clients in each country have access to, no adjustments were considered necessary for double/triple memberships. The focal point of gauging customs broker involvement in earlier expedited clearance initiatives was Canada (i.e., pre-PIP/FAST mechanisms such as PARS) so website searches concerning that gauge were restricted to CSCB members.

Findings and Discussion

Table 1 shows the results concerning involvement in PIP and C-TPAT. The results show that the current rate of PIP involvement by CSCB/CIFFA members (9.57 of the membership) is still below what the rate of C-TPAT involvement by NCBFAA members was as far back as two years ago: 35.54% of the membership. While one can speculate on the reasons for the seemingly large discrepancy (e.g., since the U.S. was the direct victim of the 9/11 attacks, its corporate citizens are more spurred to demonstrably support their government's national security initiatives), the fact is that CSCB/CIFFA support for PIP seems far below anything describable as a reasonable benchmark.

Table 1: Signatories to PIP and C-TPAT from the CSCB, CIFFA, and NCBFAA Memberships

	NCBFAA Members (C-TPAT)	CSCB Members (PIP)	CIFFA Members (PIP)	CSCB/CIFFA Members in Total (PIP)
Number of Members	709	152	167	282
Number of Signatories (% of Membership)	252[A] (35.54%)	23[B] (15.13%)	18[B] (10.78%)	27[B] (9.57%)

Notes: [A]As reported at the U.S. Customs Symposium in 2002
[B]Based on data collected up to March 4, 2005

An intriguing observation that emerged during data collection is that one of the PIP signatories from the CSCB membership, Vastera, is purely non-asset-based. In fact, Vastera does not take any physical custody of goods during their import/export journey. Vastera's value proposition is software that clients can use to automate and ensure trade compliance. A probe of Vastera's press release concerning its PIP certification reveals that Vastera views this as contributing to its

own preparedness to "*offer support to its own clients throughout the application process.*" However, beyond this stated reason, screening out restricted trade partners is a key function of Vastera's software system, which, as noted earlier, can be security risk if the system is flawed. Thus, the unique case of Vastera is instructive in reinforcing that in addition to the more obvious goals of serving clients through a consulting relationship regarding PIP (and FAST) and by protecting their import/export shipments from *direct* physical vulnerabilities, providing protection by minimizing vulnerabilities of information systems and documentary flows is also crucial. Thus, it can be argued that CSCB and CIFFA members have strong reasons to raise their PIP-participation beyond and current level of 9.57%.

Table 2 shows the status of ISO 9000 series certifications among the CSCB, CIFFA, and NCBFAA members. The results suggest that, using the NCBFAA membership as a benchmark (albeit on the basis of an 80-member sample), the externally vetted process improvement efforts of CSCB and CIFFA members seem to be, at least, on par. To wit, while 12.50% of the sampled NCBFAA members were found to have at least one ISO 9000 series accreditation (yielding a population 95% confidence interval estimate of 12.50±6.97), 15.88% of the sampled CSCB/CIFFA members were found to be so accredited, with another 1.18% undergoing the accreditation process. The above suggestion is also reinforced by the fact that via parent/acquired company relationships, a further 3 (1.76%) of the sampled members of CSCB/CIFFA are operationally –as distinct from just financially– linked to accredited companies (two being carriers, and the other a non-asset-based logistics services company).

Of the 82 CSCB members whose websites were perused, 64.63% of them *unequivocally* indicated at their websites that they are involved with at least one of the CBSA's expedited clearance initiatives (Frequent Importer Release System (FIRST), ACROSS, PARS, etc.). Not all customs brokers were explicit about their involvement. True, because their sites carried statements and bullet points that *implied* involvement (e.g., "we are EDI capable", "on-line documentation", "full on-line access to all ports across Canada"), the above statistic

may understate the actual level of involvement (e.g., it stands to reason that a broker with EDI capabilities is likely to be plugged into the electronic release mechanisms of the CBSA). However, for some brokers it would have required a leap of inference to unequivocally classify them as being involved in expedited clearance initiatives, so the analysis was limited to unambiguous expressions of involvement. Thus, since a broker's involvement in the earlier expedited clearance mechanisms is a foundation for transitioning to later initiatives such as PIP, the findings indicate that 64.63% (53) of the brokers in the resulting sample evidently have (at least some of) what is required to make the transition. Of course, the extent to which they actually continue to do so in order to add value to their clients' trans-border supply chain operations is an open question for future research. What is clear is that fewer than half of these 53 seemingly well-positioned brokers (43.40% or 23 of them) have made the transition to PIP.

Table 2: ISO 9000 series certifications among the CSCB, CIFFA, and NCBFAA Memberships

	NCBFAA Members	CSCB Members	CIFFA Members	CSCB/CIFFA Members (Total)
Number of Members	709	152	167	282
Sample size[A]	80	82	125	170
No. ISO certified (% of sample)	10 (12.5%)	18 (21.95%)	22 (17.60%)	27 (15.88%)
No (%) in ISO-certifying process			2 (1.60%)	2 (1.18%)
No. (%) Other[B]			4 (3.20%)	4 (2.35%)

Notes: [A]A random sample of 80 was chosen from among the NCBFAA members; for the CSCB/CIFFA members, the resulting comprised those with an accessible website.

[B]These were based on the following situations: (i) a claim that "our systems exceeds ISO 9002 standards" and (ii) operational link to an acquired or parent company that is ISO-certified.

Conclusions

This study presents current baseline findings to gauge the seeming readiness of Canada's customs brokers to deliver the type of security-conscious brokerage services required in contemporary trans-border trade. Gauging readiness in terms of the brokers' involvement in the government's PIP program shows that they lag behind their U.S. counterparts (whose corresponding program is C-TPAT). A second gauge, acquired ISO certificates, suggests that there is parity rather than any lag. A third gauge, involvement in earlier expedited clearance initiatives, indicates that upwards of 65% of Canada's customs brokers have adopted these initiatives in their operations. Still, fewer than half of those 65% (44%) have exploited their apparent readiness to transition to PIP.

Worthwhile directions in which the research can be extended include the following. First, in using the ISO 9000 certification gauge, the study made no distinction among the four substantive classes in the ISO 9000 series: 9001-9004. Since certification in a given class does not necessarily indicate the same level of commitment to process and quality improvements as certification in another class, it would be interesting to conduct a follow up study that provides a more refined metric of quality/process improvement efforts. A second direction is to expand the study of NCBFAA members beyond the relatively small sample of 80 members. Finally, since the paper presents baseline results, it will be worthwhile to track how the results change over time, particularly as the newer initiatives such as PIP and FAST become more entrenched among the CBSA's trans-border policies.

References

Douglas, M. (January 2002), "A strong sense of security", *Inbound Logistics*, retrieved online.

Flynn, S. (2000), "Beyond Border Control", *Foreign Affairs*, Vol. 79, No. 6, pp. 57-68.

Foster, T.A. (May 2003), "Is NAFTA at Risk", *Logistics Management*, retrieved online.

Harps, L.H. (December 2003a), "Supply Chain Best Practices: Hitting the Mark", *Inbound Logistics*, retrieved online.

Harps, L.H. (October 2003b), "Managing Imports From Docs to Docks", *Inbound Logistics*, retrieved online.

Heizer, J. and Render, B., *Operations Management*, 6th Edition, pp. 173-174, Prentice Hall, 2001.

Helferich, O.K. and Cook, R.L., *Securing the Supply Chain.* Council of Logistics Management, 2002.

Russell, D.M. and Saldanha, J.P. (2003), "The Five Tenets of Security-Aware Logistics", *Transportation Journal*, Vol. 42, No. 4, pp. 44-54

Shanahan, J. (February 2004), "Secure from the start", *Logistics Management*, retrieved online.

Sheffi, Y. (2001), "Supply Chain Management under the Threat of International Terrorism", *The Journal of Logistics Management*, Vol. 12, No. 2, pp. 1-11.

Smyrlis, L. and Toews. B. (February 2003), "Choke point", *Canadian Transportation & Logistics*, retrieved online.

The Eurasian Land Bridge – An Economic Prospective of a New Transportation Corridor between Europe and China

Barry E. Prentice[1] and Lei Fan[2]

[1]Transport Institute, I.H. Asper School of Business and [2]Department of Biosystems Engineering, University of Manitoba

Introduction

The ancient Silk Road operated from eastern China, through central Asia to Western Europe for over 2,000 years. This trade route served as a great transmission belt of culture and technology until the 16th Century, when advances in sea transportation overcame the advantages of the overland trade corridor. The dramatic growth of modern China and other Asian economies are drawing attention back to overland routes, as competitors to containerized ocean shipping. The upgrading and modernization of the Asian rail network could have important economic and political consequences for world trade in the next 50 years.

The transport corridors of the Eurasian land bridge (ELB) comprise 27 countries and regions in Europe and Asia. The members of the Organization for Cooperation between Railways (OSTD) operate 216,000 kilometers of rail lines and over 2 million railcars (Somodi, 2004). China is playing a lead role in the development of the ELB rail routes. The Chinese government has outlined a number of programs to accelerate development of these trade corridors in its Five-year and long-term objective plans. The railway mainlines have been upgraded through the continental areas of China and into Central Asia. The links through the former Soviet Union that connect with the European railways are already in place, notwithstanding differences in rail gauge that need to be dealt with.

Five main rail routes are competing for traffic from Asia to southern and northern Europe. The entire rail infrastructure in Asia is being re-examined and demonstrations with block container trains are being tested (Somodi, 2004; Guozhi, 2004). The purpose of this paper is to examine the future of intercontinental rail in Asia. The paper begins with an historical overview of the ancient Silk Road. Subsequently, the economic prospects and obstacles to the Asian rail network are examined. The paper concludes with some thoughts on the future of the overland route and its implications for world trade.

The Silk Road: A Historical Perspective

The famous Silk Road was initiated under the Han Dynasty between 200 and 100 B.C. This overland trade route connected the more advanced Chinese culture with the Roman Empire in Europe. The Silk Road reached its peak during the period of Tang Dynasty (about 700 AD), but continued to be used as a corridor of communications until China was cut off from the rest of the world by the Ming Dynasty (1368-1640).

The 7,000 km long trade route became known as the Silk Road because silk was one of the major trade commodities. Other products shipped westward from China included furs, jade, ceramics, spices, paper, iron and bronze. The returning eastbound cargoes were gold, gems, ivory, glass, textiles and perfume. Traffic was limited to valuable luxury goods because the cost of transportation was high and warring tribes made travel dangerous.

Commodities were transported by donkey cart, camel, or on foot over less than well-defined roads. Merchants that traveled under armed escorts formed caravans of over 1,000 camels. Each camel could carry between 400 and 500 pounds[i]. The map of the ancient Silk Road that is presented in Figure 1 identifies several parallel routes that took advantage of different transportation gateways to Europe, India and Africa. The northern route from Chang'an (Xian) used the Jade Gate Pass northwest of Dunhuang, following the southern foothills of the Heavenly Mountains, and the northern rim of the

Taklamakan Desert via the oasis towns of Hami, Turfan and Kashgar. The more difficult but direct route used the Yangguan Pass southwest of Dunhauang along the northern foothills of the Kunlun Mountains and the southern edge of the Taklamakan Desert to Khotan and Kashgar.

Figure 1 Map of the Ancient Silk Road Routes



Source: www.imperialtours.net

At Kashgar, caravans could go west over the Terek Pass via Samarkand to Merv. Alternatively, the caravans could cross the high Pamirs and follow the Wakhan Corridor of Afghanistan to Bactria. A third route from Kashgar went over the Karakoram Pass into India. From Merv, the westward routes to Antioch, Babylon and Damascus transited relatively easier terrain.

The old Silk Road operated like a relay system in which goods were passed from hand to hand rather than as an integrated supply chain. "The Chinese traders escorted their merchandise probably as far as Dunhuang or beyond the Great Wall to Loulan, where it was sold or bartered to Central Asian middlemen - Parthians, Sogdians, Indians and Kushans - who carried the trade on to the cities of the Persian, Syrian and Greek merchants. Each transaction increased the cost of

the end product, which reached the Roman Empire in the hands of Greek and Jewish entrepreneurs." (China: Dim Sum, Internet)

The chaos that followed the fall of the Tang Dynasty disrupted trade on the Silk Road, while subsequent climate change (melting glaciers) terminated the water sources of many oases. By the end of the 8th Century, sea routes between southern China and the Middle East began to compete with the Silk Road. Of course the sea routes had to contend with unpredictable weather and pirates.

The history of Central Asia is steeped in conflict and violence. Only the presence of a great power, like Ghengis Khan (circa 1200 AD), did trade function in a secure political regime. The pioneering of a less expensive maritime route around the Cape of Good Hope by the Portuguese finally made the Silk Road obsolete for continental trade. The overland route continued to exist in a piecemeal form that linked the population centers and industry of Central Asia, but the splendor of these opulent cities that had grown rich on trade, faded into isolation and poverty.

In modern times, the Soviet Union created a 74-year hegemony that secured the interior of the former Silk Road, and led to the construction of rail lines. Suspicion of external influences and communist economic policy impeded through trade, while the central planners directed all commerce from Central Asia to the heart of modern Russia. As a consequence, the road and rail infrastructure of Central Asia has a greater north/south, than east/west orientation (Ostuka, 2002).

Main transport corridors

The Asian nations are being encouraged to play "their role in solving corridor issues on the basis of equity and mutual-benefit, mutual respect, constructive discussion and voluntary participation". (Guozhi, 2004) A number of transport corridors have been proposed and discussed by interested nations and international institutions. Five main Asian rail corridors were summarized in the Euro-Asian

Transport Conference that was held in Russia in 2000 (Tennenbaum, 2001).

1. The Northern Corridor, from Europe via the Tran-Siberian railway to China, North and South Korea, and Japan.
2. The middle Corridor, from southern Europe via Central Asia to China.
3. The Southern Corridor, from southern Europe to Iran, serving China, Pakistan, India and Southeast Asia.
4. The TRACECA corridor, from Eastern Europe via marine over the Black and Caspian Seas to Central Asia.
5. The North-South corridor, from northern Europe to Russia by rail, cross the Caspian Sea to Iran by barge and rail to the Arabian Sea and cross by ship to India.

Figure 2 presents the main Asian rail routes between China and Europe. The detail of this map is too fine to identify the differences in railway gauge, and this is dealt with in a later section. The important point to note is that multiple routes are possible and competition is being encouraged

Major initiatives and strategies for improving the Eurasian rail transport system have emerged from international conferences. Topics include the formulation of an integrated intermodal international transport network, the facilitation of transport at border crossings operational agreements for international transport routes, and the coordination of technical standards (Declaration of the third of international Euro-Asian conference on transport, 2003).

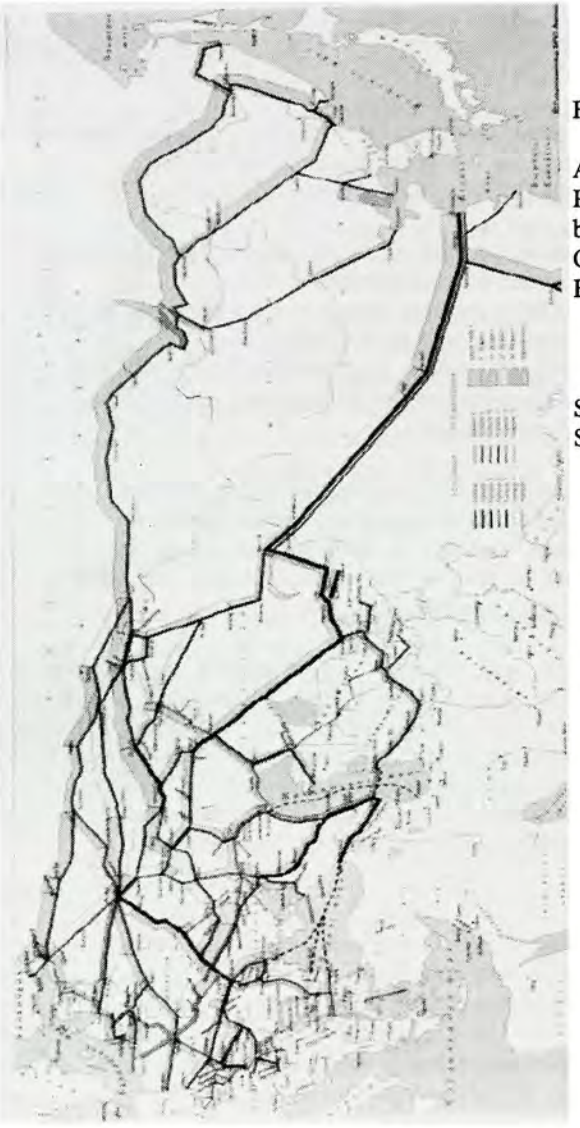


Figure 2

Asian Rail Routes between China and Europe

Source: Somodi (2004)

China, the ancient origin of the Old Silk Road, is once again leading the development of the new Eurasian land bridge. Various large-scale transport and relevant infrastructure projects have been launched. From the Chinese perspective, the ELB starts from the eastern port of Lianyungang and ends at Rotterdam in the Western Europe. The Longhai rail line that connects Shanghai in east coast to the western province of Xinjiang serves as the main artery to cross China's vast hinterland provinces. This east-west rail line crosses the border of China-Kazakhstan and splits into three separate lines, i.e. the northern, the middle, and the southern lines (Shu, 1997; Tennenbaum, 2001, Awad, 2001).

The first phase of China's new "Iron Silk Road" is becoming reality through a series of infrastructure projects. In 1990, the Beijiang Line that links Urumqi and Alashankou was completed. Some observers date this connection as the birth of the new Eurasian land bridge transport concept (Shu, 1997)[ii]. Since 1997, the Chinese government has invested 150 billion yuan/year for infrastructure program (Wen, 2001). A new railway segment was completed in western province Xinjinag, in 1999. This rail line connects Kashi in Xinjiang province to China's main railway network. The construction of a rail line through the Fergana Valley in Central Asia to connect existing track to Europe is being discussed as a next step.

Opportunities and barriers to trade

China is an important trading partner with EU. Total trade between EU and China reached US $125 billion and the trade turnover between China and central Asian counties increased to US $4 billion as of 2003.[iii] The majority of the trade between China and Europe is transported by ship. Ocean transport is continuing to increase in

productivity. Approximately 50 ships larger than 8,000 TEUs[iv] are in service in 2005. By 2008, this number is expected to reach 200.

Asian railways face a difficult challenge to compete with ocean transport, but have some advantages. Goods produced in hinterland sites have to be first transported to seaports and then shipped to Europe. The loading/unloading costs can be reduced, transport is less affected by weather, and there is demand for shorter transit time in comparison with ocean ship[v]. The sea transit time from Shanghai to Rotterdam is 28 days[vi].

Differences in railway track gauges are a major impediment to a fast rail link between Europe and Asia. Countries of the former Soviet Union use the width of 1,520 mm, while China uses the European gauge of 1,435 mm. The trains must change bogies or transfer passengers or goods to another vehicle when they cross the borders of the former Soviet Union. Transshipment and body changes require heavy equipment and cause inefficiency (Otsuka 2002).

The bane of the ancient Silk Road has its modern counterpart in the inspections and protocols associated with crossing sovereign borders. Customs formalities and inadequate facility and transfer capacity cause congestion to be an issue at border crossings. Fees and tariffs charged at border crossings add to the cost of international trade. Agreement is needed on operational aspects, as well as a tariff for combined transportation between Europe and Asia.

The problem of incompatible rail gauges in not confined to the east-west routes. Figure 3 lists the five different rail gauges that are used in Asia. The overland movement to the Indian sub-continent and southeast Asia are also impeded.

In theory, rail shipments could be much faster than marine, but in practice the transit times between Asia and Europe are relatively long. The rail route is more difficult to use than ocean shipping because no single agent can offer door-to-door service. In contrast, freight

forwarders can count on frequent, reliable and competitive sea services.

Figure 3 Railway Gauge of Asian Railways (mm)

1,000	1,067	1,435	1,520	1,676
Bangladesh	Indonesia	China	Amenia	Bangladesh
Cambodia	Japan	N. Korea	Azerbaijan	India
Laos	Philippines	S. Korea	Georgia	Nepal
Malaysia		Iran	Kazakhstan	Pakistan
Myanmar		Turkey	Krygyzstan	Sri Lanka
Singapore			Mongolia	
Thailand			Russia	
Viet Nam			Tajikistan	
			Turkmenstan	
			Uzbekistan	

In order to attract shippers, the railway supporters have recognized that a dedicated transcontinental service is needed. Somodi (2004) reports a demonstration run (December 2001) in which a trainset of 26 flatcars with 40 foot containers traveled from China, via Kazakhstan, Russia, Belarus and Poland to Germany. The distance of 6,190 km was completed in a bit over 8 days. Guozhi (2004) notes the successful debut in April 2004 of container block trains between Lianyungang (China) and Alamutu [Alma-Ata, the old capital of Kazakhstan]. This distance of 5,000 km takes 6 to 7 days.

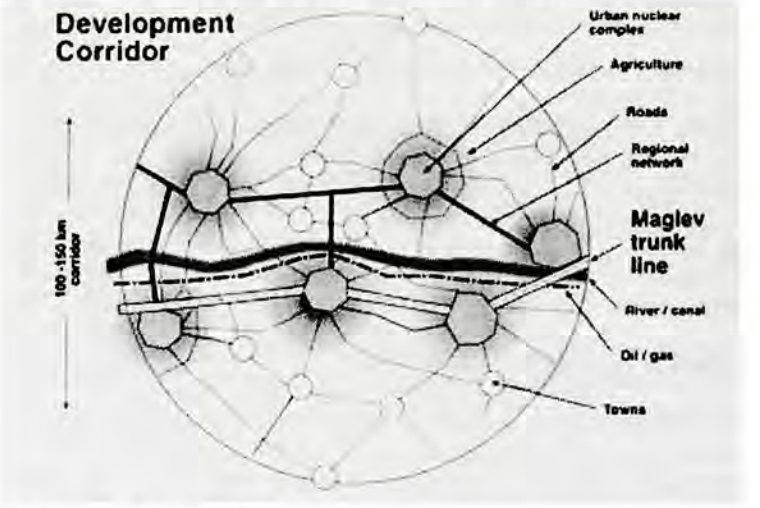
Future of the Eurasian Land Bridge

The definition of trade corridors and land bridges depends on the perspective of the participants. A corridor can vary in size from the width of a major transportation feature (e.g. railway) to the entire hinterland that is tributary to some transportation corridor. No exact designation exists, but what is common is the dependence on some major transportation infrastructure.

Logic suggests that the strength of a corridor would increase if it were served by parallel transportation infrastructures. For example, the rail, barge and highway systems associated with the Mississippi River or Rhine River valleys produce some of the world's strongest economies. Similar logic suggests that wherever transportation infrastructure can be co-located the trade corridor will be more robust.

LaRouche (1997) presents a conceptual framework for a futuristic infrastructure corridor model. The LaRouche transportation corridor that is illustrated in Figure 4 would be approximately 100 km wide and would comprise rail and road networks, a high capacity power line, natural resources pipelines, an advanced fibre-optics communication system, and ultimately a Maglev trunk line.

Figure 4 Futuristic LaRouche Transportation Corridor



Source: LaRouche, 1997

Whether the conceptual model envisioned by LaRouche emerges likely depends on the success of the Asian governments to overcome

the obstacles to efficiency and their ability to coordinate a continental-wide effort of modernization. In 2003, the first magnetic levitation train in Asia was put into operation in Shanghai. The creation of freight transport using Maglev technology anticipated by LaRouche would cost approximately US $1 trillion and take 10 to 15 years to build. At 450 to 500 kilometers per hour, a Maglev system would enable travel between Rotterdam and Beijing to be completed in one day.

The construction of a frictionless and vibration less high speed surface transportation from China to Europe may appear to be as far fetched as a bridge across the Atlantic Ocean, but it should not be discounted too quickly. A "green" technology that offers the capacity of rail with the speed of air cargo is very attractive. As trade and the economy of the Asia landmass increases, the financing of a Maglev system becomes more feasible.

Implications of the Eurasian Land Bridge

The landmass of the Central Asian republics (Kazakhstan, Turkmenistan, Uzbekistan, Tajikistan, and Kyrgyzstan) is twice the size of Europe. These countries are landlocked, which is a great disadvantage in world trade, and they are trying to re-orient their economies from the restrictions of the former Soviet Union. This area lies between two of the world's largest economies: China and the EU. Central Asia is rich in natural resources that could feed these markets, and sits on the critical land bridge that is required to the link them. One of the important implications of a new Eurasian land bridge would be the better inclusion of these countries in global trade.

In the short run, the development of land links between Asia and Europe is unlikely to have much impact on the competitiveness of countries in the Western Hemisphere. Longer term, the availability of a faster overland surface transport system between Europe and China could place North and South America at a competitive disadvantage in these markets. This is particularly the case if futuristic Maglev trains are built.

Conclusions

Everything that is old is new again. This is not just restricted to fashion and technology. Trade routes can also be reborn, but not without significant effort, investment and coordination. The overland route between China and Europe is a candidate for re-birth. Restoration of the historical Silk is a top priority in China. A rail based Silk Road from Liangyungang to Rotterdam is being implemented, albeit slowly.

About 75 percent of the Chinese section of the Eurasian railroad runs through the western region. The hinterland areas of China offer the prospect of lower cost labour, and less congestion than the coastal areas. Western regions of China are rich in natural resources and agricultural products, but economic development is held back by the lack of efficient transport links and facilities. The infrastructure arteries are considered to be essential preconditions for agricultural, industrial, and urbanization construction.

LaRouche's Eurasian Land Bridge proposal and China's Iron Silk Road program have received attention from numerous governments and international organizations. In the first instance, the Eurasian Land Bridge is likely to take the form of a containerized rail system. This could include a new section of track with a common gauge built through the former Soviet Union that connects Europe with China. In the longer term, Maglev technology offers intriguing possibilities.

Effective transport corridors are the key to the development of hinterland regions. The exchange of goods between China and Europe may provide the justification for investing in a Eurasian land bridge, but the greatest winners may be the countries of Central Asia. Opening up this land locked region to world trade will redraw the economic map of Asia and help this region capture its comparative advantage in the exchange of manufactured goods and natural resources.

References

Awad, G. S. "Proposals concerning the new silk road". *Executive Intelligence Review.* 2001: 22-26.

China: Dim Sum. "Silk Road Information". www.newton.mec.edu

Guozhi, Chang. "Promoting Eurasian Transportation Cooperation and Revitalizing Iron Silk Road." *ASEM Symposium on an Iron Silk Road.* Seoul, South Korea: 17-18 June, 2004.

Otsuka, S. Central Asia's rail network and the Eurasian land bridge. *Japan Railway & Transportation Review* 28 (2002): 42-49.

LaRouche, H. Z. "The Eurasian Land Bridge: The most important strategic question today". *Fidelio Magazine.* 1997. (excerpted in www.schillerinstitute.org)

LaRouche, H. Z. "The Eurasian land bridge as an alternative." *Seventh Conference on the Re-emergence of China. Institute of Sino Strategic Studies.* 2002.

Silkroad Foundation. www.silk-road.com

Shu, X. "The new Asia-Europe land bridge – current situation and future prospects". *Japan Railway & Transportation Review.* 23 (1997): 30-33.

Somodi, Kalman. "Eurasian Railway Freight Network in all-day operation: 50 years experience in co-operation of OSID members". *ASEM Symposium on an Iron Silk Road.* Seoul, South Korea: 17-18 June, 2004.

Tennenbaum, J. "The new Eurasian land bridge infrastructure takes shape". *Executive Intelligence Review.* 2001.

Declaration of the Third of International Euro-Asian Conference on Transport 2003. www.unece.org/trans/cd/peter3e.pdf (2004/12/2).

UIC 2002. International Freight Corridors-Eurasian Land Bridge. www.uic.asso/ (2004/11/28).

Wen, T.J. China's experience in building its economy. *Executive Intelligence Review.* 2001.

---

[i] The two hump Bactrian camel was preferred over the one hump camel. Both could carry the same weight but the Bactrian camel

could maintain a faster pace.
http://www.newton.mec.edu/Angier/DimSum/Silk.html

[ii] The Trans-Siberian Railway could claim this honour because it was completed in 1905. However, this railway served mainly intra-Russian and Soviet trade.

[iii] China Daily, March 16, 2004; August 27, 2004

[iv] A TEU is a twenty-foot equivalent unit. A 40 foot long container is counted as 2 TEUs.

[v] Abu Dhabi advertises its air-sea intermodal combination as saving one-third of the transit time from the Far East to Europe. Freight rates and volumes are not readily available.

[vi] Lykes Lines

Barry Prentice is a Professor of Supply Chain Management at the I.H. Asper School of Business, and the Director of the Transport Institute at the University of Manitoba. His research and teaching interests include logistics, transportation economies, trade and development. **Lei Fan** is a Chinese national and a graduate student in Biosystems Engineering at the University of Manitoba.

Determining Relationships between Regional Characteristics and Paratransit Demand

Peter Allaby, MASc. Candidate
University of Waterloo

Introduction

Ensuring an equal opportunity for the mobility of all citizens is a very important issue for communities across Canada. Many citizens, of all ages, have disabilities which prevent them from driving their own vehicle or being able to use fixed route public transportation. Because of this, they have no means of access to jobs, to health treatment centres, to church, or to stores to buy their basic goods. To answer this problem, and to foster community inclusion, specialized transit systems, or Paratransit systems, have been adopted. Based on a demand response system, paratransit can provide door-to-door service for persons with disabilities, by means of vehicles specially designed to meet their needs.

The use of paratransit is growing throughout North America, as agencies realize the importance of providing equal service to all users. This is becoming ever more important as our population is aging, and the number of seniors is quickly increasing. As of 2001, approximately fifty regions in Canada were providing specialized transit service. The regions range in population from 20 000 to 2.4 million, and in service areas from 36 km^2 to 3600 km^2. Information from each of these regional services can be used to determine various relationships between population, service area, and the service provided.

This paper aims to identify what community characteristics drive user demand, by developing regression relationships between independent community characteristics and measures of paratransit demand. Once the dominant characteristic is determined, regression equations can be formulated to help predict current or future demands. Also, this study aims to determine if a unique relationship exists between the elderly

population and the number of paratransit users. As our population is quickly aging, finding a relationship between age and demand could be extremely important for accurate and adequate planning in meeting the mobility needs of our seniors.

Paratransit Background

Paratransit services are provided by a variety of modes, including taxicabs, vans, mini-buses or special trolley systems, and all vehicles are modified to be fully accessible with features such as wheelchair lifts or low floor entry. Generally, paratransit offers curb-to-curb delivery, as many of the eligible users are unable to walk even short distances. Also paratransit is meant to be complementary to fixed route transit service, meaning similar service areas are covered.

Both demand-response services and ride sharing fall under the scope of paratransit. Demand-response paratransit generally operates under one of two principles – "many to many" service, which provides transportation from a user-defined origin to a user-defined destination at a user defined time, and; "many to few" service, through which transportation is provided to predefined zones or common destinations. "Many to few" service is common among the elderly for transportation to regular destinations such as local clinics, grocery stores, and pharmacies. Ride sharing refers to a service through which more than one user is accommodated in one trip. Scheduling the delivery of multiple users over one trip helps to improve efficiency of the system by reducing mileage and the number of vehicles required.

To be able to use a paratransit service, an individual must complete a needs assessment to determine their eligibility for the service. The American Disabilities Act describes paratransit eligibility as "*a functional determination of a person's ability to use the regular transit system as it currently exists, and not simply a medical or psychiatric diagnosis*" (1). An eligible individual will fall into one of three categories:

1. Individuals who are unable, because of a physical or mental impairment, to independently board or ride an accessible vehicle in a fixed route transportation system
2. Individuals with a physical or mental impairment who could use an accessible vehicle, but the vehicle is either not provided or temporarily not in service.
3. Individuals who are unable, because of a physical or mental impairment, to travel to the boarding location or from the disembarking location of a fixed route system.

The second category is the broadest and can result in some paratransit users only being temporarily eligible or eligible only for certain services and not others. For example, certain routes may provide accessible vehicles while others do not. It is important to note that in the United States, eligibility standards are controlled by the American Disability Act, a nation wide standard. In Canada, accessibility standards and criteria are controlled on a provincial basis. Therefore, eligibility standards may differ from province to province.

An Aging Population

In North America, the current status of our population profile shows an increasing median age and drastic increases in older age bracket populations (3). This trend can be attributed to high birth rates during the 1920s, WWII, and of course the baby boom between 1946 and 1965. In Canada, the number of people age eighty and older has increased by 41% since 1991, and is expected to increase by another 43% by 2011. Baby boomers who are now between the ages of 39 to 58, will certainly add an even greater increase over the coming years as they near the elderly age bracket (See Figure 1). The aging of our society is a cause for concern when planning for the future strain on health care, home care, and social security; and the future strain on the mobility needs of our senior citizens is no exception.

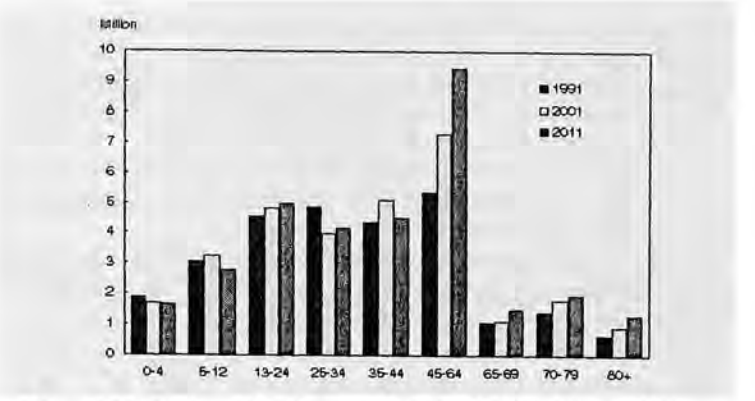


Figure 1: Population, selected age segments, Canada, 1991, 2001 and 2011 (Stats Canada 2001)

Senior citizens have several options for modes of transportation. They can use private automobiles, either as drivers or passengers, public fixed route transportation, paratransit, or other forms of specialized transportation services that may be provided by community organizations or through governmental assistance programs. The use of private vehicles remains the most common mode of choice, as many seniors retain the capability to drive themselves or have friends and family who can drive them; but, for those who are not capable of driving, or who do not have an income which can sustain the cost of maintaining an automobile, public transit is their main alternative.

Intuitively, one assumes that an elderly person is more likely to acquire a physical disability than a younger person and thus the proportion of those paratransit-eligible among seniors should be higher. A Paratransit Growth Management Study performed by the Orange County Transportation Authority (OCTA) in California (4) states that nearly half of the seniors in Orange County have or will develop two or more chronic illnesses affecting their mobility. Also, the study predicts that 10% of the senior population will require specialized transit services. The study summarizes that over the next

twenty years, the aging baby boomer generation will increase the potential need for transportation assistance.

If the above discussion is correct, a unique relationship should exist between the elderly population and the number of paratransit users; however, although a higher proportion of seniors may be eligible for paratransit, this does not mean they will choose paratransit. A study by Franklin and Niemeier in 1998 (5), provides results relevant to this issue. The study examined the choice elasticities for elderly and disabled travelers between fixed-route transit and paratransit in Sacramento, California. Four variables common to both systems were chosen to analyse their effect on choice of mode. These variables were sex, age, fare, and income of user. The results of the study showed that fare and income produced negative inelasticities, and age was found to be negatively elastic, with an elasticity of -1.612. This is a significant result. It means that as individuals age by 1%, the probability that they will choose paratransit decreases by 1.6%. Also, the study found that females are more likely to choose paratransit than males.

Considering the discussed research findings, it is difficult to determine conclusively whether an aging population alone is means for increases in demand. Clearly, higher numbers of seniors means more individuals requiring mobility assistance, but will paratransit be their choice of service? The following section of this report will examine empirical evidence to show what relationships exist between Canadian community profiles and their respective paratransit demand, and determine if a unique relationship is found between the elderly and registered users.

Method of Analysis

At the onset of this study, a database was obtained containing a vast amount of information on the paratransit systems in Canada. The database (compiled by the Canadian Urban Transit Association) included approximately fifty different regions, and reported on information such as regional population and service area, as well as on over 150 categories of paratransit service information. Service information categories most important to this study were:

- Total Users – Users registered for the paratransit service.
- Revenue Vehicle km – Annual vehicle kms traveled by revenue vehicles during regular passenger revenue service.
- Total Passenger Trips – Annual number of passenger trips.

The following assumptions were made at the onset of the study.

- Provinces set similar accessibility standards.
- All paratransit services provide an acceptable quality of service for their regions. In other words, basic service is being made available to all registrants that meets their needs. This assumption had to be made to ensure that all regions can be analysed on the same playing field.

The first stage of the study was to determine the basic relationships between the regional characteristics and the respective paratransit demand information, to find the most influential independent factor across all systems for determining demand.

The second stage of the study was to draw relationships between the age profile of a region and the registry of its paratransit system. To obtain the age profile information, relevant data was extracted from the Statistics Canada 2001 Canadian Census (3). Population data for those over 65 years were obtained, separated into 10 year age brackets, and separated by sex. The aim of this stage was to try to identify whether the elderly population had a stronger correlation with users than did other ages.

All analyses were performed through simple linear regression. The system demand categories were chose as the dependent variables, while the independent variables were regional characteristics, ie. population, service area, or service area population density. The main output sought through the regression analyses was the coefficient of determination, r^2. A perfect correlation is indicated by $r^2=1$, while smaller r^2 values (i.e. below 0.4) indicate a weak linear relation (6). The regression equation can also be noted, as it can be used to make general projections relating to the future growth of regions.

Results and Discussion

Community Characteristics vs. Demand Categories

The regression analyses performed for the first stage of this study produced correlation coefficients which are displayed in Table 1. These are described in more detail below.

Table 1: R^2 Values from Regression Analyses

		Dependent Variables		
		# of Users	Passenger Trips	Rev Veh. Kms
Independent Variables	Population	0.83	0.90	0.73
	Service Area	0.09	0.09	0.10
	Service Area Population Density	0.38	0.44	0.30

1. Number of Users: Prior to performing the regressions, it was expected that the number of users registered in each region would relate strongly to its population. This was expected because it is fair to assume that across a single, developed country, the proportion of persons with disabilities will remain constant. The r^2 value of 0.83

indicates a fairly strong relationship, one much stronger than the r^2 value of 0.38 for population density.

2. Passenger Trips: The relationship between passenger trips and population is very strong at 0.90, not far from a perfect correlation. Unlike other public transit services, where population density is considered to affect ridership, the 0.44 r^2 value indicates a fairly weak relationship for paratransit. This could be attributed to the fact that a paratransit system relies on the volume of people registered in the system, and not on the concentration of users living in the vicinity of the transit corridor.

3. Revenue Vehicle Kms: The r^2 values show that mileage has a reasonably strong relationship with population, but very weak relationships to population density and service area. This may be surprising since one might expect areas with sparcely distributed users to lead to higher mileage. To further inspect this relationship, another test was performed, between population density and kilometers/passenger trip. It was felt that using this variable may be more appropriate in identifying a relationship. The results did show a downward trend in mileage for higher densities; however an extremely weak correlation of 0.06 exists. This reinforces the previous results that mileage does not depend on population density or service area.

From the results, population can be considered to be the most influential factor in predicting paratransit usage. The regression equations derived from each analysis can be applied to a given region in Canada to predict the expected requirements of a paratransit system. The equations are displayed in the table below.

Table 2: Derived Regression Equations

Dependent Variable	Regression Equation
# of Users	$E(Y\|x) = 553.4 + 0.009847x$
Passenger Trips	$E(Y\|x) = -2269 + 0.6459x$
Revenue Vehicle Kilometres	$E(Y\|x) = 49567.51 + 3.2746x$

It is very important to note that the validity of these equations only applies to regions with populations in excess of a lower limit, roughly 10 000 to 20 000 people. In an attempt to avoid a negative intercept for passenger trips, another equation was calculated, forcing the y-intercept to zero; however, this produced unreasonably inflated expected values, so the equation was deemed inappropriate and the original equation remains. The above problems suggest that at the lower end of the data (smaller populations), the line of best fit may not be linear, but resembles more of an exponential shape. The problems may stem from having no data available for regions with less than 20 000 people.

Having equations such as the ones above could be useful to transit planners, who are (a) proposing to introduce a new paratransit system into a region, and wish to have an idea of what to expect on par with the rest of Canada; or (b) planning for the future demands of an existing system, resulting from population growth.

Age vs. Paratransit User Registry

In this stage, for the purposes of simplicity, the only paratransit demand measure tested was the number of registered users.. In addition, the only population characteristics considered were age and sex, as aging individuals have a higher likelihood of developing problems with mobility, and previous research has shown women are more likely to choose paratransit than men.

Age data was extracted from the Statistics Canada website. Since some of the regions' service areas exceeded or did not coincide with the municipal boundaries set forth in the census, several regions had to be omitted from this section of the study. Compatible information was found for 38 regions. The census data were divided by sex, and into the following age brackets:

- Ages 65-74
- Ages 75-84

- Ages 85 and over

The same procedure was used for this analysis as in the previous stage. Regression calculations produced an r^2 value, which was used to determine the strength of correlation between variables. Age and sex were independent variables, while number of users was considered dependent. Analyses were performed not only for each individual age bracket but for combined brackets, in order to find the range that has the strongest correlation. The table below summarizes the results obtained from the analyses.

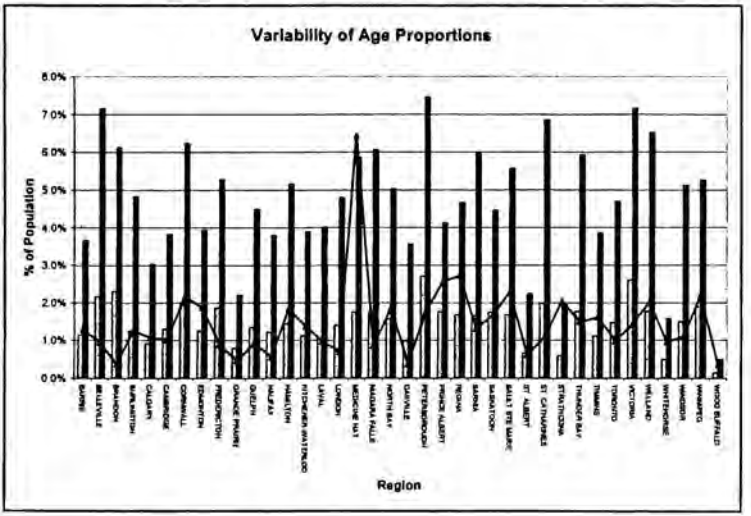
Table 3: R^2 Values for Age/User Relationships

Age	Male	Female	Overall
65-74	0.85	0.85	0.85
65-84	0.85	0.86	0.86
65 and over	0.86	0.86	0.86
75-84	0.85	0.87	0.86
75 and over	0.86	0.86	0.86
85 and over	0.84	0.84	0.87

The r^2 values in the table above show a strong relationship with user registry across all age brackets. There is little discernable difference in r^2 values between any of the brackets, or between sexes. Also, there is only a slightly stronger relationship between these age populations and registered users than the relationship for the entire population. To further investigate the comparison between the elderly age bracket and the rest of the population, regression analyses were performed for the age brackets of 45-54, and 55-64. The results indicated r^2 values of 0.87 and 0.86 respectively. This information suggests that there is insufficient evidence to conclude that a unique relationship exists between the elderly and the number of paratransit users; additionally, the results indicate the lack of a unique relationship between sex and paratransit users. This does not agree with the results from Franklin and Niemeier (5) which state females are more likely to choose paratransit.

It was thought that a possible cause of the lack of relationship corresponding to age could be attributed to constant proportions of age populations among communities, or in other words, insufficient variability of age bracket proportions across Canada. For example, if all communities experienced an 85+ population of 5%, then one should expect no deviation from the correlation found between total population and user registry. To examine the existence of age proportion variability across regions, the proportions of age brackets 85+ and 75-84 were plotted for each region. The resulting graph is shown below in Figure 2.

Figure 2: Age Proportion Variability across Regions



- **Blue Columns: Age 85+ Proportion,**
- **Red Columns: Age 75-84 Proportion,**
- **Black Line: User Proportion**

Since there appears to be a reasonable degree of variability in age bracket population proportions across all regions, further possibilities

should be explored which explain the lack of relationship between age and user registry. These possibilities may include the following:

- Elderly individuals with mobility problems may have other sources of transportation that better suit their lifestyle.
- Younger individuals, who seek to be more active, may be more apt to use paratransit services, thus evening out the age scale of users.
- Elderly people sometimes have the attitude of "not wanting to be a bother" and therefore do not register with paratransit on that principle.
- Paratransit service operators may not be sufficiently reaching out to the elderly population who are sometimes timid of using new systems.
- The approach of analysis in this study may be too simplified to identify a unique relationship. More complex approaches may be necessary.

All but the last suggestion support the results from the study by Franklin and Niemeier – although seniors may have greater mobility needs, they are not necessarily choosing paratransit as a transportation option.

Conclusions

The goal of this study was to examine the state of the paratransit system in Canada and to determine the most influential regional characteristics on paratransit demand categories. At the onset of the study, it was thought that a unique relationship may exist between the elderly population and the number of paratransit users. This relationship would be useful in predicting future strains on paratansit in Canada as our elderly populations are rapidly increasing. The following summarizes the results of the study:

- A very strong correlation exists between population and paratransit demand measures such as registry, passenger trips, and revenue vehicle kms,

- Service area population density showed weak correlation with paratransit demand measures.
- It is inconclusive to suggest a unique relationship exists between the elderly population and the number of paratransit users.
- It is inconclusive to suggest a unique relationship exists between male/female populations and the number of paratransit users.

The results of this study can be quite interesting. The strong influence of population (and not population density) may pique the interest of rural areas proposing to implement paratransit, where population density is a concern. The weak dependence of paratransit on population density may be a cause for further investigation into the role of population density in planning specialized transit services.

As population growth trends are somewhat easy to project, regression equations could be very useful for projecting future demand and costs for paratransit; however, prior to seriously employing any of the regression equations derived from this study, a more in depth analysis should be performed on the lower end of the regression line to determine if a linear approximation is accurate. Also, an important factor for transportation planning not included in this study is land-use. The layout of regions may play in integral role in paratransit services. For example, for many-to-few services, how close are residential areas to health clinics, pharmacies, or grocery stores? Land use cannot be quantified like population or service area, but is very important and may need to be examined on a case by case basis to determine the strength of its influence.

Finally, the fact that age did not seem to play a stronger role in predicting user regisitry was somewhat unexpected. As mentioned in the discussion of results, a more complex approach should be investigated in relating the age and user information. The approach in this study may not have modeled the data in a way to capture a unique relationship. The issue of our aging population is still a very

important one, and more work should be performed using similar data to develop more reliable relationships and prediction models.

REFERENCES

1. Paratransit Eligibility – Draft. ADA Transportation Series, Federal Transit Administration [Internet]. [Cited October 25, 2004]. http://www.ada-infonet.org/documents/transportation/paratransit-eligibility.asp

2. Denson, C.R. 1998. *Transitioning to Fixed-Route Services.* Transportation Research Record No. 1623. Transportation Research Board, National Research Council, Washington D.C. Pg 37-44.

3. Statistics Canada 2001 Census. *Profile of the Canadian population by age and sex: Canada ages* [Internet]. July, 2002. [Cited October 24, 2004] http://www12.statcan.ca/english/census01/Products/Analytic/companion/age/contents.cfm

4. Orange Country Transportation Authority. *Paratransit Growth Management Study* [Internet]. [Cited October 21, 2004] http://www.octa.net/busrail/access/part1.pdf

5. Franklin, J.P. and Niemeier, D.A. 1998. *Discrete Choice Elasticities for Elderly and Diabled Travelers Between Fixed-Route Transit and Paratransit.* Transportation Research Record No. 1623. Transportation Research Board, National Research Council, Washington D.C. Pgs 31-36.

6. Khisty, J.C. and Mohammadi, J. 2001. *Fundamentals of Systems Engineering.* Prentice Hall, New Jersey. Pg 275.

Public Transportation: A Solution to Social Exclusion

Sean Nix

Edited for the CTRF – February 2005

Research originally conducted for the School of Environment and Development, Sheffield Hallam University (United Kingdom) April 2004

Executive Summary

Social exclusion is a newly emerging term given to specific groups of people who suffer from a series of social problems. This secondary research paper will first establish the definitions of social exclusion and those who suffer from this problem, then look at the relation between social exclusion and a lack of public transportation policies in the United Kingdom. The paper will also discuss what needs to be improved to resolve this growing problem through reviews of specific recommendations by Transport 2000 (UK), the independent national body concerned with sustainable transport; the Social Exclusion Unit, a division of the Office of the Deputy Prime Minister; and various academics. Modest successes of local Canadian examples are mentioned and discussion is provided on how they could be applied to the United Kingdom model as an interim and incremental solution to social exclusion.

Introduction

The term *social exclusion* can be a factor and a result of many of society's great problems. When it comes to conventional public transportation, social exclusion is a result of poor public transportation policies and yet, at the same time, is an important factor to address when planning for public transportation. Improvements to public transportation can provide various benefits towards dependant social groups. A properly planned conventional public transportation system can be an effective solution to social exclusion both in rural and urban areas. However, when taking social exclusion into account as one plans a public transportation system, it is essential to identify the specific transportation-related problems of those who suffer from social exclusion in order to establish appropriate solutions.

This paper will provide definitions to the term *social exclusion*, then discuss the various transportation-related problems associated with social exclusion and those who suffer from it, noting the distinction between problems in rural areas and those in urban areas. Some simple yet effective solutions to those problems will also be reviewed.

Social Exclusion Defined

There are many different definitions of social exclusion. The most common associated definition of social exclusion relates to poverty; a tenure described as the 'have not' social groups. Such groups are "excluded" from society's main elements such as education, employment, healthcare, shelter, leisure, retail, culture, and religion. The Department for Transport Local Government and the Regions (DETR) in England provides a recent definition which asserts that social exclusion can happen when individuals or areas suffer from a combination of problems such as unemployment, poor skills, poor housing, high crime environments, bad health, and family

breakdowns.[1] Transportation plays an important role in the solution to social exclusion in that it provides *access* to these elements. The main options for the tens of millions of people without access to a car – public transportation, walking and cycling – are often expensive, inadequate, unpleasant or dangerous (as traffic volumes have increased, cycling and walking have become more dangerous), while distances to services have also increased over the last two decades – as witnessed by the growth of out-of-town superstores at the expense of local shops.[2] In order to have access to these and other of society's essentials, everybody must have ready access to adequate and affordable transportation, both private and public.

The Socially Excluded

It is not enough to simply provide transportation without careful consideration of what the mode or form is addressing. It is important to identify the individual groups that are considered as socially excluded and their specific needs, especially those without access to proper transportation. Such groups include the elderly, the physically disabled, students, children (and their mothers or carers), women in general, the unemployed, and those living in areas where the local transportation system does not serve them adequately. The elderly and physically disabled cannot walk as fast or far, so proximity to adequate public transportation is vital. These general problems become everyday ones when aged people have to use (transit) systems that, in many cases have been designed without consideration

[1] McKoy, Keith. "The Social Impacts of Transport: Access Needs and Mobility Characteristics of Different Social Groups (The Transport Deprived)". Transport Growth and Change Lecture Series – Sheffield Hallam University. 19 Mar. 2004.

[2] Friends of the Earth. *Environmental Justice: Mapping Transport and Social Exclusion in Bradford.* London: 2001.

for their special needs.[3] Students and the unemployed usually require the cheapest mode of transportation; in many cases public transit is that cheap alternative. Lack of information about public transit opportunities can reduce the take-up of job interviews and employment amongst the unemployed, while unreliable services can also reduce take-up of employment opportunities situated on the urban periphery amongst those who do not have access to a car.[4] Children and youth under sixteen are physically (and mentally) unable to drive, so public or shared private transportation is the only mode of transportation available. Poor public transportation, especially school transportation, in an education policy environment of 'parental choice' can reduce the educational choice and opportunities of children in those households without a car to the most local; a problem if the school is 'failing' through the impact of the local social environment.[5] Women and infant carers have special transportation needs that usually require transit away from the traditional rush hour periods when public transit is more frequent. A key difference in the developed world is the propensity of women to combine a set of activities relating to their extensive range of household tasks within the overall structure of one journey period (trip chaining) whereas men are more likely to make a single purpose trip.[6] While the majority of households own a private automobile, at least 50% of each of those households is likely to fall into the listed social groups, and thus require some alternate mode of transportation. A properly planned and managed public transportation system can be the appropriate solution.

---

[3] McKoy, Keith. "The Social Impacts of Transport: The Transport Deprived". The Geography of Transport Lecture Series – Sheffield Hallam University. 21 Jan. 2004.

[4] Quinn, D.J.. "Accessibility and Job Search: A Study of Unemployed School Leavers". *Regional Studies*; 20:2 (1994): 163-173.

[5] Turner, J. and M.S. Grieco. "Gender and Time Poverty: The Neglected Social Policy Implications of Gendered Time, Transport and Travel." Proceedings of the International Conference on Time Use – Luneberg, Germany. 22-25 April, 1998

[6] Ibid

Social Exclusion in Rural Areas

In rural areas retail is being lured away to the larger modern shopping complexes and city centres, forcing these residents to travel further to obtain essential everyday items. It is much more difficult to serve the socially excluded by public transportation in rural areas due to the dispersed rate of destinations from one particular origin point. Typical suburban public transit systems prefer to operate from one nodal development to another, where groups of passengers can be captured onto the bus and individual passengers can be served along the way. Unfortunately, most rural transit services operate at a bare minimum frequency which does not offer an attractive service to anybody, nor is it efficient for the socially excluded. This is particularly a problem among rural areas in Canada where, in certain cases, public transportation linking rural communities is the responsibility of the private coach industry. To contrast, public transportation in urban areas in the United Kingdom is solely the responsibility of private bus companies, but the government will establish and fund rural services through a tendering process with those bus companies. Two specific United Kingdom bodies, Transport 2000 and the Social Exclusion Unit (a branch of the Office of the Deputy Prime Minister), believe that demand-responsive transport with door-to-door options is the answer. Transport 2000 recommends a dial-a-ride taxi-bus service for rural areas where passengers phone a booking line 30 to 60 minutes before they wish to travel. These services run from early in the morning until late until late at night or around the clock seven days a week, and regular bus fares apply in most cases.[7] The Social Exclusion Unit suggests introducing flexible routes into the network where demand is more dispersed and directed to where users live or want to go within a defined area, all of which can be supplied by buses, community transport vehicles or taxis.[8] The Social Exclusion Unit also believes

[7] Transport 2000. *Rural Transport Futures: Transport Solutions for a Thriving Countryside (Summary).* London: 2003.

[8] Social Exclusion Unit. *Making the Connections: Final Report on Transport and Social Exclusion.* London: Office of the Deputy Prime Minister, 2003.

in making use of the existing lack of rural services by integrating schedules between services; adjusting schedules to allow all trips of any or all modes to connect at a central location. One Canadian town has an efficient example of these recommended services. In the Greater Toronto Area suburban Town of Oakville, Ontario, the local transit authority has recently introduced a dial-a-bus system on Sundays. At 25 minutes past the hour, they arrive and depart from the main rail station to guarantee connections with the hourly commuter rail service to and from Toronto. Using "airport style" shuttle buses, *Zone Express* dispatch develops unique routes based on customer pick-up and drop-off points; passengers are picked up at the closest bus stop to their origin and dropped off at the closest bus stop to their destination.[9] Zone Express has proven to be quite popular with passengers since it can at times provide a more direct service between their origin and destination, so popular that Oakville Transit replaced its weeknight and Saturday evening route-based service (with the exception of two popular routes) with Zone Express. The service is also much cheaper to operate since only four mini-buses are used during these low-demand periods, while covering the same area as multiple standard-sized buses on a route-based system. This is just one simple, yet effective way to address transportation-related social exclusion problems in rural or suburban areas where larger-scale public transit initiatives are either infeasible or unaffordable by local governments.

Social Exclusion in Urban Areas

Urban areas have a more established and accessible public transportation system. However, they still lack some basic provisions for those who have special needs in that they are usually planned for those who are physically able and financially well off. In order for

[9] Oakville Transit. "Zone Express: Zone Bus System" [Online]. (Mar. 2004). Available: http://www.oakvilletransit.com/schedules/zonebus.asp (Accessed on March 30, 2004).

public transportation systems in urban areas to address social exclusion, they must be easily accessible and affordable. In all urban areas of significant size, the public transportation system needs to be multi-modal, using attractive, modern low-floor buses feeding passengers to and from high-quality mass-transit light rail or modern tramway lines serving the main arterial corridors.[10] Buses and trains are often inaccessible to the elderly and disabled because they are not level to the platform or sidewalk, and distances between each mode can sometimes be great. With the recent vehicle renewal movement, bus and train operators are building their fleets with new low-floor vehicles. In Canada, the Province of Ontario introduced the *Ontarians with Disabilities Act, 2001* which requires public transportation organizations in the province to prepare an accessibility plan and consult with persons with disabilities and others in preparing the plan.[11] Many public transit agencies in Ontario have responded to this act by purchasing low-floor wheelchair accessible buses or train coaches, many of which compliment the transit agencies' special wheelchair shuttle services. Low-floor (buses and trains) are convenient for wheelchair users and for the blind and partially sighted as there are no steps to negotiate or fall over when entering the bus (or train) from the pavement.[12] Proper terminals, intermodal facilities, and transfer/ connection points can also be beneficial to the elderly and disabled. Not all trips on public transportation in urban areas are made on single modes or services; connections from bus to bus, bus to train, train to taxi, and many others are very common, and therefore must have proper facilities to ease these transfers and connections, particularly to those who would choose to drive as a result of these lack of proper connections but

---

[10] McKoy, Keith. "The Meaning of Transport Integration". Transport Growth and Change Lecture Series – Sheffield Hallam University. 13 Feb. 2004.

[11] Bill 125: *Ontarians with Disabilities Act, 2001*. Royal Assent, December 14, 2001, 37th Legislature, 2nd Session, 2001, Legislative Assembly of Ontario, 2001.

[12] Parry, Tamsyn. "Transport for the Visually Impaired: A Look at Urban Britain". *World Transport Policy and Practice*; 1:4 (1995) 25-28.

cannot due to their own disabilities. Modal (and service) integration can enable easy transfer between different modes because of the close physical location.[13] Another way to improve connections between trips of a single mode or trips of multi-modes is integrated ticketing. Integrated tickets allow passengers to use the same ticket on services operated by different transit agencies; it makes travel more affordable as well as saving time for passengers.[14] In jurisdictions where a separate fare is required, this fare collection system can be beneficial to those with lower incomes, such as students and the unemployed, who cannot afford to pay additional fares when changing services to continue a trip. All of this can contribute to a fully intermodal accessible urban public transportation system.

Conclusion

Improvements to conventional public transportation can contribute great lengths to the reduction of social exclusion. Everybody has different needs, and the individual needs of society need to be addressed when providing social services to people, even in public transportation. It is not enough to merely fix the largest problem (such as wheelchair accessibility) and call public transportation an efficient social service; if certain people are unable to use public transportation for any given reason, they are not only denied access to public transportation, but also to the various services and destinations that public transportation provides access to. People deserve access to society's necessities, and the suggested improvements discussed simply serve as small incremental steps towards an overall, fully integrated public transportation system which can greatly increase

[13] McKoy, Keith. "The Meaning of Transport Integration". Transport Growth and Change Lecture Series – Sheffield Hallam University. 13 Feb. 2004.

[14] Social Exclusion Unit. *Making the Connections: Final Report on Transport and Social Exclusion*. London: Office of the Deputy Prime Minister, 2003.

access for the socially excluded. As more improvements are made, public transportation can be accessible to many more people, and they will be able to access all of society's essentials through the improved public transportation system. The reduced effects of social exclusion will lead to a better quality of life for all members of society.

Developing Media Partnerships for Crisis Communications During Transportation Disasters

Jill A. Lynn, Assistant Professor,
Aviation Technology, Purdue University
lynnj@purdue.edu; 317-381-6075

Abstract

A modern challenge faced by transportation professionals today is our relationship with news media. The 24-hour news cycle and the growth of "fringe" media are phenomena that are here to stay. As transportation professionals we must learn to control, or at the very least, effectively manage these phenomena with confidence and finesse, especially when relates to transportation incidents and accidents. The majority of media should not be viewed as the enemy, but as potential partners who can assist with disseminating critical information to the general public and specific populations (victims and their families, governments, etc.) who have an interest and/or stake in the event that has occurred. Currently viewed by many organizations as "just another function" of their public relations department, research in the areas of communications and mass media has discovered that the communication processes and techniques during disasters (crises) of any type are highly specialized. Traditional public relations skills established one's base knowledge and skill, but specially trained crisis communications personnel are needed to create and maintain cooperative partnerships with the media and public, to control the message in an effort to reduce unsubstantiated accusations, misinformation (guesses), and panic.

A Purdue University aviation technology researcher has been with working communications and media professionals to develop corporate advocacy and crisis communication programs and processes to handle the delicate and demanding area of disaster (crisis) communications for the transportation industry. The paper outlines the affects of using corporate advocacy programs to develop positive and trusting relationships with media (and other populations)

during times of normal operations, and how these partnerships are used to handled crisis communications in a controlled and positive manner. In addition, the need for specialized training in the area of crisis communications, and programs to provide this training is discussed. Crisis communications cannot be left to chance or to the general skills of public relations departments to establish effective and positive outcomes. Due to the 24-hour, on-demand news we must take a proactive approach in establishing communication processes and programs to provide accurate and timely information to maintain the public's trust during a time of crisis. The inability to properly manage crisis communications can lead to public mistrusted, unwarranted suffering, and devastating economic problems for an organization.

Introduction

Crisis communication has become more prominent for corporations and other organizations over the past two decades. In the past, most incidents and / or accidents would generate local to regional media interest only, however with the emergence of the 24-hour news cycle and various types of news mediums being created at a quick pace the luxury of minimal or limited response by organizations is gone. All organizations must develop a crisis communication program (which includes training for personnel) in order to meet the increasing demands of the media and the "public's right to know" doctrine that is the common battle-cry of today's journalists.

No matter what an organization's image, if they become involved in a disaster their actions and responses to the public and media during this time of crisis will have long term effects that can be positive or negative. This issue of image in crisis is not just a matter of disclosing information and assisting in the recovery and healing process, but also one of maintaining public trust. If an organization attempts to protect its image during a crisis by choosing to withhold information it is making a tremendous mistake. Any attempt to withhold information generally back-fires for an organization by

causing the public and media to lose confidence in them through the creation of a feeling of mistrust.

Establishing and maintaining a media relationship built on trust is an important factor in the process of maintaining an organization's image during not only good times, but in times of crisis and disaster. Therefore it is necessary for organizations to develop crisis communication plans and establish mutually beneficial partnerships with the media so in times of crisis the media will assist in informing and protecting the public.

Media – Mass Communication

Mass communication "occurs when a small number of people send messages to a large anonymous and usually heterogeneous audience through the use of specialized communication media" (Mass Communication Context, 2001). Therefore anytime a person turns on the radio or television, reads a newspaper, or logs onto a news website they have become engaged in the mass communication process and system.

Mass communication is the process of transmitting information to the public and other interested parties. The media (in the form of journalists) is the messenger and it determines what information is shared and communicated to the public during the communication process. Mass media is a means of public communication which reaches a large number of people in an effort to inform and protect the public. The challenge for many organizations today is to determine which of the hundreds of media organizations are dedicated to the traditional values of public information (public's right to know) and public welfare as opposed to the sensationalist media who are after the "story" of the day and high ratings.

Reputable media sources for this paper are considered to be news entities which have established themselves in the area of informational and factual reporting of events, plus they have a proven track record in unbiased reporting of events and facts. The reputable

media sources used in this paper are, for electronic media: ABC News; CBS News; NBC News (and their local affiliates); CNN; and MSNBC; for print media: major local, regional, and national newspapers. Sensationalist (tabloid) and entertainment media sources were not used in this paper, due to the fact that the majority of these media outlets present consistently biased opinions in an attempt to sway public opinion and generate increased market share and higher ratings. While elements of factual presentation are seen in tabloid media, the predominance of sensationalism makes them an unreliable source for consistently factual information and accurate presentation of events. Fox News was not used due to the fact this source is in a transitional phase from entertainment news to a reputable major news entity and resource. For this paper, the term "media" is used as it relates to the definition of reputable media above.

When working with the mass media, one must be aware of the three primary units of mass communication which are: message; medium; and audiences. In any mass communication plan, these elements and their interactivity must be considered and planned for.

Selected Mass Communications Theories

It is important to understand the mass communication theories that will affect the development of an organization's crisis communication plan. There are twelve mass communication theories, however only three of these are relevant to crisis communication planning and the establishment of media partnerships. The three theories most relevant for crisis communication plans are agenda setting, functional approach to mass communication, and the media equation.

The Agenda-Setting Theory states that the media (primarily the news media) are not always successful at telling the public what to think, but they are quite successful at telling the public what to think about. Due to the media's desire to present information to the public as soon as possible, they will go off on tangents that distract from the event and actions of the organizations involved. An organization must have processes (plans) to counter this development and strategies on how

to address the concerns raised in these side-stories in an effort to return the media's focus to the actual occurrences and activities taking place. An organization cannot control the side-stories that develop, but contingency plans and strategies must be developed to minimize the public's interest and demand for information in these tangential areas.

The Functional Approach to Mass Communication Theory states that there are five functional approaches the media serves related to users of the media: surveillance; correlation; transmission; entertainment; and mobilization. Surveillance is the means by which the media presents the information given. Correlation is where the media presents the information to the public after they select, interpret, and criticize it. Transmission refers to how the media reflects the public's beliefs, values, and norms. Entertainment provides an escape from everyday life, and mobilization refers to promoting a society's interest. For crisis communication planning the only function that does not need to be addressed is the entertainment function. It is important for crisis communication planners to adapt elements of surveillance, correlation, transmission, and mobilization into any and all communication plans. When these items are addressed in the appropriate manner, an organization will be able to maintain a positive image by promoting understanding and cooperation concerning the coverage of the crisis.

The Media Equation Theory attempts to predict why people respond unconsciously and automatically to a communication medium as if it were human. The public will react to the media presentation as if they were talking to a person one-on-one. This tends to make the public an active participant in the event that is unfolding instead of an inactive observer. They want information, and the media is their venue for receiving it. This element must be addressed in order to assist the media in organizing the knowledge base required for the public to make informed decisions and to influence the action(s) the public might take, for example patience versus aggression, or evacuation versus remaining in the area.

Crisis Communication versus Traditional Public Relations

Traditional organizational public relations focuses primarily on the delivery of the message, persuasion (advertising), preservation of image, and facilitation of open dialogue. Most of these elements are used in crisis communication with the exception of persuasion. Any attempt to use persuasion during a crisis is viewed negatively by the media and public. Additionally, attempts at persuasion will hurt an organization's image and trustworthiness in reference to the media's and public's perception of the organization's willingness to help.

As with traditional public relations, a crisis communication professional needs to know what the message is, how to deliver it quickly and concisely, and who will transmit it to the media and public. The challenge for an organization in a crisis situation is to provide a balance of information. Too much information will cause speculation and side-stories to run wild. Too little information and the phrase "cover-up" will be used to describe an organization's response to the event. In a disaster (crisis) situation an organization must be prepared to be truthful and provide accurate disclosure on the events that occur. Hiding facts and other elements of the investigation and recovery efforts being taken only succeeds in hurting the organization as these efforts progress. A key component for any effective crisis communication plan is the professional an organization has "up-front" and / or coordinating the communication process, and information dissemination to the media, public and stakeholders related to the occurrence. This individual must be confident, calm, and have an ability to communicate effectively and empathically with various types of individuals and groups.

For crisis communication planning, the preservation of an organization's image is not about promoting the latest product (service) or promotional image, but is related to handling the dissemination of information in a timely and effective manner. Truth, honesty, and competence, are the elements that preserve an organization's image in a time of crisis. The ability for the media and

the public to receive actuate information from a reliable source (the organization involved) provides them with a positive feeling that the organization is in control and will keep the channels of communication open. Also, the personnel presenting the information need to remain professional and empathic to the media, the situation, the victims, and the victims' families in order to establish trust and cooperation from the various parties involved in the crisis.

The facilitation of open dialogue is essential in keeping communication channels open. In traditional, public relations this open dialogue format provides an opportunity for building consensus or establishing an opinion related to a specific topic of interest to the organization. When open dialogue is used in crisis communication it is not to gain consensus, but to provide information related to public safety and the various issues the media, victims, and victim's families have brought up as important. While to some the issues raised by the media, victims, and victims' families might seem mundane and unimportant, they must be addressed effectively and empathically due to the emotional nature of crisis communication. Avoiding these items or "brushing" them off, creates the impression of an uncaring organization that is non-responsive to their customers, employees, the media, and the public.

Crisis Communication Training

An organization needs to develop training to address the specific needs and skills related to crisis communication planning. Through an informal survey with air carriers that have family assistance plans, communication professionals and instructors, and crisis management professionals, the following five items were cited as key components for crisis communication personnel training.

1. Planning skills
2. Public speaking skills (experience)
3. Interpersonal communication skills
4. Media relations skills
5. Organizational skills (multi-tasking)

An individual must possess effective planning skills in order to present a competent and caring image, plus provide actuate information for dissemination by the media. Failure to have defined elements ready for use in the event of a crisis was listed as a major failure in crisis communication planning. Persons who will be responsible for or involved in the actual communication process during a crisis, need to establish plans outlining who will address the media on what topic areas related to the event, frequency and location for addressing the media, victims, and victims' families, plus determine the necessary facilities and equipment to have available to the media for filing reports. Making the media or the public wait and search out information hurts the communication process by portraying a disorganized and uncaring image on the part of the organization. This negative perception is translated subconsciously by persons outside the organization to be the reason why the event occurred – disorganization and apathy on the part of the organization involved.

Public relations professionals are trained in public speaking however they are most familiar with the traditional public speaking style which is delivery of a specific message which may or may not include a question / answer section after the message is delivered. In times of crisis, the normal situation of the audience waiting patiently for the delivery of the message is gone. Generally, the professional entering the room to address the audience is bombarded with questions and photographers, there is little or no time to take a breath and present a message interruption free. The crisis communication professional is trained to entered the "press room" in the midst of chaos and gain control of the room with a simple, confident introduction and brief update (opening) statement. In addition, they are taught to present the information in a clear, concise manner and control the questions asked by setting forth a "game plan" to the members of the media and others present for these information sessions.

Traditional interpersonal skills are stressed in crisis communication training due to the highly emotional and personal aspects involved in

dealing with the outcomes of disaster (loss of life, destruction of property). The crisis communication professional will have interaction with victims and their family members; therefore they must be trained in skills such as active listening and empathic response. People cannot be taught to "fake" emotional responses therefore it is essential for persons involved to be genuine. If the media or public detects false (non-genuine) emotion, the organization will be viewed poorly concerning their treatment of the accident and those involved.

Crisis communication training also involves media relations through role plays. One of the trickiest aspects for any public relations plan is the press conference. A person is in front of a room of media professionals each of whom has their own agenda, and each wants to discover the cause or solution for the situation before anyone else. Examples of political press conferences are reviewed due to the effective way in which these press conferences present a public relations professional's response to the media under various conditions. The person's treatment of the media and their responses to questions and statements made by the media demonstrate key components for success and failure. In addition, the training teaches crisis communicators to respond to questions with simple and direct answers, and stresses that a person should not expand their responses with more information. The courtroom "ask-and-answered" model is used for this portion of the training. Providing factual information and / or an update on the activities related to the crisis are the main elements of a press conference (and family / press briefing), allowing or creating movement to side-stories and tangents creates uncertainty and speculation.

Organizational skills and the ability to multi-task are essential for a professional in crisis communication. Most crises will occur without warning therefore the individual(s) must be able to mobilize a plan and the other people involved within a matter of hours (less for the first press statement). Students learn how to move from their normal, daily routine to full implementation of a crisis communication plan, including preparation of press / family briefing area, and delivery of

press packet materials. Timelines are addressed in order to demonstrate efficient and supportive response by the organization. Items such as press packets, technical and organizational information sheets, and personnel information are addressed as pre-crisis preparation activities, so dissemination of the information is timely when a crisis occurs.

Media Partnership Program

Organizational - media relationships are similar to other types of relationships. One cannot expect a helpful and trusting relationship to exist if there is no effort to establish and maintain a relationship prior to a time of crisis. The majority of organizations have a "love-hate" relationship with the media. The media has the ability to be a great friend and ally, but in the blink of an eye they can become a formidable enemy.

The key to building a relationship with the media is to work with them on a regular basis during times of normal operations. Establishing an organization as cooperative and an industry resource has led many journalists to assist and support an organization in a time of crisis. An organization cannot rely on this goodwill relationship to prevent side-stories or "tough" questions, but without it journalists tend to be more adversarial than helpful.

According to media professionals (journalists) the best way to establish a positive relationship is to work with the journalist(s) on various topics prior to a crisis. This includes regular meetings with the journalist(s) (monthly or bi-monthly) to establish a level of trust and understanding. Also, a willingness to assist the journalist(s) in areas of interest to their audience related to the organization and / or their industry establishes a cooperative link. For example, an air carrier may provide a journalist with information concerning the purchase of new aircraft including technical, environmental, and enhanced public service information.

Journalists surveyed also stated that when organizations provide a "disaster handbook" or guide they are more likely to disseminate the information presented to them by the organization rather than use the information gathered from outside sources. Information from the organization is preferred to outside "expert" information, but a journalist will use whatever information they can find to provide information to the public related to the event. This handbook should have detailed information concerning the route of flight (with departure and arrival times), technical information on the aircraft involved, a listing of the number of persons involved (no names), the organization's safety and training program, and other information such as briefing schedules.

Conclusion

The 24-hour news cycle creates many challenges for transportation professionals, but the situation is not hopeless if an organization takes the time to plan and prepare for crisis situations by developing effective media relations (partnerships).

How an organization handles a crisis and disseminates the information to the media and public is key to maintaining a positive corporate image. Failure to respond with confidence and empathy is an invitation for failure. Crisis communication professionals need to be developed within transportation organizations to provide the unity an organization will need during times of crisis and disaster.

Establishing a positive, working relationship with the media is one of the most important elements in any crisis communication plan. The media at first look may seem like the enemy, but research and existing media relations programs demonstrate the power the media has to assist organizations in times of crisis and trouble. Reputable media sources want to provide the public with truthful and factual information, they prefer to work with organizations directly in order to reduce the amount of bad and speculative information released to the public.

An organization's development of a crisis communication plan and strong media partnerships prior to needing either can be equated to buying an insurance policy. Like an insurance policy a crisis communication plan and media partnerships are put in place with the hope that they will not be needed, however should the need arise, having them in place is reassuring and comforting during a difficult time.

REFERENCES

Binns, Tristan B. (2003). FEMA: federal emergency management administration. Chicago: Heinemann Library.

Englehardt, Elaine, E., Barney, Ralph D. (2002). Media and ethics: principles for moral decisions. Stamford, CT: Thomson Learning.

Erikson, Kai T. (1994). A new species of trouble: explorations in disaster, trauma, and community. New York: Norton.

Fensch, Thomas. (1990). Associated press coverage of a major disaster: the crash of delta flight 1141. Hillsdale, NJ: Lawrence Erlbaum Publishers.

Kingston, Jeremy., Lambert, David. (1979). Catastrophe and crisis. New York: Facts on File.

Lane, D.R. (January 2001). Mass Communication Context. Retrieved on December 1, 2004, from www.uky.edu/~drlane/capstone/mass.

LaValla, Patrick. (1999). Search is an emergency: field coordinator's handbook for managing search operations. Olympia, WA: ERI International.

Steinberg, Theodore. (2000). Acts of God: the unnatural history of natural disaster in America. New York: Oxford University Press.

Van Ruler, B. (2004). The communication grid: an introduction of a model of four communication strategies. *Public Relations Review*, 30 (2), pp. 123-143. Amsterdam.

Williams, Robert Leonard. (1993). Starting Over. Waco, TX: WRS Publishing.

Wright, Ted. (1993). Wright's complete disaster survival manual. Norfolk, VA: Hampton Roads.

Potential Demand for Household Alternative Fuelled Vehicles in Hamilton, Canada: A Stated Choices Experiment and Survey

Dimitris Potoglou, Pavlos S. Kanaroglou
Centre for Spatial Analysis (CSpA), School of Geography and Geology, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada (e-mail: potogld@mcmaster.ca)

1. Introduction

The development of alternative fuelled vehicle (AFVs) technologies is considered a key implementing strategy towards environmental sustainability and improved air quality in cities. Analysis of the role of vehicle technology in fulfilling sustainability targets requires estimates of future vehicle demand. A fuller understanding of the demand is constrained by the inability to observe actual (revealed) car type preferences, which has led researchers to the use of stated choice methods.

The aim of this paper is twofold. First, we review the existing literature on household automobile demand modelling for both conventional and AFVs. Second, we describe an internet-based data collection effort for automobile demand that focuses at the urban level. The survey consists of two major stages; the first asks about transaction of vehicle make and models, within a time horizon of eight years (1997-2005) and the second employs a discrete choice experiment regarding households' future vehicle purchases including both conventional and alternative fuelled vehicles. This research makes use of contemporary data collection techniques and has the potential to offer empirical evidence on preferences regarding alternative fuelled vehicles in the Canadian urban context.

In the context of assessing potential demand for AFVs, it is necessary to address prominent vehicle technologies and their major characteristics. The following section provides a brief discussion on "cleaner" technologies focusing on their advantages and disadvantages as well as their overall development trends.

2. Alternative Fuelled Vehicle Technologies

AFV technologies are fuel-efficient, low- or zero-emission solutions to the problems of the internal combustion engine (ICE). This section following briefly describes the prominent near- or long-term technological advances including battery-powered, fuel cell and hybrid electric vehicles. The discussion draws heavily from the works of Dearing (2000), Wang (1999) and Walwijk et al. (1999).

Battery Electric Vehicles

The battery electric vehicle (BEV) is the ideal case of zero noise and exhaust emissions. Advanced electronic systems both for generators and electric motors can increase the mean energy efficiency of an electric drive train to approximately 70 percent. Also, a considerable amount of the input energy can be recovered by a regenerative braking system and other improvements such as high efficiency motors, inductive recharge systems and aerodynamic designs. However, most BEVs today suffer from limited performance, compared to gasoline vehicles, and the high battery costs. Also, major limitations of BEVs are the travel range between refuelling and the recharge time of the battery. The range goes from 95 to 250 km, while the recharging time could last from several minutes to hours.

Fuel Cell Vehicles

Fuel cell vehicles (FCVs) are characterized by high-energy efficiency, zero tailpipe emissions (depending on the type of fuel cell) and low levels of noise (Walwijk et al., 1999). Fuel cell vehicles achieve efficiency levels between 45 and 55 percent, which increases under moderate loads.

The storage, production and distribution of hydrogen fuel are of the most important constraints to wide adoption of fuel cells. The fuel cell system requires pure hydrogen as fuel in a relatively low-temperature. Hydrogen can be stored directly (direct systems) or produced onboard the vehicle (indirect systems) by reforming a liquid fuel such as gasoline, methanol or synthetic liquids produced from natural gas. A fuel processor can effectively process these fuels into hydrogen prior to conversion in the fuel cell.

Hybrid Electric Vehicles (HEV) with ICE

The operation of hybrid electric vehicles (HEVs) is based on the combination of an electric motor and a battery with a heat engine and a fuel tank. The primary engine and the electric motor are used in series or in parallel.

In the series design, the resulting energy of the heat engine is converted to electrical energy through the generator, which either separately or jointly with the battery, powers a single drive train. In the parallel design, the engine and the battery are capable to drive the vehicle together or separately, achieving the same power output from a smaller engine, with a slight reduction to overall efficiency. The independent battery system provides auxiliary power, accepts regenerative breaking energy, and restarts the engine in city traffic (McNicol et al., 2001).

Although a hybrid vehicle might have great overall complexity, individual components may be smaller and simpler than in a fully internal combustion or battery-powered vehicle. In particular, hybrid vehicles require less on-board electrical storage capacity than battery vehicles, which a flywheel or ultra-capacitor could provide. They also provide an essentially unlimited range from the existing gasoline fuel infrastructure (Dearing, 2000).

AFVs' Development Trends

Future vehicle technologies are generally classified into short- and long-term technologies. Short-term technologies include mainly hybrid electric and compressed natural gas vehicles. Being already available in the market, short-term technologies guarantee reduced GHG emissions, while their fuel efficiency will continue improving. Such technologies are the intermediate step that would help to bridge the gap between existing and long-term technologies (Wang, 1999). The latter includes EVs, HEVs and FCV. HEV and FCV can reduce the GHG emissions over 40 percent and under special conditions can reach 80 percent (Wang, 1999). Figure 1 shows the relative position of many powertrain technologies. Fuel cell electric vehicles are the most efficient and produce the lowest emissions. The gasoline internal combustion engine is a relatively low emission, low

efficiency power source. However, a natural gas vehicle is a cleaner solution with no aromatic compounds and lower fuel costs (Walwijk et al., 1999). Hybrid electric vehicles are more efficient and emit lower emissions. Electric vehicles, on the other hand, require 24-35 percent lower total energy, while they achieve a reduction by 55-85 percent of total oil use and emit 24-37 percent lower GHG. However, battery production may increase pollutants such as SOx, TSP and heavy metals.

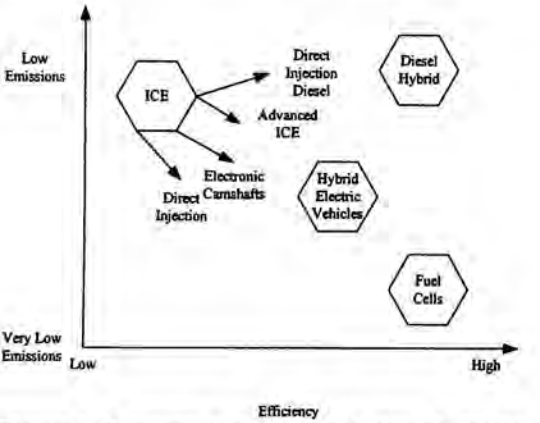


Figure 1: Powertrain technologies (Smith, 2001)

From the FCV's perspective, the use of gasoline, methanol and natural gas are the fuels receiving most of the attention. The development of FCVs is largely driven by the various stakeholders particularly car manufacturers and oil companies. While they are both engaged in FCV research, car manufacturers tend to emphasize on the benefits of the "ethanol route" while oil companies on the "hydrocarbon" route, thus efforts have remained divergent (Dearing, 2000).

3. Modeling of Automobile Demand: A literature review

Early automobile demand models focused on aggregate estimates of car-ownership per capita or per household using area-wide time-

series or zonal-based regression models (Hensher et al., 1992 ; Purvis, 1994). The aggregate modelling approach is considered relatively cheap in terms of computation and data requirements, but suffers from limited representation of the process to a non-dynamic context, that is the household and vehicle characteristics remain unchanged during the study period, and from the significant level of collinearity between explanatory variables (de Jong et al., 2004). Such models have been used for the assessment of "what if" scenarios, but they are impractical for projecting demand over several years (Train, 1986).

In contrast, disaggregate approaches are structurally stable and capture the causal relationship between car ownership level and its potential determinants. Early approaches focused on car ownership level models (one, two cars, etc.), which assume that a household associates a utility value for each car ownership level and selects the level with the maximum utility (Lerman and Ben-Akiva, 1976). However, recognizing the interdependence between vehicle-type choice and the number of vehicles owned, a number of vehicle-type choice models have been developed (Bhat and Pulugurta, 1998). Main explanatory variables include cost related and vehicle attributes, socio-economic characteristics of car owners and manufacturer country of origin (McCarthy and Tay, 1998). Both car ownership level and vehicle type choice models are also static models and they are limited to the new-vehicle market. In this context, such models do not provide policy implications regarding any dimension of the automobile ownership (e.g., the effect of fuel price on the market shares of different vehicle sizes) (Tardiff, 1980).

In support of the previous limitations, disaggregate vehicle-holdings models assume that households retain or revise their current number of automobiles and types at finite equal intervals. Vehicle-holdings models use data from cross sectional surveys to incorporate modelling aspects such as the number, type of vehicles in the household fleet and the usage rate of each vehicle in a discrete-time fashion (Mohammadian and Miller, 2003). Modelling of these decisions is theoretically consistent with the process of joint utility maximization. As Figure 2 shows, it is based on a three-level decision sequence in which the first level refers to the choice of "number of vehicles to own", the second is the choice of the "class" of the vehicle

and the third level corresponds to the "make/model" and sometimes vintage of the vehicle (Bunch, 2000).

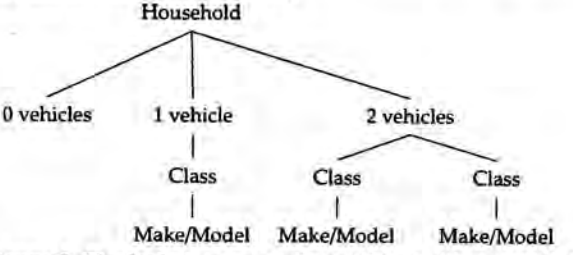


Figure 2. Typical structure of a vehicle-holdings model

The main concern about automobile-holdings models is their static nature. A dynamic framework would consider the "decision to transact" prior to "vehicle type-choice". The dynamic approach is also attractive from a behavioural point of view, because it might be more consistent with the actual process of vehicle-choice within a real-world dynamic setting (Berkovec and Rust, 1985 ; Bunch, 2000). In satisfying such requirements, automobile transaction models consider the vehicle type-choice conditional on the decision to be active in the market within a fixed time-period (de Jong, 1996 ; Yamamoto et al., 1999), while they fundamentally recognize that the processes of buying and selling vehicles are heterogeneous. Within a pre-specified time-interval, a household faces four decisions: (a) to purchase, (b) to replace, (c) to dispose a vehicle or (d) do-nothing. Each of the outcomes is determined as a function of socio-economic and demographic characteristics, vehicle attributes and/or an "event" such as vehicle failure. From a utility maximization perspective, one of the transactions is triggered when the household's net utility gain of transacting exceeds a threshold (Bunch, 2000). Additional choice levels can incorporate vehicle-type choice (make/model/vintage, see figure 3). Dynamic automobile transaction models are more data intensive requiring longitudinal data on vehicle transactions and vehicle choices, which are modelled either using discrete choice or hazard duration models.

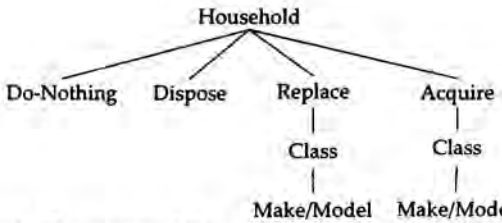


Figure 3. The automobile-transactions modelling approach

4. Demand for Alternative Fuelled Vehicles - Stated Choices Experiments

Conventional methods with actual market data (revealed preferences, RP) have proven to be insufficient for assessing future demand for non-gasoline vehicles (Train, 1986 ; Beggs and Cardell, 1980). Alternative fuelled vehicles exhibit performance characteristics and attributes, which are either unavailable or limited in the current market. For instance, fuel availability, travel range and lower vehicle emissions deviate from present day vehicles. Also, the variation of existing vehicle attributes is inadequate for providing robust modelling outcomes for alternative fuelled vehicles.

Stated Choices (SC) or Discrete Choice Experiments (DCE), borrowed from the field of marketing research, consist an attractive approach for assessing demand of a new product, where new attributes of these products are absent and therefore actual choices have not been observed yet (Hensher et al., 1999). Such data are generated by a systematic design procedure (experimental design), in which the attributes (variables) and their levels (values) are pre-defined without measurement error and varied to create choice alternatives (Louviere et al., 2000). Each alternative option is represented as a "package" of different attributes and respondents simply choose the most preferred option. The main feature of the SC approach is the ability to obtain repeated measurements (treatments) that is each respondent can contribute with a number of observations, typically between eight and sixteen.

In the context of alternative fuelled vehicles, attributes describing hypothetical vehicles are divided into three categories: monetary, non-monetary and environmental (Bunch et al., 1993). Monetary

attributes include the purchase price of the vehicle, fuel and maintenance costs whereas non-monetary are the availability of the fuel, the range between refuelling or recharging, vehicle performance, refuelling time and convenience, and interior space. Finally, environmental attributes usually correspond to emission reductions because of their contribution to improved air quality.

Studies in early- and mid-eighties focused on the demand for battery-electric vehicles and they showed that limited range and long refuelling times kept the demand for these vehicles low (Beggs and Cardell, 1980 ; Calfee, 1985). A later study (Bunch et al., 1993) in the state of California included clean-fuel vehicles such as methanol, ethanol, propane and natural gas. Characteristics such as purchase price, fuel operating cost and vehicle range between refuelling were found to be the most significant attributes. Introducing the similar hypothetical vehicles (conventional, electric and fuel efficient), Ewing and Sarigollu (1998) examined the vehicle-technology choice of suburban drivers in Montreal, Canada. The main purpose of that study was to identify the most important factors influencing the choice of lower-and zero-emission vehicles. Alternative policy measures encouraging the purchase of cleaner vehicles were based on economic instruments such as commuting cost and time. The results revealed a large potential for fuel-efficient and electric vehicles among Montreal car commuters.

Contemporary research (Brownstone et al., 2000) has focused on integrating stated choices with revealed preferences to estimate market penetration of AFVs and assess alternative policy scenarios. This requires joint estimation of SC and RP data, which allows forecasts for both new and used vehicles. Also, an important aspect of assessing demand for alternative fuelled vehicles is evaluation of potential environmental benefits through emission reductions (Kazimi, 1997). The latter has been examined only in the case of the Los Angeles Area using a microsimulation model of automobile ownership (Ren et al., 1995) to estimate total-vehicle emissions.

5. The CIBER-CARs Survey for Hamilton CMA

The Choice Internet - Based Experiment for Research on CARs (CIBER-CARs)[1] is a self-administered online survey for household automobiles in the Census Metropolitan Area (CMA) of Hamilton. The objective of this data collection effort is to build an understanding of households' behaviour regarding purchases, replacements and disposals of their automobiles. Also, this study seeks to assess the importance of several vehicle attributes and household characteristics that may shift choices to "cleaner" vehicles in the near future. The CIBER-CARs survey consists of two stages.

Stage 1 – Longitudinal Automobile Survey

The first stage is a longitudinal (revealed preferences) questionnaire that collects information on household automobile transactions, makes, models and age of the vehicles as well as the history and the structure of the household retrospectively (see, Table 1).

Table 1. Core questions in Stage 1 of the CIBER-CARs survey

Item	Description
Household	location, type of building, number of members and vehicles, total household income, date household formed, date moved in the Hamilton area.
Person(s)	gender, age, work and student status, licence, education level
Vehicle(s)	make, model, year of production, main user of the vehicle, ownership, fuel type
Future Vehicle Purchase Plans	vehicle class, money spent if purchased, usage rate (kilometres/year)

To prevent measurement error and typos, the survey employs a number of skip patterns based on the previous answers of the respondent and drop-down menus. For instance, the answer on the make/model/age of the household vehicle is simply selected from an

[1] Available at: http://www.science.mcmaster.ca/~potogld/php/index.php

embedded database that integrates all the make/models recorded on the Fuel Economy Guide (all models between 1972-2005). The last step of this stage asks respondents if they would purchase a vehicle in the next five years. In case of a positive answer, additional questions ask about the class, the annual usage rate and the amount of money a household would spend to buy that vehicle.

Stage 2 - Stated Choices Experiment

The second stage is a stated choices experiment regarding future vehicle purchases. Vehicle choice alternatives are labelled based on the fuel type and class of the vehicle. The fuel type label includes a conventional gasoline, a hybrid-electric and an alternative-fuel vehicle labelled also with vehicle class categories such as subcompact car, compact car, mid-size car, large car, pick-up truck, van or SUV.

Each respondent completes eight exercises, which involve the choice of one out of the three fuel type options (Figure 4). The conventional gasoline vehicle is the base alternative and it is customized per respondent based on information elicited in Stage 1 (see Table 1, Future Plans).

	Vehicle A	Vehicle B	Vehicle C
Fuel Type	Gasoline	Hybrid	Alternative Fuel
Vehicle Class	Compact	SubCompact	Compact
Price	$10,000	$11,000	$9,000
Fuel Cost per Year	$1,500	$1,300	$1,100
Maintenance Cost per Year	$3,000	$5,000	$4,000
Acceleration	6 sec	9 sec	12 sec
Fuel Availability	4 out of 4 stations	4 out of 4 stations	2 out of 4 stations
Incentive	NOT APPLICABLE	Carpool Lanes	None
Tailpipe Emissions	85% OF PRESENT DAY AVERAGE CAR	50% OF PRESENT DAY AVERAGE CAR	10% OF PRESENT DAY AVERAGE CAR
CHOICE			

Figure 4. Example of a choice exercise (stage 2)

The other two options (i.e., the hybrid-electric and the alternative-fuel vehicle) are constructed based on the scenarios of an experimental design (Louviere and Street, 2000). The attributes describing each of the vehicle choice alternatives in this experiment were collected from the existing literature (Bunch et al., 1993 ; Ewing and Sarigollu, 1998 ; Adler et al., 2003), pilot testing and request of

experts' opinions. In general, it is necessary - for stated choice experiments - to minimize the number of attributes, because of the overall cognitive complexity and respondent's fatigue. Therefore we tend to include the most important and relevant attributes. Each attribute in this experiment uses four attribute levels to cover the range of interest (Table 2).

Table 2. Attributes and levels in the DCE

Attribute	Levels
Purchase Price	(1) -20%, (2) -10%, (3) Base, (4) +10% than the Base
Annual Fuel Cost	(1) - 80%, (2) - 60% (3) - 40% lower, (4) -20% than Base
Annual Maintenance Cost	(1) - 50%, (2) -25 %, (3) Base, (4) +25% than Base
Fuel Availability (For AFVs only)	(1) 75%, (2) 50% (3) 25% (4) 10% of existing stations
Acceleration	(1) 6, (2) 9 sec, (3) 12, (4) 15 sec
Incentives (For Hybrid Electric and AFVs only)	(1) None, (2) Carpool Lanes, (3) No Parking or Metered Fees, (4) No Purchase Taxes
Pollution Levels (For Hybrid Electric and AFVs only)	(1) 10%, (2) 25%, (3) 50%, (4) 75% of present day average car

The values (levels) of attributes assigned to the hybrid-electric and the alternative-fuel vehicle options of a choice exercise are drawn from a design matrix. The design of the matrix is based on the general approach for SC experiments for developing discrete choice models (Louviere et al., 2000 ; Kuhfeld, 2005) and includes 128 scenarios. The first 64 scenarios were derived from a 4^{13} orthogonal main effects design and the remaining 64 scenarios were generated from a 2^{13} "endpoint" main and interactions (within-the-alternatives) effects. Of the eight scenarios assigned per respondent, the first four are taken randomly without replacement from the first 64 scenarios (in blocks of 16 respondents) and four are taken randomly from the second 64.

Such design allows for parameter estimation of both non-linear main effects and also two way interactions of attributes within the alternatives.

Data Collection Procedures

Potential respondents are invited to participate in the survey through an e-mail invitation. The e-mail serves as an introduction letter, which - in addition to the internet address of the survey - includes some general information and also asks recipients to forward the invitation to other residents of the area. E-mail lists have been obtained through McMaster University and the City of Hamilton. Eligible to participate are household heads or member who participate in household decisions regarding vehicle purchases. So far, two pilot data collection campaigns have taken place for both review and general testing of the survey instrument and data storage purposes.

6. Conclusions

The assessment of preferences for AFVs by the general public is very important for the design of policy that would address environmental concerns in relation to the transportation system.

This paper, after a brief description of AVF technologies, provides an overview of existing automobile ownership modelling and describes a data collection effort currently undertaken in the Hamilton CMA. Contemporary approaches in automobile demand focus on agent-based (household, individual) modelling, combining both revealed and stated vehicle transactions and type-choice of automobiles (Brownstone et al., 2000). The data collection of the urban-scale automobile survey for the Hamilton CMA, described in this paper, is based on concepts from this type of research.

Modelling of such data using discrete choice models will contribute to the development of an automobile modelling system for determining the market share of alternative and conventional vehicle technologies per vehicle class under several scenarios. Implication can be derived to assess the environmental benefits of "clean" vehicles on the road network of the Hamilton CMA.

References

Adler, T., Wargelin, L., Kostyniuk, L.P., Kavalec, C. and Occhiuzzo, G. (2003) Incentives for Alternative Fuel Vehicles: A large-scale stated preference experiment In *10th International Conference on Travel Behaviour Research*Lucerne, Switzerland.

Beggs, S.D. and Cardell, N.S. (1980) Choice of smallest car by mutli-vehicle households and the demand for electric vehicles *Transportation Research Part A,* **14A** (389-404.

Berkovec, J. and Rust, J. (1985) A nested logit model of automobile holdings for one vehicle households *Transportation Research B,* **19B (4)**, 275-285.

Bhat, C.R. and Pulugurta, V. (1998) A comparison of two alternative behavioral choice mechanisms for household auto ownership decisions *Transportation Research Part B,* **32 (1)**, 61-75.

Brownstone, D., Bunch, D.S. and Train, K. (2000) Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles *Transportation Research Part B,* **34** (315-338.

Bunch, D.S. (2000) Automobile Demand and Type Choice. In: *Handbook of Transport Modeling* (Eds, Hensher, D.A. and Button, K.J.), Pergamon, New York, pp. 463-479.

Bunch, D.S., Bradley, M., Golob, T.F., Kitamura, R. and Occhizzo, G.P. (1993) Demand for clean-fuel vehicles in California: A discrete choice stated preference survey *Transportation Research A,* **27 (3)**, 237-253.

Calfee, J.E. (1985) Estimating the demand for electric automobiles using fully disaggregated probabilistic choice analysis *Transportation Research,* **19B** (287-302.

de Jong, G. (1996) A disaggregate model system of vehicle holding, duration, type choice and use *Transportation Research Part B,* **30 (4)**, 263-276.

de Jong, G., Fox, J., Daly, A., Pieters, M. and Smit, R. (2004) Comparison of car ownership models *Transport Reviews,* **24 (4)**, 379-408.

Dearing, A. (2000) Technologies supportive of sustainable transportation *Annual Review of Energy and the Environment,* **25** (89-113.

Ewing, G.O. and Sarigollu, E. (1998) Car fuel-type choice under travel demand management and economic incentives *Transportation Research D,* **3 (6)**, 429-444.

Hensher, D.A., Louviere, J. and Swait, J. (1999) Combining sources of preference data *Journal of Econometrics,* **89** (197-221.

Hensher, D.A., Smith, N.C., Milthorpe, F.W. and Barnard, P.O. (1992) *Dimensions of Automobile Demand: A Longitudinal Study of Household Automobile Ownership and Use,* Elsevier, Amsterdam.

Kazimi, C. (1997) Evaluating the environmental impact of alternative-fuel vehicles *Journal of Environmental Economics and Management,* **33** (163-185.

Kuhfeld, W. (2005) *Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint and Graphical Techniques,* SAS Institute Inc., Cary, NC, USA.

Lerman, S.T. and Ben-Akiva, M. (1976) Disaggregate behavioural model of automobile ownership *Transport Research Record,* **569** (34-51.

Louviere, J.J., Hensher, D.A. and Swait, J.D. (2000) *Stated Choice Methods: Analysis and Application,* Cambridge Press.

Louviere, J.J. and Street, D. (2000) Stated-Preference Methods. In: *Handbook of Transport Modeling* (Eds, Hensher, D.A. and Button, K.J.), Pergamon, New York, pp. 131-143.

McCarthy, P.S. and Tay, R.S. (1998) New vehicle consumption and fuel efficiency: A Nested logit approach *Transportation Research Part E,* **34 (1)**, 39-51.

Mohammadian, A. and Miller, E.J. (2003) Dynamic modeling of household automobile transactions *Transportation Research Board - 2003 CD-Rom.*

Purvis, C.L. (1994) Using 1990 census use microdata sample to estimate demographic and automobile ownership models *Transportation Research Record,* **1443** (21-29.

Ren, W., Brownstone, D., Bunch, D.S. and Golob, T.F. (1995) A vehicle transactions choice model for use in forecasting demand for alternative-fuel vehicles *Unpublished Manuscript.*

Tardiff, T.J. (1980) Vehicle choice models: Review of previous studies and directions for further research *Transportation Research A,* **14A** (327-335.

Train, K. (1986) *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand,* MIT Press.

Walwijk, M.V., Buckmann, M., Troelstra, W.P. and Elamn, N. (1999) *Automotives Fuels Survey Part 2,* IEA Autom. Fuels Inf. Service, The Netherlands: INAS.

Wang, M.Q. (1999) Fuel-cycle Greenhouse gas emission impacts of alternative transportation fuels and advanced vehicle technologies *Transportation Research Record: Energy and Environment,* **1664** (9-17.

Yamamoto, T., Kitamura, R. and Kimura, S. (1999) Competing-risks-duration model of household vehicle transactions *Transportation Research Record,* **1676** (116-123.

Hydrogen Hybrid Internal Combustion Engine Bus Demonstration

Allister Hickson, Ph.D., CFA
Christy Sokol, B. Comm. (Hons)

University of Manitoba Transport Institute

Introduction

Hydrogen fuel cell powered buses have often been described as the gold standard for public transportation for the future, offering quiet, zero emission transportation. In spite of their promise, hydrogen fuel cell vehicles have remained stubbornly as the technology for the future.

Most automakers view fuel cells as a viable technology only in the 2020+ time frame[1]. As of 2005, there are very few fuel cell vehicles in operation. According to Adamson, Baker and Jollie[2] there have been only about 800 fuel cell vehicles produced cumulatively, with around 65 of these being fuel cell buses[3].

Some of the factors that have been cited as reasons for the slow rate of adoption of fuel vehicles include[4]:

1. The technology remains developmental.
2. Work needs to be continued on the integration of on board systems, especially in hybridized vehicle applications.
3. Fuel cells remain very costly.
4. Supportive infrastructure such as fuelling stations have yet to be developed in any meaningful manner.

While the technical issues will be overcome through further research and testing, the last two issues fall into the classical chicken and egg problem. Before costs of fuel cell vehicles can be driven down through mass production, consumers have to be willing to purchase the technology. However, consumers will not purchase the technology without accessibility to fuel supplies similar to that available for today's vehicles. Fuel suppliers will not invest in hydrogen fuelling infrastructure until there are sufficient vehicles being produced and resultant derived demand to make that infrastructure economically viable.

One approach to break this circle is to adopt a bridge technology that provides most of the emission benefits of hydrogen fuel cells, yet is more economically competitive with current engine systems (gasoline and diesel). Such an alternative is the Hydrogen Internal Combustion Engine (HICE). The hydrogen internal combustion engine is a matching of off the shelf existing engine technologies using hydrogen as a fuel. In hybridized form it is a matching of the hybridization technology deployed in fuel cell vehicles, primarily regenerative energy capture and storage, with a more or less standard approach to electric generation. This coupling would result in Hydrogen Hybrid Internal Combustion Engine (HHICE) vehicles.

The cost differential between a hydrogen hybrid powered vehicle and a regular diesel or gasoline-powered vehicle is much smaller than the leap to fuel cells. For example estimates of the cost of a HHICE bus is about 30-40% of the cost of a fuel cell bus. It is about 1.5 to 2 times more expensive than a conventional diesel powered bus. The smaller differential between HICE technology and conventional technology increases the potential for breaking the fuel supply problem for hydrogen.

Due to these advantages relative to fuel cell buses the Governments of Canada, California and Manitoba, along with the United States Department of Transportation and other partners[5] funded the development and testing of a HHICE bus in 2004/05. The winter testing was performed in Winnipeg during February and March 2005. This paper describes the test vehicle, and the test and provides advise in the way of a description of some of the knowledge gained form the project.

The HHICE Bus

The HHICE bus is a New Flyer Glider low floor bus, hybridized as follows:

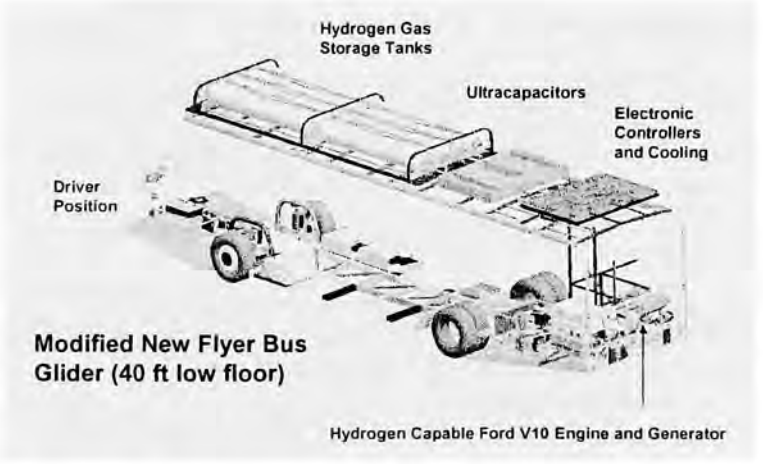
- The engine is used to generate electricity.

- Drive power is by electric motor/ generators.
- Energy recapture occurs with storage in ultracapacitor banks.
- The ultracapacitors provide burst energy for acceleration.

The engine used to generate the electricity is a Ford Power Products V-10 designed to operate on hydrogen. The engine generator combination creates about 150 kilowatts of power.

A schematic of the bus design is as follows:

Figure 1: HHICE Bus Design



The engine electric motor/generator was designed to fit in a cradle replacing the existing diesel engine and transmission. A bank of eight

hydrogen tanks were stationed on the roof of the bus, along with the ultracapacitor pack and associated electronic controllers and coolers.

The HHICE bus provides significant pollution reduction compared to regular diesel transit buses as shown in the following table.

Table 1: Comparative Emission Levels[6]

Bus Drive Type	NOx (g/km)	CO2 (g/km)	PM (g/km)
Diesel	19	1600	.15
Hybrid Diesel	9	1500	.01
HHICE	.3	0	0

The only emission of any significance is NOx that is less than 2% of that emitted from a regular diesel transit bus.

THE TEST

Overall, the HHICE bus test program was a success, with the bus operating in regular for fare service at Sunline Transit Agency starting in December 2004. From December until mid January when it was shipped to Winnipeg the bus performed without breakdown for 6000 kilometres. The service cycle in Manitoba was less, however, when in operation the bus met expectations.

The Winnipeg test began in mid February and run through to mid March 2005. Upon arrival a number of activities were performed in order for the bus to enter service. These included:

- Safety inspections.
- Testing all systems on board to ensure no damage occurred during shipping.
- Commissioning the hydrogen electrolyzer and re-fueller system.

- Safety training for drivers, persons doing the re-fuelling and mechanics that maintained the bus.

This time period was selected since the daily average temperature –14 degrees Celsius. This compares to a daily average temperature of 17 degrees Celsius in Palm Desert for the December to January time period.

Similar to the Palm Desert portion of the test the bus was operated in regular for fare service. The route selected crossed Winnipeg from the Northwest to Southeast. The total length of each trip was 32 kilometres return.

The bus was shipped back to Sunline Transit Agency at the end of thee test and returned to service.

Preparation for the test required the building, pretesting, shipping and putting into service of the bus, as well as the acquisition of an electrolyzer to create hydrogen and the renting of a suitable storage site. Since hydrogen is deemed an explosive material storage sites must meet substantially higher building code than typical of most service garages and transit bus storage sites.

The overall project was divided into two major projects. The first focused on all activities related directly to the bus. The second related to infrastructure, namely, the electrolyzer and storage. Principle components of the first project, in addition to the building of the bus were: arranging for importation with Transport Canada and the Canada Customs and Border Services Agency, creating the legal frameworks needed, training the City of Winnipeg Transit Department drivers, conducting safety training and testing consumer response to the vehicle. The second dealt with arranging for the electrolyzer and refueller to be stationed at Red River College and acquiring a suitable fire commissioner approved storage and maintenance site.

The project was managed through an overall steering committee with representation from Natural Resources Canada, the Government of Manitoba, the Vehicle Technology Centre, the City of Winnipeg Transit Department, New Flyer Industries, Red River College, ISE Corporation, and the University of Manitoba. From start to completion the project took 16 months.

KNOWLEDGE GAINED

The technology used in the bus proved to work well in both warm and cold climates and is ready for greater commercialization. The concept also proved, in a general sense, that HHICE technology may be the bridge technology to enable hydrogen engine systems and related fuelling infrastructure to become more commonplace.

One of the key lessons learned was that projects dealing with new technologies such as those used in the HHICE bus and with hydrogen fuels need long lead times. In addition to the time needed to construct the vehicle, time is needed to find suitable storage locations, which are not commonplace, and to access the limited electrolyzer capacity in North America. While 16 months seems like a relatively long time period, the project team members moved many mountains in order to complete the project in that time period.

In total there were thirteen direct partners involved in the project to test the bus in Winnipeg. This excludes partners such as the U.S. Department of Transport and the South Coast Energy Management Board that were indirectly involved through the test in Palm Desert. Partnerships require legal agreements describe the terms of the relationships. The more partners the more complex the agreements, in terms of both content and number. For this project the agreements ranged from relatively simply letters of understanding to complex lease agreements, requiring legal counsel. The framework used for this demonstration is shown in Appendix A. Overall due to the number of partners and generally the assignment of signing authorities at senior organizational levels it took about 9 months for

the legal agreements to be completed. Reducing the number of partners and assigning authority at the operational level facilitates the legal agreement process.

Related to the legal agreements was the need to implement a risk management strategy for the project. For the HHICE bus cold weather demonstration, risk was managed in the following ways:

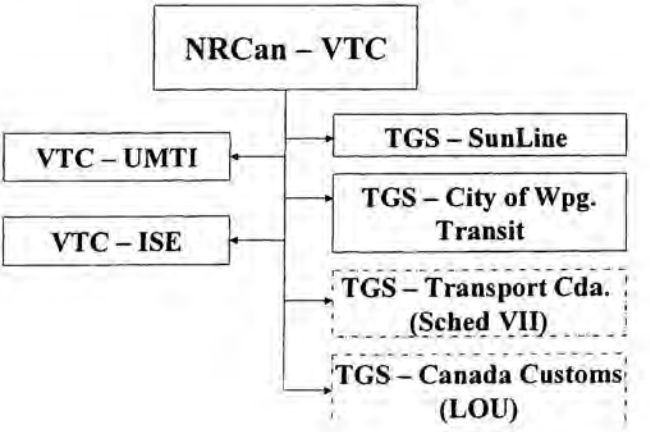
- Reasonable timelines were negotiated. While relatively tight, the actual timeline was significantly longer than when the project was first envisioned.
- Training with respect to safety of hydrogen, safe operation of the hydrogen systems, and operation of the bus was provided before the bus was put into operation.
- Access was limited to the storage site and the electrolyzer/refuelling site.
- Insurance was acquired to protect all parties involved in the project from liability actions, plus to protect the physical assets of the project.
- Project managers spoke uniformly with a "no" response to any activities that bordered on breaching any of the agreements.

Insurance coverage due to its customized nature, the risk related to hydrogen, and the number of partners covered was quite expensive

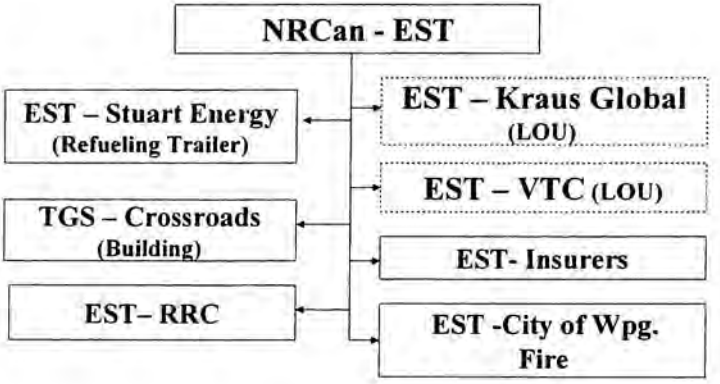
Appendix A

Legal Agreement Structure

Bus Conversion & Operation



Infrastructure



Legend:

NRCan = National Resources Canada
VTC= Vehicle Technology Centre
TGS = Manitoba Department of Transportation and Government Services
LOU = Letter of Understanding
UMTI = University of Manitoba Transport Institute
ISE = ISE Corporation
EST = Manitoba Department of Energy, Science and Technology
RRC = Red River College

Endnotes

[1] Advance Automotive news, October, 2003, http://www.rmi.org/sitepages/pid1065.php

[2] Fuel Cell Systems: A Survey of Worldwide Activity, Kerr- Ann Adamson, Alexandra Baker and David Jollie, Fuel Cell Today, December 2004. Page 4 and 5. http://www.fuelcellstoday.com/FuelCellToday/FCTFiles/FCTArticleFiles/Article_924_WorldWideSurvey2004.pdf

[3] Fuel Cell Market Survey: Buses, Kerri- Ann Adamson, Fuel Cell Today, November 2004. Page 1http://www.fuelcellstoday.com/FuelCellToday/FCTFiles/FCTArticleFiles/Article_916_BusSurvey%20Final%20Version.pdf.

[4] ABI Research, New York, September 2004. http://www.abiresearch.com/pdfs/AFC04pr.pdf

[5] The list of project partners were as follows: Natural Resources Canada, Government of Canada Western Economic Diversification, United States Department of Transportation Federal Transportation Administration, Government of Manitoba Transportation and Government Services, Government of Manitoba, Energy, Science and Technology, California Energy Commission, South Cost Air Quality Management District, Sunline Transit Agency, City of Winnipeg Transit Department, Ford Motor Company, ISE Corporation, New Flyer Corporation, Stuart Energy, Kraus Global, Vehicle technology Centre, Red River College, University of Manitoba Transport Institute.

[6] ISE Corporation: Proposal for a Hydrogen Hybrid internal Combustion Engine (HHICE) Bus Demonstration Program, Page 2.

An Analytical Model for Determination of Road Safety Index

MAHMOUD SAFFRZADEH
Associate Professor, Tarbiat Modares University, Tehran, Iran
Email: Saffar_M@Modares.ac.ir
MAGHSOOD POORYARI
Research Associate, Transportation Institute, Tehran-Iran
pooryari@rahiran.ir
HAMIDREZA BAHRAMIAN
Research Associate, Transportation Institute, Tehran-Iran
bahramian@rahiran.ir
Tel: 009821-8889979 - Fax:009821-8889979

ABSTRACT - In this paper, different variables affecting the rate of highway accidents were studied and a mathematical model for the assessment of highway safety was proposed. Field data as well as necessary information was collected for about 580km of main roads including those of different classes and various terrain conditions throughout the country. The modeling process was followed using linear functions, based on the collected data from the variables. Comparing the correlation coefficients of these models, the Poisons Regression Model was perceived as the most appropriate one, which was then calibrated and verified using the information for a six-year period. The dependent variable was defined as an accident index, AI, representing the number of accidents and their severity. Formation of commercial and residential land uses by the roadside considerably increases the rate of accidents. The mountainous roads are thirteen times more dangerous than the level roads Moreover, increasing the rate of heavy vehicles by 10%, will increase AI by 60%. Also changing the function of a two-lane highway to a one-lane results in a decrease of AI by 60%. Addition of one lane to the roadway, and increasing its shoulder width by one meter will improve safety index by 50 and 60 percent, respectively.

Keywords: *Safety Index, Traffic Accidents, Roadside Hazard,*

1- Introduction

Nowadays, accidents prediction models are considered as powerful tools in accident analysis and are used in identifying and analyzing roads with high accident rates. These models, which are obtained based on assessment of statistics and information on roads and accidents, are not only effective in geometric improvements and administrative evaluations, but also make it possible to easily identify and analyze high accident locations.

As some of the factors affecting road accidents are qualitative and happen in nature at random (such as driver behavior and weather condition), the real correlation between accident and the underlying factors is empirical and probabilistic.

For example, a completely ideal road from design perspective, can be unsafe in case of having careless drivers. Therefore, modeling correlations between those factors and accidents needs a statistical parameter which encompasses the scattered nature of accidents, while describing the interactions of reciprocal factors.

In order to obtain real and reliable results from the model, proper accident data, traffic, geometric design and terrain from both qualitative and quantitative aspects are needed. One of the reasons that makes the development of scientific models of accident predictions difficult, is the lack of expertise and scientific works in collecting and recording data and its improper classification.

In this research, it is tried to offer a model to predict accidents on different rural roads in Iran in order to estimate safety conditions of roads. One of the complexities of transportation in Iran is the high accident rate and fatalities.

Comparing accident rates shows that the index of fatalities in Iran for the year 1999 was 33 per 10000 vehicles, while this rate in most developing countries is less than 20 and in developed countries is less than 5 [1].

2- Research Background

Studies conducted in last decades show that design components are effective in road safety. Design components include characteristics of Lane cross section, horizontal profile, vertical profile, shoulders, cross section design (intersection and interchange),

lighting, accessibilities and methods of access control and quality of pavement. In most of studies carried out and models offered in this field for finding the correlation between accidents rates and design components, regression analysis is used. Some of which will be discussed as follows.

By using Highway Safety Information System (HSIS) data base and accidents data reported from 1985 through 1990 following predictive equation was drawn [2]:

$$Y = 0.0002(DVMT)^{1.073} EXP(0.131X_1 - 0.151X_2 + 0.034X_3 + 0.163X_4 + 0.052X_5 - 0.572X_6 - 0.094X_7 \; 0.003X_8 + 0.429X_9) \quad (1)$$

where:

Y: predicted annual accidents,
DVMT: Daily Vehicle Miles of Travel,
X_1: Average roadside hazard rating (1 to 7 from less dangerous to the most dangerous),
X_2: Access control (no control =0, Partial control=1),
X_3: Driveways per mile,
X_4: Intersections with turn lanes per mile,
X_5: Intersections without turn lanes per mile,
X_6:Function class (rural Principal arterial=1, rural minor arterial/collector=0),
X_7: shoulder width (feet),
X_8: median width (feet),
X_9: Area Location type (rural municipal=1) rural non-municipal=0,

In this equation, width of road pavement, median type and percentage of heavy vehicles were parameters, which had trivial effect in accidents and were not considered in the model.

Neumann has proposed an accident model using components of road cross section in sections of straight road according to Equation 2 [3]:

$$A = 0.0019(ADT)^{0.8824}(0.8786)^{W} (0.9192)^{PA}(0.9316)^{UP}(1.2365)^{H}(1.3221)^{Ter2} \quad (2)$$

where:

A: Number of annual accidents per mile (related accidents),
ADT: Average daily traffic volume in both directions,
W: Lane width in feet,
PA: Width of paved shoulder in feet,
UP: Width of unpaved shoulder (gravel, clod) in feet,
H: Distance of dangerous objects from width of road with index 1 to 7 from hazardless to hazardous,
Ter_2: One for mountainous roads and zero for other roads.

Dart used a regression analysis to determine the amount of interference of variables in the total accidents with respect to dry and wet road condition and in day or night time, after calculating correlation coefficient and square of 10 variables and came up with the following regression model [4]:

$$Y = 41.32 - 1.23X_1 - 0.54X_2 - 0.67X_6 + 0.03X1 \times X_2 + 0.03X_2 \times X_6 - 0.0009X_2 + 0.026X_{11} - 0.12X_4 \times X_{11} + 0.009X_5 \times X_9 \quad (3)$$

where:
Y: Total of accidents for 100 million vehicles per mile,
X_1: Percentage of commercial vehicles,
X_2: Ratio of peak traffic volume to service level of 13,
X_4: Lane width (feet),
X_5: Shoulder width (feet),
X_6: Super elevation,
X_9: Horizontal,
X_{11}: Traffic conjunctions.

- Road Surface Width (RSW)

This variable expresses width of the road (total sum of both traffic directions) in chosen sections and its unit is in meters. In roads with divided lanes, width of the road in only one direction is taken into account.

- Shoulder Width (SW)

This variable, takes into account shoulder width in different sections of the road and its unit is in meter. This includes paved and unpaved surface and in case of divided lanes, width of the right shoulder is taken into account.

- Percentage of Commercial Vehicles (PCV)

Considering the unique movement of commercial vehicles, this variable is considered in its involvement in road accidents.

- Access per Kilometer (APK)

This variable includes cross roads and ramps in sections under study and number of access to commercial centers, offices, complexes and etc... For this reason, the number of cross roads, exit and entrance ramps are taken into account.

- Terrain Conditions (TC)

This variable is used to consider geographical condition of the area where the road passes through. Due to the unsafe conditions of, mountainous roads and because of restrictions imposed by geometric criteria, this variable is entered into the equation separately. In this regard, using dual variables one and zero for mountainous and other roads are assigned as values.

- Roadside Hazard (RH)

Since a considerable amount of accidents results from run-off-road, this variable deals with safety of the roadside. In this research, percentage of accidents involving vehicles crashing another vehicle, fixed object, turning over or run-off-road is considered.

- Area Location Type (ALT)

One of the important issues in road safety is access control. Formation of commercial, residential and agricultural land uses along the highways results in the mobility of non motorized vehicles and pedestrians in the road areas adding to the considerable insecurity of a road. This variable is used as a qualitative one assigning values one or zero considering population density in the area or density of land use.

- Functional Class (FC)

The rate and type of accident in a divided highway is different from that of a one way road. For this reason, two variables one and zero are used for two way and one-way roads, respectively.

- Accident Index (AI)

This variable indicates the accidents occurred in a section of the road. Since accidents in a road are three types, property damage, injury and fatality, research resulted from Transportation Research Laboratories of England [6] regarding developing countries has been used to equate these three types. Equivalent index used in Equation 1 is calculated as follows:

$$EDA = DA + 3 \times ID + 12 \times FA \qquad (4)$$

EDA: Accident damage Equivalent

DA: Accident damage

ID: Injury damage

FA: Fatal accident

Then the variable AI is defined as:

$$AI = EDA / L \qquad (5)$$

where:

L is the length of section of the road in kilometers. One of the major advantages of this model is consideration given to severity of accidents.

One of the main shortcomings of the accident prediction model is the use of different variables such as the number of accidents or rate of accidents and no differentiation is made in the kind of accidents.

Hence in this research, this shortcoming is corrected and different kinds of accidents are equated to take into consideration the relevant elements and severity of accidents.

- Metrological factors

As a result of different analyses from accidents, the number of accidents varies in different times of the year as the result of traffic volume and any increase in the number of accident is due to the increasing traffic volume. Because of lack of major difference in the number of accidents in bad weather conditions, no variable is introduced for it. Results of the collected data for the model are given in Table 1.

5- Modeling

The modeling process is shown in Figure 1. To analyze the model, multi linear regression is used with the help of software SPSS (Statistical Package of Social Science). To obtain the model, linear regression method is used and different function forms is used in a linear fashion. After preliminary analysis, appropriate model is obtained with respect to control criterion. After evaluating and examining different conventional functions for accident prediction, the three following types of functions were chosen for the evaluation of road safety.

Results of the analysis of these functions are given in the following table, in which the equation gives good explanation of different variables used in the model.

Table 1: Specification of used variables in model analysis (1994-1999)

AI	RH	ADT	PCV	ALT	TC	APK	FC	RSW	SW	
13.4	4.8	364.7	6.1	0	0	0.3	0	11	2	1
6.5	10.3	6551	33	0	0	0.1	0	11	3.25	2
4.1	2.9	19190	72	0	0	0.1	1	8.6	4.3	3
5.9	9.4	4364	42	0	0	0.08	0	11	2	4
7.1	8.7	4893	47	0	0	0.08	0	11	2.25	5
9	1.1	39088	30	1	0	0.1	1	22	3	6
8	3.8	3923	41	0	1	0.3	1	7	1.5	7
4.2	3.2	5137	51.4	1	0	0.22	1	7.3	4	8
38	1.4	16680	25	1	0	0.12	1	7.3	3	9
24	3.9	8368	50	0	1	0	1	8.8	2.25	10
0.6	6.2	2557	27	0	1	0	1	7	3	11
17.8	4.8	35724	8.5	0	0	0.3	0	11	2	12
9	10.3	5929	33	0	0	0.1	0	11	3.25	13
6	2.9	11681	80	0	0	0.1	1	8.6	4.3	14
4.9	9.4	5137	53	0	0	0.08	0	11	2	15
4.6	8.7	5015	42	0	0	0.08	0	11	2.25	16
7.6	1.1	22110	56	1	0	0.1	1	22	3	17
9.6	3.8	4086	41	0	1	0.3	1	7	1.5	18
2	3.2	3921	60	1	0	0.22	1	7.3	4	19
36.1	1.4	24630	18	1	0	0.12	1	7.3	3	20
19.4	3.9	8377	50	0	1	0	1	8.8	2.25	21
1.2	6.2	2092	34.4	0	1	0	1	7	2	22
22.2	4.8	44684	6	0	0	0.3	0	11	2	23
12.6	10.3	8655	37	0	0	0.1	0	11	3.25	24
3.9	2.9	16986	88	0	0	0.1	1	8.6	4.3	25
7.6	1.4	21459	50	0	1	0.1	0	22	3	26
5.7	3.2	4382	66	1	0	0.2	1	7.3	4	27
73	1.4	21518	34	1	0	0.1	1	7.3	3	28
49.7	3.9	9712	57	0	1	0	1	8.8	2.25	29
2.3	6.2	2176	41	0	1	0	1	7	2	30

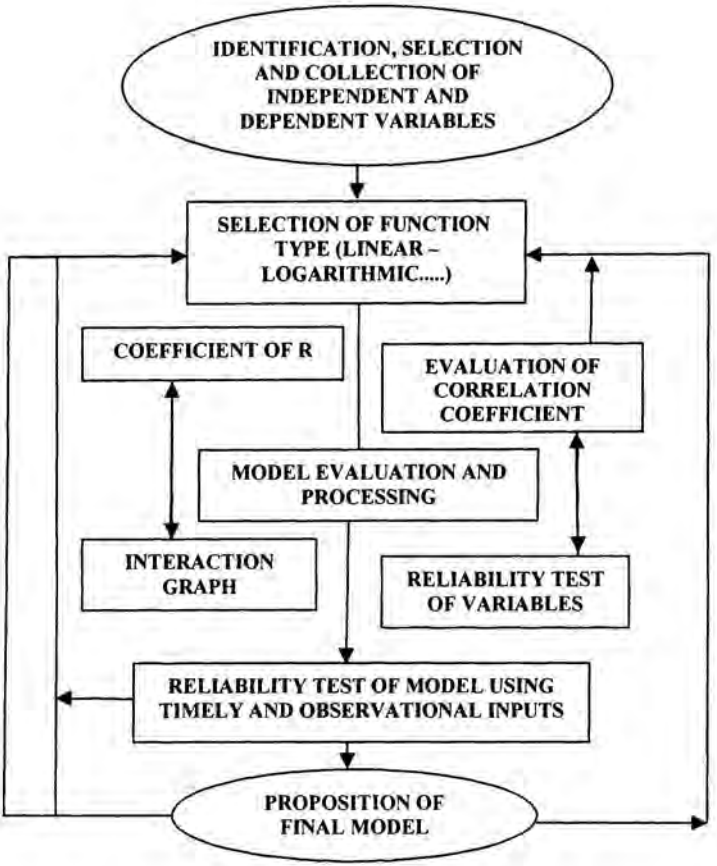


Figure 1: Process of research modeling

General form of equation to evaluate the model is based on the following terms

$$Y = A + \sum (B_i . X_i) \quad (8)$$

$$LnY = A + \sum (B_i . X_i) \quad (9)$$

$$Y = A . X^B . EXP(\sum_{Bi . xi}) \quad (10)$$

where:

Y = contingent variable (number of annual accidents in unit of road (AI)),

X = independent variable, Average Daily Traffic (ADT),

X_i = Other variables are independent and are mentioned in Equation 5

A, B, B_i = Fixed variables.

The results of preliminary processing of models are shown in Table 2.

Table 2: Results of preliminary regression analysis of model's functions

Model	R	R^2	R^{2a}
8	0.774	0.598	0.418
9	0.687	0.472	0.234
10	0.929	0.863	0.801

Variable (R^2) as a correlation coefficient to justify dispersal of relevant variables is used as the result of independent variables. The high values of R^2 in the above table, is indicative of proximity of data from observation and model's prediction of median inputs. Coefficient (R^{2a}) is another primary research quantity which is used to explain the dispersal of relevant variables as a result of independent variables.

In this regard, the analysis of model in Equation 1 is an appropriate model.

Results obtained from the correlation coefficient (R) between dependent and independent variables are shown in Table 3.

Table 3: Matrix of correlation coefficients between dependent and independent variables (Ln AI)

	Ln AI	SW	RSW	FC	ACC	TC	ATL	COV	RH	LN ADT
Ln AI	1.00	-0.21	0.05	-0.16	0.16	-0.14	0.19	-0.29	-0.28	0.61
SW	-0.21	1.00	0.00	-0.09	-0.09	-0.47	0.41	-.59	-0.27	0.14
RSW	0.05	0.00	1.00	-0.05	-0.05	-0.15	0.11	-0.05	-0.14	0.48
FC	-0.06	0.29	-0.32	-0.19	-0.19	0.35	0.46	0.39	-0.69	-0.14
ACC	0.16	-0.09	-0.05	1.00	1.00	-0. 31	0.15	-0.32	-0.18	0.31
TC	-0.14	-0.47	-0.15	-0.31	-0.31	1.00	-0.39	0.03	-0.11	-0.42
ATL	0.19	0.41	0.11	0.15	0.15	-0.39	1.00	0.00	-0.58	0.22
COV	-0.29	0.56	-0.05	-0.32	-0.32	0.03	0.00	1.00	-0.16	-0.26
RH	-0.28	-0.27	-0.14	-0.69	-0.18	-0.11	-0.58	-0.16	1.00	-0.49
Ln ADT	0.61	0.14	0.48	-0.14	0.31	-0.42	0.22	-0.26	-0.49	1.00

The degree of correlation between accidents and amount of daily traffic variable shows the importance and effect of this variable in the model. Low correlation between independent variables shows that each one of these independent variables had no effect on the other

variables and the value of both independent variables is protected independently, which is the result of choosing proper variables.

Since regression coefficient provides an intuitive understanding of the importance of variables, the low value of some dependent coefficient variables with dependent variables, does not reduce the reliability of variable, but it causes doubt about the importance of the variable which must be tested again in (t-student). Variable (t) expresses whether or not a variable is important in predicting accidents?

With respect to the result of analysis of experiment (t), it was shown that except for the variable of access per kilometer (APK), the rest of the variables have reliability of 90 percent and higher. By eliminating the variable of access per kilometers (APK), the obtained coefficients from the final regression model is as given in Equation 11.

$$AI = 0.01144 \times (ADT)^{2.41} \times EXP(0.45 \times RH + 0.047 \times COV + 3.41ALT + 2.626 \times TC - 0.83 \times FC - 0.203 \times RSW - 0.735 \times SW) \quad (11)$$

Results of this analysis are included for evaluating safety condition of roads with different functional class and terrain condition is simultaneous effect of effective parameters. In order to identify which variables had the most effect on the model, using a software in modeling process, we started from a model which only had fixed value and in every step added a variable to the model which caused the most change in (R^2). Results indicated that average daily traffic (ADT), width of shoulder (SW) and width of road surface (RSW) had the most effects in accident rate, respectively. ADT variable increases the accident rate and variables SW and RSW cause to reduce accident rates. Results of evaluation of model are presented in Figure 2.

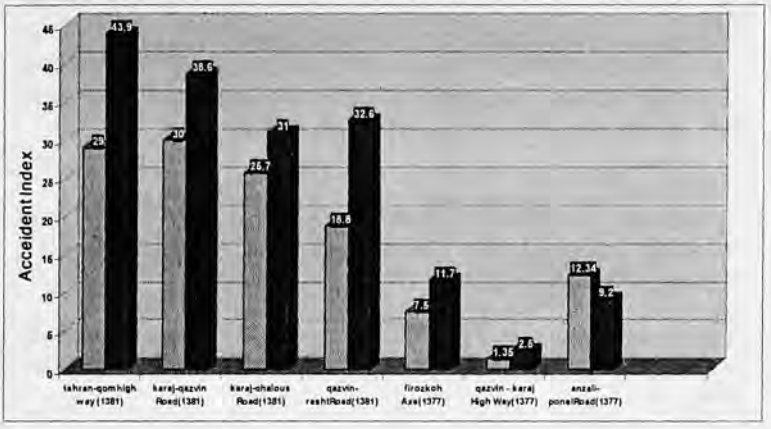


Figure 2: Result of evaluation of model in accident prediction

6- Conclusion

Regarding the high degree of variable (ADT) in the obtained model, the influence of accidents from this variable is considerable. In other words, human factors and transportation fleet are major elements of road safety.

For example, based on the model. Adding 33 percent to traffic volume, increases accident possibilities by 100 percent, adding one meter to width of the shoulder will improve road safety by 50 percent. Adding one lane to traffic lanes will result in 50 percent reduction in accidents. But the effect of these variables outside standards given in the model is not clear. Formation of any land use in road environment will add significantly to road unsafely.

In the case of road function (FC), it is worth to emphasis that separation of traffic lanes in each direction from road to highway improves safety and reduces accidents by 60 percent.

Based on results obtained, commercial vehicles are one of the major causes of road safety. Assuming a 20 percent reduction in commercial vehicle fleet, accidents will be reduced by 60 percent.

Shoulders is another factor in road accidents. use of road restraints, signs and sign boards especially in mountainous roads and

high density accident areas is very important. Also the effect of road environment is significant in accidents and mountainous roads are 13 times more unsafe than other roads. Moreover, in this study it was shown that adverse weather conditions and the number of cross roads per kilometer do not have major effects in rural roads.

7- References

[1] Bureau of Information Technology, Road Accidents: A Hidden Disaster, Transportation Monthly in One Glance, No. 43, Transportation and Terminal Organization, 2003.

[2] "Safely Effects of Cross Section Design on Rural Multilane Highways", Publication No. FHWA- RD-97-027, Issued September 1997.

[3] Neuman, Timothy R., Glenon, John, Sagg. James "Accident Analysis for Highway Curves", TRR 923, Transportation Board, Washington, D.C. 1996.

[4] Dart, Olin K. Jr., and Lawrence J. Mann. "Relationship of Rural Highway Geometry to Accident Rates in Louisiana", Louisians State University Library, 1995.

[5] David M. levinson, Adib Kanafani, David Gillen "The Costs of Intercity Passenger Transportation", University of California at Berkeley, November 1997.

[6] "Towards Safer Roads in Developing Countries", TRL 2441, 1st. Edition 1991.

[7] Ruediger Lamm, Basil Psarianos, "Highway Design and Traffic Engineering Safety Handbook"/ Theodor Mailaen der, 1999.

[8] "Safety Effects and Operational Practices of Highways and Streets", Report FHWA A-RD-78-218, US Department of Transportation, Feb. 1978.

TRAVEL DEMAND MODEL IN RAILWAY TRANSPORTATION NETWORK, USING THE VECTOR AUTOCORRELATION ROUTINE (VAR)

M. Saffarzadeh[1], A. Rezaee-Arjroody[2] and P. Bazdar-Ardabili[2]
1. Department of Civil Engineering, Tarbiat Modares University, Tehran, Iran
2. Transportation Research Institute, Tehran, Iran

ABSTRACT: In this paper, the travel demand in the railway of the Islamic Republic of Iran is estimated through the vector autocorrelation routine and convergence during 1971-2009. The calculations show that application of above mentioned methods will benefit the economy. The alternatives used in the paper include: per capita trips, per capita income, price of train ticket, length of railway main lines, etc. Variance analysis methods, traumatic reaction functions, and stability rate are all used to study the model dynamic behavior.
Results reveal that the travel demand will increase up to 17,000,000 in 2009. The average travel growth rate and Gross National Production during 1971 to 2009 will be %5.01 and %4.72, respectively, and the total travel growth and Gross National Production during the said period will be %195.5 and %184, respectively.

1. INTRODUCTION

Today, movement and transfer are among the most important requirements which are led to economic results. In economics, services that cause transfer and relocation of production resources,

have economic value and are considered as part of production process. Therefore, transportation is an economic requirement, which requires attention and consideration more than ever. Transportation system can lead to centralization or non-centralization, prosperity or stagnation, employment or unemployment and many other effective changes in economic activities and generate remarkable added value, which is one of the criteria for assessing economic growth and development of countries.

Comparing different transportation systems, one can find countless advantages for railway transportation. Among these, we can mention such issues as security, comfort and convenience of travel, easy access to ticket, cheapness, saving in fuel consumption, less air pollution, high ability and potentiality for relocating passengers and cargo, saving in national income and less land occupation [Malekpour,2000] .

During 1971-2001, the number of passengers increased from 4,000,000 people to 13,000,000, which shows an increase of % 225 and the average annual growth rate of % 5.4. Some of the effective factors in railway trips are stable - aside the conditions of time and location - and they are a function of economic and social factors such as income, population, price and ... There are some other factors, that encounter change because of location and time conditions, among which we can mention social category of the applicants, tastes of the applicants, technology and welfare facilities provided for the customers. In this study, considering the long-term relation between per capita travels and per capita income and using VAR methods and autocorrelation, the travel demand function is estimated within 1971-2001 and finally the travel demand is anticipated till 2009 [Islamic Republic of Iran Railway, 2001].

2. TRAVEL DEMAND MODEL

The number of passengers relocated by railway is a function of the number of passengers and also the average distance (Km) for each passenger to be transferred. Therefore, any factor to influence one of

these alternatives will change the number of passengers transferred by railway. At the beginning of this section, factor influencing these two alternatives are studied, and in the next stage, through an analytical study, these two alternatives are studied from quantitative point of view [Gojarati, 1999].

2.1. Effective Factors on Number of Trips

The first factor- i.e. number of trips- depend on population, level of income, population and common level of prices. The impression mechanism of these factors will be discussed in details.

2.2. Population

Along with population increase, travels are more likely to take place. So, even in conditions where other factors are unchanging, the growth of population will lead to increased demand for travel, in general, and increased demand for traveling with railway increased in particular. Therefore, the population growth and number of trips that are resulted by the factor of population have been considered as equal, then for studying the influence of other factors through the use of per capita alternatives in the country, the influence of population has been deleted.

2.3. Income and National Production

When people's income increases, the expenses of travel in total budget of the family increases too, specially since trip is not an essential requirement, the share of travel expenses will grow in the total income. So, we can predict that the travel demand elasticity is bigger than one and the speed of increase of travels is more than the speed income growth. On the whole, we can state that business trips are directly influenced by total production and non-business trips depend indirectly on total production and directly on individual consumption cost, but in theoretical discussions about the individual

consumption in economics, individual consumption finally relates to total production, therefore the alternative of total production can explain both factors of the number of trips. In this paper, the national gross production at a fixed price is used as the state production index. Also for elimination of the influence of population on the national gross production, the corresponding transactions have been used as per capita [Hodge,1999].

2.3.1. Prices

Price is one of the most important factors in determining the demand rate. In fact, it is the most significant economic indicator, on the basis of which the economists make their decisions. For a named good, the price increase, reduces the demand, and when the prices fall, the demand increases. The demand rate for one good, not only depends on its price, but also it is a function of the price of other goods. Depending on the type of relationships between goods, there are two kinds of relation between the considered good and other goods as substitution and supplementary.
In principal, for selecting the means of travel, there are different choices, that are replaceable with each other, the most important of which are as follows; airplane, train, road transportation means, including bus, minibus and ...
Another important alternative which has a remarkable desirability is the use of private cars. The demand for traveling with trains has a reverse relation with price of the ticket and has a direct relation with competitors prices. In fact, we can say that demand for train has a relationships with price of the ticket and other alternatives [Simon, 1999].

2.3.2. Factors Effective on Length of the Travel

The factors which affect the length of travel time and as a result on average length of the trip include: Demographic distribution, form and volume of country development, and existence of travel attractions in the country. The more demographic distribution in the

country, the longer destinations people have to travel. As a result, the average of travel lengths will increase. Another factor is the existence of travel attractions, for instance if in a country, there is an important pilgrimage or tourism spot, which attracts many visitors, then the population of the country will locate in the average length of the affected geographical location and within the distance of this spot from other centers. The third factor which is effective on the average length of the travel, is the form of railway network. Certainly if there were no Tehran-Isfahan-Kerman line or Tehran-Ahwaz line in Iran's present network, or the railway lines were restricted to Tehran-Rey line, the travel applicants were not able to choose railway alternative for long distances of their trips. So, following the increase of the length of railway lines, it is expected that the average length of each journey increases too. In this study, hypothesizing the equal growth of population in all centers and low rate of the volume of immigrations as compared with total population, the population distribution has been deleted from the calculations [Mahmoodi, 2002].

The two factors, the effects of which on the average length of the travel have been considered, are the passenger attractions and form of the network. Generally, the effect of passenger attractions can also be deleted like the factor of population distribution, because of stability of these centers. If for any reason the passenger attractions change during the study period, we must take them into consideration. In Iran, since the beginning of the imposed war, the average number of travels has increased due to transportation of military forces to west and south of the country by railway in the routes, the lengths of which have been mostly more than the travel average during the past years. In order to show this process, in this study an alternative (a virtual alternative) has been used, amount of which has been zero and one during the years before and after war, respectively, and this shows an increase in travel average after beginning of the imposed war. In order to show the effect of the form and breakage of the network on the travel average, the alternative of length of railway lines during the past twenty years has been used. In principal, through studying the situations of railway lines and the average of travel length in the last twenty years, we come to see that whenever the length of railway lines

is increased, the travels average has been also raised and in other times, the average travel length has been almost fixed and proportionate to the length of the lines [Kamanta,1998].

3. STRUCTURE OF THE MODEL

In this section, we will use econometric analyses to determine the quantitative relations existing among the economic alternatives. Basically, the methodology of econometric indicates that the fulfillment of this target, first of all requires a mathematical model, which may be used to explain the relation between above-mentioned alternatives in mathematical language, and secondly it requires proper statistical methods that can help estimating the existing parameters in that mathematical model. Such statistical methods, which are commonly called "estimating methods" are based on statistical observations. Also in order to estimate all models, the computer package called "Eviews" has been used. The model used in this paper is as follows [Mazreati, 2000]:

[1] $Pa = f(GDP, P, Lra, Dum, T)$

In which:
Pa : Per capita travel,
GDP : Per capita income,
P : Ticket price,
Lra : Main airway lines,
Dum : Virtual alternative (war)
T : Time.
It should be noted that the alternatives are used in the real model (at the fixed price of 1990) and in the form of logarithm.
The Vector Auto Regression (VAR) model is in fact a linear relation between the dependent alternative and pauses in all alternatives, present in the system of equations. The number of pauses is

determined experimentally through Model-maker. The following equation system generally describes the VAR system with n dependent alternatives (n equations):

$$\begin{bmatrix} y_{1t} \\ y_{2t} \\ \cdot \\ \cdot \\ \cdot \\ y_{nt} \end{bmatrix} = \begin{bmatrix} A_{11}(L) & A_{12}(L)... & A_{1n}(L) \\ A_{21}(L) & A_{22}(L)... & A_{2n}(L) \\ & \ddots & \\ & \cdot & \\ & \cdot & \\ A_{n1}(L) & A_{n2}(L)... & A_{nn}(L) \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ \cdot \\ \cdot \\ \cdot \\ y_{nt} \end{bmatrix} + \begin{bmatrix} C_{1t} \\ C_{2t} \\ \cdot \\ \cdot \\ \cdot \\ y_{nt} \end{bmatrix} + \begin{bmatrix} C_{1t} \\ C_{2t} \\ \cdot \\ \cdot \\ \cdot \\ C_{nt} \end{bmatrix} + \begin{bmatrix} E_{1t} \\ E_{2t} \\ \cdot \\ \cdot \\ \cdot \\ E_{nt} \end{bmatrix}$$

[2] $$y_t = A(L)y + \varepsilon_t$$

In which:
L : Represents lapse,
$Ci, i = 1,...,n$: A part of width from the equations origin,
ε_t : Accidental intervention.

A matrix is defined as follows [3]:

[3] $$A_{ij}(L) = \sum_{k=1}^{k} L^k a_{ijh}$$

In which:
i : Number of the equation,
j : Number of alternative present in the equation,
k : Number of lapses.

4. EXPERIMENTAL RESULTS

In time series models, before any estimation, we must answer two principal questions, otherwise the results obtained from estimation will not be trustworthy and reliable. The questions are:
1- Are the model alternatives constant?
2- In case the alternatives are not constant, are they auto-correlated? (i.e is there meaningful long-term relation between them?)

4.1. Model Alternatives' Stability Test

In order to implement this test, we use the Dicky-Fouller generalized unit root test. The test method about the ideal series of y_t case is as follows. If the unit root test is implemented in the data levels, we should do the following estimation:

$$[4] \qquad \Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta y_{t-1} + \varepsilon t$$

In which:
Δ : Difference of first turn,
β_1 : Constant amount and t is the time trend,
ε_t : Error clause.

If this test is done for the first turn difference, we should perform the following estimation:

$$[5] \qquad \Delta\Delta y_t = \beta_1 + \beta_2 t + \delta \Delta y_{t-1} + \alpha_i \sum_{i=1}^{m} \Delta\Delta y_{t-1} + \varepsilon_t$$

In both above-mentioned cases, the zero hypothesis indicates that the time series y_t has a common root, or it is not inert (with passing of time, its variance and covariance increases and also its average is

dependent on time and is not a constant amount). According to AIC (Akaike Information Criterion) we determined the number of vector auto-correlated model pauses, the optimal pause is then defined when the amount of AIC is at minimum. On the basis of this criterion, the number of pauses has been selected as P=2.

Table 1. Root Unit Test

	(1) level of data	(2) First turn difference
Alternative	**ADF (2)**	**ADF (2)**
Pa	-2,52	3,15
GDP	-2,52	* 3,64
Lra	* 3,91	-1,3

- Meaningful on security level 95% of Macinson Crisis Amount
  Internal alternatives: Pa, GDP, P, Lra
  External alternatives: T, Dum

This test has been implemented with two pauses, and constant amount of © and existence of trend, the results of which have been presented in Table 1.

4.2. Co-Accumulation

According to unit root test, the statistic figures of Dicky-fouller, per capita travel alternatives and Gross Domestic Production and also the price of train ticket were unstable. On this basis, the Co-accumulation test was carried out, the results of which indicated the accumulation of the said alternatives and/or showed the existence of stability in the linear configuration of related alternatives. Because LR test, has proved the co-accumulation of the said alternatives in 0.05 level, so

we can use the said series without implementing differentiation in VAR equations.

4.3. Estimation of Auto-Regressive Vectors

Auto-regression equations have been done with two pauses for Lra, P, GDP, Pa alternatives, the results of which have been reflected as follows:

[6] PA= -13/27 – 0/27PA (-1) – 0/19PA (-2) + 1/99GDP (-1)- 0/98GOP(-2)-0/05P(-1) – 0/006P(-2)-1/11LRA(-1) + 1/5LRA(-2)- 0/07DUM +0/001T

[7] GDP= -0.71 – 0.04PA(-1) – 0.07PA(-2) – 1.33GDP(-1) – 0.52 GDP(-2)+ 0.032P(-1) – 0.013P(-2) – 0.23 LRA(-1) + 0.44LRA(-2) – 0.042 DUM + 0.0008T

[8] P= -23.79 – 0.26PA(-1) + 0.053PA(-2) 0.67GDP(-1)+ 0.37GDP(-2)+ 0.59 P(-1)+ 0.08P(-2)-2.7LRA (-1)+4.66LRA(-2) – 0.05 DUM – 0.03T

[9] LRA= 0.77+0.005PA(-1) +0.009 PA(-2) +0.13GDP(-1)- 0.08GDP(-2)+0.029P (-1) +0.006P(-2)+0.35LRA (-1)+0.52LRA(-2)- 0.004DUM +0.009T

Since the alternatives have been used as logarithms in the model, we can describe the coefficients as elasticity. According to the first equation we can say that one percent increase in the alternatives including, P, GDP, Pa and Lra, with one or two pauses will change -0,27, -0,19, -0,1,99 , -0,98, -0,05 , -0,006, -1,11, and 1,5 percent in Pa alternative respectively.

The explanation of the other three equations will be similarly the same.

4.4. Analysis of Variance

According to the statistics, tables and charts, we can state that during ten rounds, the highest amount of changes in Pa alternative will be explained by the alternative itself. Of course this explanation will be reduced with increasing of the length of the round, in a way that from 100% in the first round will decrease to 50% in the third round, and then it will remain constant. The amount of explanation of Pa alternative will be increased through the changes in GDP alternative; from 0% in the first round will reach 40% in the 3rd round and will be fixed thereafter. The amount of explanation of Pa by P and Lra is very little, so that it will be ignored [Mahmoodi, 2002].

Table 2. Variance Decomposition of PA(%)

period	LRA	P	GDP	PA	S.E.
1	0.000000	0.000000	0.000000	100.0000	0.109710
2	1.354134	0.129285	38.36439	60.15219	0.141819
3	2.444316	0.115731	43.84047	53.59948	0.150261
4	2.408285	0.228068	44.48737	52.87628	0.151395
5	2.789713	0.299824	44.25121	52.65925	0.151800
6	2.934118	0.385150	44.14670	52.53403	0.152039
7	3.171275	0.467406	44.00097	52.36035	0.152291
8	3.318113	0.565108	43.97264	52.14413	0.152619
9	3.481025	0.677009	44.12636	51.71560	0.153278
10	3.628257	0.812314	44.47704	51.08239	0.154254

4.5. Forecasting Travel Demand

One of the main goals in estimation of a regression model is to predict the changes in the function alternative per special amounts of the explanatory alternative. Due to the resulted estimations in the

previous section and figures of the basic year in 1990, and also the forecasts of the third development program regarding the population growth and growth of Gross Domestic Production, we can forecast the volume of travelers for the future.

The hypotheses used for these forecasts include:

- According to the forecasts of the Third Development Plan of the Islamic Republic of Iran, the Gross Domestic Production will have a growth of 6%.
- According to the forecasts, the population will have an annual growth of 1.65% annually.

The results derived from travel demand forecasting are illustrated in table 2 and also in figure 2. It is seem that the number of travelers has reached from 4,000,000 in 1971 to 17,000,000 in 2009. The results of variance analysis and traumatic reaction functions show that the impact of shock in per capita income will increase the per capita travel and will not have a remarkable impact on the price of train ticket and length of the main railway lines. The volume of travel and Gross Domestic Production will increase to 5.01% and 4.72%, respectively. Also through the end of travel demand forecasting period, the travel demand and Gross Domestic Production will have a growth of 195.5% and 184%, respectively.

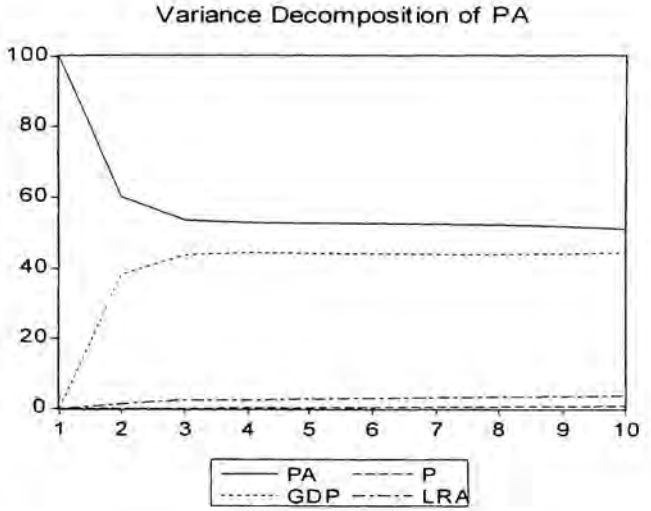


Figure 1. Variance Decomposition of PA

Number of travelers(million)

year

Figure2. Forecasting Travel Demand(1971-2009)

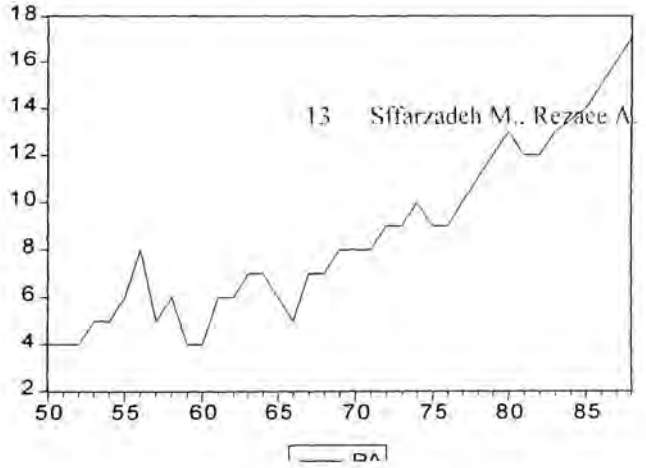


Table2. Travel Demand Forecasting on the Base of GDP

	Forecast		Growth rate (%)	
Period	Number of travelers (million)	Gross Production (Billion Rials)	Number of Travelers	Gross Production
1971	4	17039	-	15
1972	4	20263	-	19
1973	4	22332	-	10
1974	5	25853	25	16
1975	5	28496	-	10
1976	6	33501	20	18
1977	8	33162	33	-1
1978	5	30602	-37.5	-7
1979	6	29413	20	-4
1980	4	26596	-33	-10
1981	4	25472	-	-4
1982	6	27856	50	9
1983	6	31874	-	14
1984	7	32291	17	1
1985	7	32560	-	0.8
1986	6	29306	-14	-11
1987	5	28774	-17	-2
1988	7	27301	40	-5
1989	7	28765	-	5
1990	8	32788	14	14
1991	8	37357	-	14
1992	8	38454	-	3
1993	9	39281	13	2
1994	9	40123	-	2
1995	10	42679	11	6
1996	9	44810	-11	5
1997	9	45900	-	2
1998	10	46841	11	2
1999	11	49248	10	5
2000	12	51749	9	5

2001	13	53527	8	3
2002	12	56738	-8	6
2003	12	60143	-	6
2004	13	63751	8	6
2005	13	67577	-	6
2006	14	71631	8	6
2007	15	75929	7	6
2008	16	80489	6	6
2009	17	85313	6	6

5. CONCLUSION

Because of economic, environmental and traffic advantages, today the railway transportation system has been considered as one of the most favorable methods for transportation. As a result, the governments have started to focus on increasing their railway transportation share as a secure and safe means of transportation.

In this paper, the travel demand has been forecasted through the use of alternatives like per capita income, price of the train ticket and length of the main railway lines. The results reveal the external effects of the price of the ticket and railway main lines. This means that the travel demand is more likely to be affected by income rather than other alternatives of the model. The proposed models have special and important applications, particularly in forecasting and analyzing the economic periods. The growth rate of annual travel demand and the annual growth rate of Gross Domestic Production will be 5.01% and 4.72%, respectively. The results obtained from the stability of the model reveals that the relation achieved between the travel demand and other alternatives are constant, and they are statistically trustworthy for forecasting. During 1971-2009, the total travel growth will be 195.5% and the total growth of Gross Domestic Production will be 184%. So the results show that the travel demand will be increased from 4,000,000 in 1971 to 17,000,000 in 2009.

Finally, in order to increase the travel demand, the following suggestions are proposed for improving the railway system:

1- Investment in increasing the capacity of railway system and purchasing rail transportation devices.

2- Provide proper conditions for raising the level of responsibility, securing the rights of customers and using customer-focused methods.
3- Increasing productivity, qualitative development of human resources and improvement of the quality of services in railway transportation.
4- Using equipment and fleet with modern technology in railway transportation

6. REFERENCES

Gojarati, Damoudar, 1999, *Economics evaluation Basics*, Translator: Hamid Abrishami,Third Edition,Tehran, University of Tehran.

Hodge.J.Tyler, 1999, *A co-integration Analysis of U.S. Energy Demand Elasticities*, France.

Islamic Republic of Iran Railway, 2001, *Statistics of The activities of The Islamic Republic of Iran Railway.*

Kamanta, Yian, 1998, *Economic Evaluation Basics*, Translator : Kambiz Hozhabr Kiani, First Edition , Tehran, Center for University Publication.

Mahmoodi, Ali, 2002, *Economics of Transportation*, Tehran- Institute for Commercial Studies & Researches, Eghtesad-e-No Publishing Co.

Malekpour, M.R., 2000, *Privatization and structure of Iran's Transportation System*, Tehran, Iran.

Mazreati, Mohammad, 2000, *Comparison of performance in Forecasting VAR & BVAR Models*, PhD Dissertation, University of Tehran.

Simon, David, 1999, *Transport and development in third world*, England.

SHIP TRAFFIC SIMULATION IN THE PORT OF A DEVELOPING COUNTRY: EVIDENCE FROM NIGERIA

Mobolaji S. Stephens & Innocent C. Ogwude
Dept. of Transport Management Technology,
Federal University of Technology Owerri
Nigeria, 460008

ABSTRACT

This study set out to establish that land transport system in the country of destination determines the turn-around time, capacity utilisation of port infrastructure, facilities and cargo-handling equipments and general port performance. Of particular interest is the contribution of road transport infrastructures and system to the magnitude of turn-around time, port performance and general economic growth and development. This is true and significant for developing countries of Sub-Saharan Africa, where transport systems are poorly developed and sparsely integrated; but the economies are import oriented.

The port's rates of capacity utilisation was determined over a period 14 years (1990 – 2003) and traffic simulations were done for 156 vessels calling at the port and 19, 296 ladened road vehicles leaving the port between 1 December 2002 and 31 March 2003.

The result showed that over-utilisation of road transport resulted in under-utilisation of several port infrastructures, port congestion, longer turn-around time and general poor performance of the port. The need for a well-integrated transport system was therefore underlined.

Keywords: Capacity utilisation, turn-around time, berth effectiveness.

1 INTRODUCTION

Nigerian ports are presently owned and operated by Nigerian Ports Authority. The ports are located in Calabar, Lagos, Port Harcourt and Warri. Lagos has 4 ports, namely: Container Terminal; Lagos Port Complex; Roro-Port and Tin-Can Island Port. While Port Harcourt has 2 ports: Onne Lighter Terminal and Port Harcourt Port. Lagos Port Complex accommodates Container Terminal port. The working and operational arrangements of these ports are similar, hence a study done with one can be generalised for all. This study has Lagos Port Complex (LPC) has its case study. The over-utilisation of the road mode as a medium of conveying cargo out of the LPC, which is due to impatience and the need to save costs (cargo handling and rents) on the port users making them to opt for the use of direct cargo discharge method has resulted to poor port performance.

The total handling capacity of cargo throughput at the port is 9.5 million tones per annum. There are twenty-two berths and the total berth space in length is 3277 metres long of which 2110 metres (general cargo), 680 metres (dry bulk cargo), 200 metres (general cargo and passenger combined) and 287 metres (general cargo and fish combined). The total transit shed space available is 513588.57m^3 and 473677m^3 for warehousing space. The total capacity in tonnes of cargo handling plants, equipments and machinery is 6964 tones.

1.1 *Cargo Discharge Methods*

In the Lagos Port Complex, LPC there are two methods of cargo discharge, namely direct and indirect. Direct cargo discharge means the transfer of 90% of the content of a vessel, straight from that vessel onto waiting road vehicles and/or rail wagons (provided by the shipper(s) or his/their agent(s)) and out of the port. The remaining

10% is (later) thoroughly inspected by the customs. Note that customs were equally present or represented when the initial 90% was loaded. On the other hand, indirect discharge means that the cargo onboard a ship will be transferred and stored/stacked in a shed/stacking area for a later evacuation from the port.
Port operators in the bid to increase ship traffic prefer direct delivery particularly at times of heavy ship/cargo traffic to avoid congestion. The carriers also prefer this method of cargo delivery at the port as they assume that less time will be spent at such ports. Also, the port users prefers this method in their bid to minimise port cost particularly cargo handling costs and rents. This preferred method of delivery results in an immediate increase in demand for lands transportation at the LPC.

2.0 LITERATURE REVIEW

De Langon (1988) considered time as the most important factor in measuring the efficiency of any transport system. Pyre and Briggs (1989) suggested that, *"For any steady economic development, there must be an adequately and efficiently operated seaport"*. Benathan et al (1979) stated, *"The time a ship spends in a port affects cost per ton of cargo delivered, since port costs are part of total shipping cost. Reducing cargo-handling cost will have a great impact in reducing cost of commodities, improving world economies (particularly for economies that are import-oriented) and world's people standard of living."* All these underlined the importance of improving port performance and optimally utilising capacities.

According to Laing and Hecker (1989), the main justification for port investment is the reduction of ship's waiting time in over-crowded ports. The authors used estimates of berth occupancy from traffic and productivity forecast to determine waiting-to-service time ratio (w/s) of a port. They found that the cost of waiting time or time spent in the queue dominated total operating cost of a shipping company and

suggested a w/s ratio of 0.1 for ports with 15-20 berths. UNCTAD (1982) recommended an average queuing time of 19% of discharging or loading time as being close to the limit to which a port should be designed.

Wolfhard (1989) advised developing countries against expansion of facilities like berths or relocation of ports as a solution to port congestion, rather than ensuring that existing capacities are optimally utilised. According to him provision of more berths could lead to over-supplying of berths and later to under-utilisation.

3 METHODOLOGY

Road Traffic within the port, particularly at the exit points was observed with the assistance of field research workers –covering the three exit and the emergency lane. They recorded the service start times and end of service times for all out-going loaded road vehicles (Mondays - Fridays between 1100 – 1700 hours) for four (4) months December 2002 – March 2003. A total of 19, 296 such vehicles were recorded. The researcher during this period recorded the arrival time for each of the road vehicles.

Ship Traffic for vessels calling at the port during the same period was recorded daily from Ship Traffic Data Report, Quay Record Book and Ship's Record Folder, all raw data. The extracts collected were arrival times, berthing times, service start times, ship/berth idle times, end of service times, ship departure times, berth worked/used, ship length, cargo throughput and berth effectiveness. A total of one hundred and fifty-six vessels were observed.

For road traffic simulation, out-going traffic was studied because road use and conditions of roads outside the port can not be controlled or affected directly by decisions taken within the port either by the port operators or the shippers and their agents. Apart from this, points of

destination vary for different cargo and consignees.

The simulations were done manually and covered the followings: ship arrivals and berth allocations (noting the times); berth and quayside operations (including cargo discharge methods, volumes and times, noting delays); mode of conveyance of cargo out of port; arrival time and service time at exit points for road vehicles. Poisson distribution was used.

Databases were accessed to provide information on capacity utilisation of berths, storage facilities (sheds, warehouses and stacking areas), rail, road, pipeline and suction pump infrastructures.

4 DISCUSSION OF RESULTS

4.1 *Simulations*

Table 1 shows the results of the road traffic and ship traffic simulations. Road traffic simulation showed that an arrival rate of 0.66 vehicles per minutes was calculated. A service rate of 0.18 vehicles per minute was recorded. Each vehicle waited for at least 10.42 minutes to be served. A very high traffic intensity of 0.92 was recorded. The probability of no road vehicle leaving the port is 0.008675. Average time such a vehicle spent leaving the port from the point it joined the queue is 7 minutes. While the average time a road spends in the queue is just 1.062 minutes. A careful look at the total time spent leaving the port i.e. 7 minutes and time spent in the queue is just 1.062 minutes reflects the corrupt practices and laziness that exist at the gates as the vehicles are being served. More time is therefore wasted at the gates than on the queue. At any point in time, about 13 road vehicles are leaving the port i.e. at the gate area. While at least 10 road vehicles will always be on the queue leaving the port. The probability of queuing on arrival at the gates is 0.828627. This is rather high and suggests a considerable degree of congestion.

On the other hand, ship traffic simulation has arrival and service rates of 1.36 ships per day and 0.118 ships per day respectively. Arriving vessels wait on the average for 3.2 hours to be berthed. A traffic intensity of 0.48 was recorded meaning that vessels traffic is low. This contradicts the high level of cargo congestion at port. The probability of no vessel at the port is 0.0000167. Average time a vessel spent at the port (i.e. queuing and service time) is 7.353 days. Average time spent in queue (including time there is no queue) by ship is 0.04 hours. Average number of ships in the system (i.e. at the port at any point in time) is 11 ships. It means that at any time only 50% of the berth capacity is being used, as there are 22 berths at the port. Average number of vessels in the queue is 0.0048378, while the probability of not queuing on arrival at the port is 0.998.

From the simulation results, it seems obvious that the use of direct cargo discharge method unto waiting road traffic led to a longer time in port of 7.4 days. This was due to berth idle time that resulted from time loss while waiting for road vehicles to return for re-loading at the berths. This is well illustrated on figures 3 and 4 where berth-working time was compared with berth idle time and turn around time was shown on monthly basis.

Table 1: Calculations from Simulations

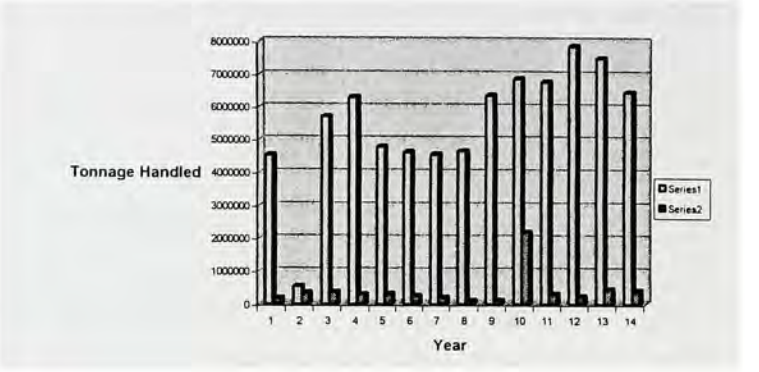
	Road Traffic	**Ship Traffic**
Arrival rate λ = No. of vehicle served / Time b/w arrival	19296 / 29232 = 0.66 veh/minutes	151/2663.7 = 0.057 ships/hour or 1.36 ships/day
Service rate μ = No. of vehicle served / Service time	19296 / 105856 = 0.18 veh/minutes	151 / 32650.7 = 4.625×10^{-3} ships/hour or 0.118 ships/day
Ave. waiting time = Total waiting time / No. of vehicles	201120 / 19296 = 10.42 minutes	484.2 / 151 = 3.2 hours
Traffic intensity ρ = Arrival rate / Service rate x C. Where c is no. of service points	C = 4; $\rho = 0.66/0.18 \times 4 = 0.92$ meaning traffic is very high	C = 22 $\rho = 0.57/(5.38 \times 10^{-23}) \times 22 = 0.48$ meaning traffic is relatively light

Probability of not queuing on arrival, P_0	$P_0 = 8.675 \times 10^{-1}$	$P_0 = 59739.2$
Average time spent in the system. $\frac{(\rho C)^2}{C!(1-\rho)^2 C\mu} \times P_0 + 1/\mu$	i.e. time spent leaving the port = 7 minutes	i.e. time spent at the LPC (queuing and service time) = 7.353 days
Average time spent in queue (including time there is no queue) $\frac{(\rho C)^2}{C!(1-\rho)^2 C\mu} \times P_0$	= 1.062 minutes	= 0.04 hours
Average number of lorries/ships in the system $\frac{\rho(\rho C)^c}{C!(1-\rho)^2} \times P_0 + \rho C$	= 13 vehicles	= 11 ships
Average no. of vehicles on queue $\frac{\rho(\rho C)^c}{C!(1-\rho)^2} \times P_0$	= 10 vehicles	$= 4.8378 \times 10^{-5}$ ships
Probability of queuing on arrival $= \frac{(\rho C)^2}{C!(1-\rho)^2} \times P_0$	= 0.828627	$= 2.42 \times 10^{-1}$
Probability of not queuing on arrival $= 1 - \frac{(\rho C)^2}{C!(1-\rho)^2} \times P_0$	= 0.17	= 0.998

4.2 *Methods of Cargo Discharge*

Figure 1 shows the methods of delivery of cargo used by the shipper. It showed that level of demand for storage facilities is very low. This mean that sheds, warehouses and stacking areas are not put to adequate use. About 24% of annual cargo traffic passed through the storage facilities.

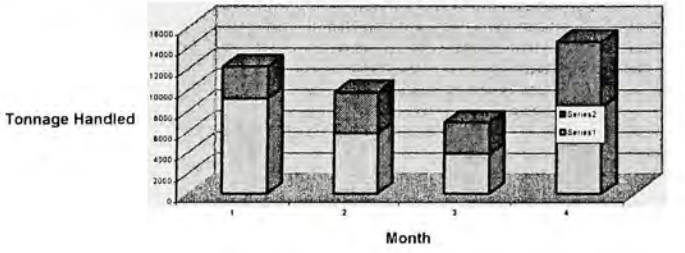
Figure 1: Cargo Delivery Methods



Series 1: Direct Delivery; Series 2: Indirect Delivery

Figure 2 derived from the ship traffic simulation equally showed that shippers preferred direct delivery during the period of ship traffic simulation.

Figure 2: Cargo Throughput Showing Cargo Delivery Method.

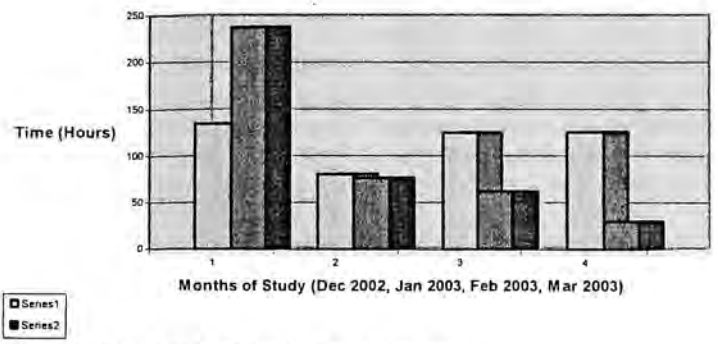


Series 1: Direct Delivery, Series 2: Indirect Delivery

4.3 ***Berth Working Time and Berth Idle Time for Ships at Port***

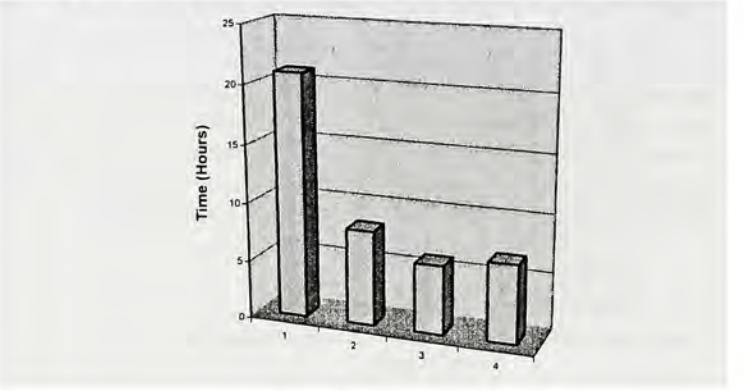
Figure 3 shows a comparison between berth working time and berth idle time. It showed a similar trend with the method of cargo delivery. December had most proportion of the cargo delivered by direct delivery and that same month has the highest ship berth idle time. The same month as shown by figure 4, had the highest turn-around time. It would appear that the higher the berth occupancy, the higher the turn-around time. However, high berth occupancy at the LPC seems deceptive as one-third of berth times are made up of idle time.

Fig 3: Comparing Berth Working Time & Berth Idle Time for Ships at port



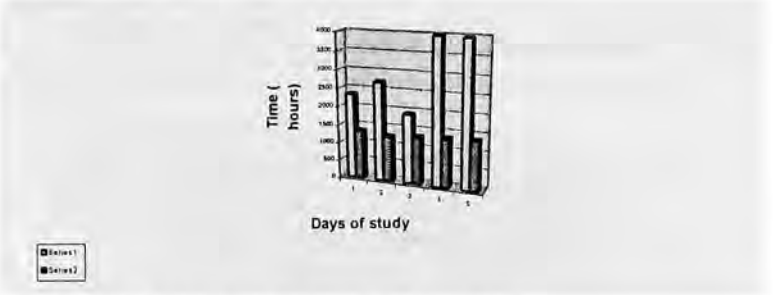
Series 1 Berth Working Time, Series 2 Berth Idle Time for Ships at port

Fig 4: Turn-round Time at for Ships at the port



A comparison was made between waiting/service times and inter-arrival times of road vehicles leaving the port. This was captured in figure 5. While study the result, it was obvious that the effect of direct delivery led to serious congestion at the exit points during December resulting in higher berth idle times throughout the month.

Fig 5: Comparing Waiting time & Service time for Road Vehicles leaving at the port between December 2002- March 2003



Series 1: waiting Time; Series 2: Service Time for road vehicles leaving the port.

Table 2 shows the distribution of cargo handled based on transport infrastructures used at the port. From the table it is clear that the rail mode is marginally used, indicating the relative unimpressive presence of rail transportation in Nigeria. This makes direct delivery method of cargo handling to pull so much demand on the road infrastructure. For this reason, direct delivery method has been considered rather attractive for shippers thereby putting enormous pressure on the road infrastructure.

Table 2: Capacity Utilisation Based on Tonnage Through Storage Facilities

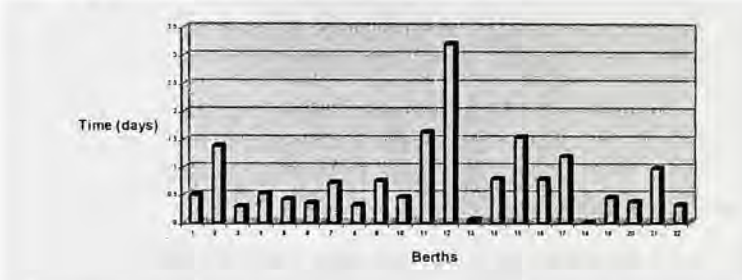
Year	Tonnage		
	Rail	Road	Pipeline
1990	128300	815912	3138164
1991	95600	1259	3652002
1992	--------	2043127	3981504
1993	63133	1011364	4049728
1994	--------	1303468	3819479
1995	64566	771346	3306711
1996	--------	506336	3369254
1997	8550	759219	2648380
1998	6259	713674	3496614
1999	5088	1003266	938780
2000	4566	987512	4792511

2001	7845	878452	6598745
2002	9546	645781	5437877
2003	6234	815461	6675498

Source: Nigerian Ports Authority

4.4 *Analysis of Berth Operations*

Fig 6: Inter-arrival time at port for calling vessels

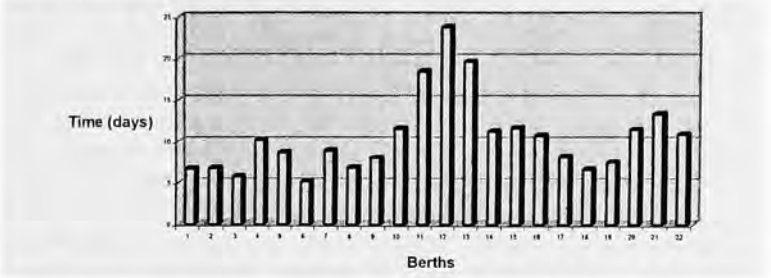


Figures 6 – 8 show how individual berth performed when the ship traffic simulation was conducted. Figure 6 shows inter-arrival times at LPC as vessels/shippers or their agents book for berths. Berth 18 has the lowest inter-arrival time while berth 12 has the highest (over 3 days) upon booking a berth vessels wait on a queue in the deep sea for over 3 days before approval to berth 12. Berths 14, 11 and 2 recorded the highest waiting times of over 9.6 hours, while berths 7, 9 and 21 recorded almost no waiting time before berthing.

Figures 7 and 8 showed berth 12 had the longest turn-around time and the worst berth effectiveness. Berth effectiveness was calculated by dividing the tonnage handled over at individual berths over the total time spent at such berths. Berths with impressive turn-around time had relatively good berth effectiveness scores. Berths 1-3 serves liquid dry bulk cargo berths where suction pumps are used for cargo handling. This makes the use of human labour limited and the time lost to bad/unfavourable weather is eliminated. Berths 4 –5, 12, 13, and 19-22 serves as dry bulk cargo berths. They use suction

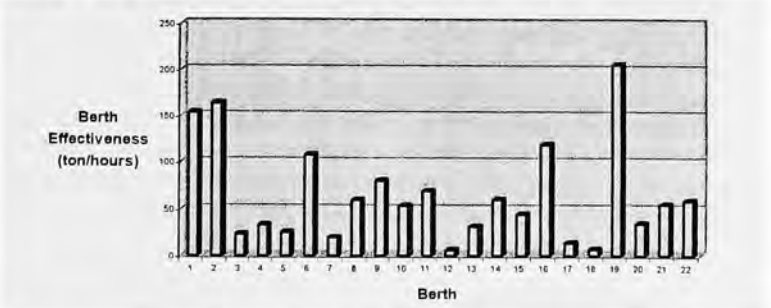
pumps and conveyor belts for cargo handling. Weather condition does not also affect these utility berths. However, berth 12, which records the worst turn-around time and berth effectiveness and longest inter-arrival time, is used for salt that are bagged and treated in the adjoining shed. Note that salt consumption is low when compared with other dry bulk (household commodities) cargo.

Fig 7: Turn-round Time at Port



Berths 20 posted relatively poor berth effectiveness because most times vessels that are berthed in berth 19 take are so long that they use up space in berth 20.

Fig 8: Berth Effectiveness (Throughput/service time)



Since the need for road vehicle use is not necessary at these dry and

liquid bulk cargo berths operation in those berths is a mixture of direct and indirect cargo handling process.

5 CONCLUSION

From the study, the use of direct cargo discharge from the vessels by virtually all the port users led to high vehicular traffic at the exit points which in turn resulted in longer berth idle time and time spent in port, poor capacity utilisation of storage facilities and other port infrastructure. High berth occupancy was recorded but this was quite deceptive as the over 30% of time spent in berths were idle berth times. Nigerian ports are known to have frequent congestion problems but from the study it was obvious that ship traffic to the port is low and berth capacity utilisation is about 50%. This could only mean that the port's cargo handling procedure and discharge method is somewhat inadequate, resulting in non-utilisation existing storage facilities, general poor port performance and the unwarranted congestion.

To solve this problem better integration of the transport modes with each other at the land end should be encouraged. Cargo storage should be encouraged at the port and direct cargo discharge should be discouraged. Means of encouragement should be devised. Should direct cargo discharge method still be favoured by port users, then efforts must be made to divert traffic off the road mode to other modes: rail and inland waterways. For diversion to be successful, the port users and consignees must see the rail and in-land waterways modes to be efficient, fast and widespread. Companies that have access to railway tracks in their factories should be mandated to use this mode for moving their consignments out of the port.

REFERENCES

Benathan E, and Walters A. A. (1979), "Port Pricing and Investment

Policy for Developing Countries" World Bank Research Publication, Oxford University Press.

Bird J. (1970), Seaports and Seaports Terminals, Dept. of Transport Geography, University of Southampton, England.

Briggs M. (1989), "Developing a Major Port Facility" *Developing World Transport,* vol. 1, pp 265-267.

Ikporuko (1977), Spartial; and Efficiency in a Physical Distribution System, Ph.D. Thesis, University of Ibadan.

Laing F. T. and Hecker (1989), "A Note on Optimal waiting-to-service Time Ratios in Large Ports" *Maritime Political Management Review,* vol. 16, No. 1 pp 85-86.

Levinson M. D., (1988),"On the Tolls Falls: A Model Network Financing" a Ph.D. Thesis, University of California, Los Angeles

Pieter De Langen (1998), "Time centrality in Transport", paper No. 1149, 8th World Conference on Transport, Brussels, Belgium.

Pyre C. J. (1989), "Maritime Infrastructure-An Overview", *Developing World Transport,* vol. 1, pp. 260-264.

UNCTAD/SDA (1982), Impacting Port Performance: Management of General Cargo Operations, Drake Educational Associates Ltd., Cardiff, Wales.

Wolfhard A. (1989), "Port Problems in Developing Countries" *Developing World Transport,* vol. 1 pp 268-270.

Impacts of road improvements on generated traffic, congestion and emissions: Case of the Hamilton CMA, Canada

Hejun Kang, Darren M. Scott, Pavlos S. Kanaroglou, Hanna Maoh
School of Geography and Geology, McMaster University, 1280 Main Street West, Hamilton, Ontario L8S 4K1

Abstract

Many recent studies have provided persuasive evidence for generated traffic. However, there remain many problems in such research: the unclear separation of induced traffic from diverted traffic, failing to incorporate some induced traffic factors, or not controlling for confounding influences such as socio-economic factors. This study clarifies and addresses these issues by a case study of two new highway projects in the Hamilton CMA, whose construction has been debated for over half a century. Our research used IMULATE, an integrated land use and transportation to evaluate scenarios: a base scenario and two highway scenarios, over a 20-year time horizon. Specifically, there are three main objectives in our research: 1) to gauge induced travel and diverted travel separately, due to new road infrastructure proposals; 2) to examine the impact of new highways on parallel links or upper/lower (feeder) links flows; and 3) to examine the impact on traffic congestion, and surrounding environment. Results show that the estimated number of total trips increases, which provide evidence of induced travel. At the same time, the hypothesis of diverted travel is supported, including route switching and mode switching. The results also showed that, at the metropolitan level, the estimates of the congestion levels and total emissions decrease after new highway construction. However, the environment of areas surrounding these highways is heavily affected.

1. Introduction

Few issues in the urban transportation field have threatened proposed road projects and sparked as much controversy as claims of induced traffic (Cervero, 2003). Highway critics have argued that it is a futile exercise to build new roads or expand existing ones to relieve traffic congestion—these improved roads simply spur additional travel (induced travel) or diverted travel, both of which are collectively known as generated traffic (Litman, 2004). Such behavioral responses to highway capacity expansions are categorized by Hills (1996) and by Litman (2004) into two classes: short-term and long-term impacts.

Over the short run, increased capacity leads to the decrease of traveling time, which is followed immediately by three behavioral responses. Downs (1992) calls them "triple convergence": a) route shift from unimproved roads to the newly improved roads; b) mode shift from transit to auto driving; and c) temporal shift from off-peak periods to the peak. All of the above responses are defined as diverted traffic by Litman (2004). Additionally, formerly suppressed trips are now made, which can be to farther destinations or new trips, under the same travel-time budgets due to increased accessibility. Both of these new and longer trips are defined as induced traffic or latent demand (DeCorla-Souza and Cohen, 1999). Over the long run, people and firms locate close to new highways, to exploit the accessibility benefits created (Cervero, 2003).

In this paper we seek to identify and differentiate the impacts of road improvement projects with the help of a large-scale simulation models, known as Integrated Transportation and Land-Use Urban Models (IMULATE). Our objectives are: to gauge induced travel and diverted travel separately, due to new road infrastructure proposals; to examine impacts of new highways on parallel links or upper/lower (feeder) links flows; and to examine impacts on traffic congestion, and surrounding environment.

The reminder of this paper is organized as follows. In the next section we summarize empirical findings from earlier studies, and review some research challenges. This is followed by our research objectives and context. A section describing the methods is then

presented, followed by the discussion of results. A concluding section summarizing the results is also included.

2. Literature Review

2.1 Previous research findings

Beginning in the late 1990s, studies have provided persuasive evidence for induced travel, and the principle has been acknowledged by leading transportation researchers or agencies (Transportation Research Board, 1995; Transportation Research Circular, 1998) and by the U.S. Environmental Protection Agency (EPA, 2000). Most of past works measure induced travel as the elasticity with respect to lane kilometers (VKT) or travel time (VMT--vehicle minutes traveled). Results find that the majority of estimates are in excess of 0.5 (VKT) (Hansen and Huang, 1997; Fulton et al. 2000; Marshall, 2000; Noland and Coward, 2000), or about –0.5 and up to –1.0 (VMT), suggesting that there is a significant relationship between new road capacity and increases in traffic (Scott, 2002). In other words, the most improved road capacity is "absorbed" by induced traffic.

Although there is no doubt about the existence of induced traffic, its estimates (e.g., elasticity of VKT or VMT) have been found to be within a wide range. Several factors impose effects on such difference. First one is the time span calculated. Long-run elasticities are considerably higher than short-run ones, which have been proved by many studies. Noland (2001) finds elasticity values ranging from 0.3 to 0.6 in the short run but from 0.7 to 1.0 in the long run. An analysis of nationwide metropolitan level data by Noland and Cowart (2000) tells the same story.

Another important factor of VKT elasticity is the level or resolution at which researches have been conducted (facility level, county level or metropolitan level) (Cervero, 2002). The induced-demand effect generally becomes larger as the unit of analysis increases in size, especially when diverted trips are netted out of calculations. The effect of resolution is represented by the work of Hansen et al. (1993): at the road segment level, estimated elasticity of VKT is in

the range of 0.15 to 0.30; at the county level and metropolitan level, elasticity is higher, in the range of 0.32 to 0.50 and 0.54-0.61 respectively. However, Cervero and Hansen (2002) further note that, to gauge newly generated traffic, or induced demand, researches should focus on changes in VKT at a county or metropolitan level versus the road segment level.

Cervero (2002) also identify that VKT elasticity estimates may change with types of road improvements. All things being equal, a brand-new highway can be expected to come up with more new traffic than the expansion of an existing facility. Such contention is supported by the findings of Ruiter et al. (1980) who estimate substantially lower elasticities of VKT with respect to road capacity for the expansion of older segments of the same freeway, in comparison to the extension of a freeway linking Oakland, California, to its eastern suburbs.

2.2 Challenges in practice

Given the overwhelming evidence, there is no question about the existence of induced travel and diverted travel. Therefore research has shifted to another important issue: whether the policy makers or decision-making tools could reflect the generated travel effects while planning. Cervero (2002) reports that in many metropolitan areas of North America, traditional four-step travel-forecasting models or urban transportation modeling systems (UTMS) fail to embody "induced travel" effects (Transportation Research Board, 1995). Specifically, most forecasting models do not have an accessibility component, allowing them to forecast land-use changes due to accessibility improvements (Scott, 2002). Therefore, forecasters run the risk of understating air pollutant emissions, but overstating estimates of speed and travel time savings of future road proposals (DeCorla-Souza and Cohen, 1999).

Another challenge of testing the induced travel hypothesis is controlling background factors like demographic and socio-economic changes (e.g., population growth, decreased household size, and increased female work force, income increase as well as the

decentralization of metropolitan areas). One direct outcome is that estimates of induced traffic are overstated.

Furthermore, it is hard to separate induced travel from diverted travel using UTMS. The challenge is coming from the discrepancy between definitions and measurements of both types of travel. On one hand, it is commonly believed that induced travel should include only new and longer trips; on the other hand, most researchers measure induced travel to be an increase in VKT (Fulton et al. 2000; Hansen and Huang 1997; Noland and Cowart 2000), which could include both new and diverted trips (Scott, 2002).

3. Research Context and Objectives

This paper is devoted to solving the three challenges discussed in section 2, by using the case study of Hamilton CMA, Canada. Before we go further, it is necessary to introduce briefly about our research context. The Red Hill Creek Highway project was first proposed in 1956 and had been debated for several decades. Until 1987, this project was passed and would proceed with two sections of an expressway. The first section, the Lincoln Alexander Parkway (Linc: 11km), was open in 1997 (Friends of Red Hill Valley, 2004). The second section is called the Red Hill Creek Expressway (RHE: 8 km) (Peace, 1998). Upon completion, these two sections will connect the lower city with Queen Elizabeth Way and extending east to west across Hamilton Mountain (Peace, 1998), as shown in Fig. 1.

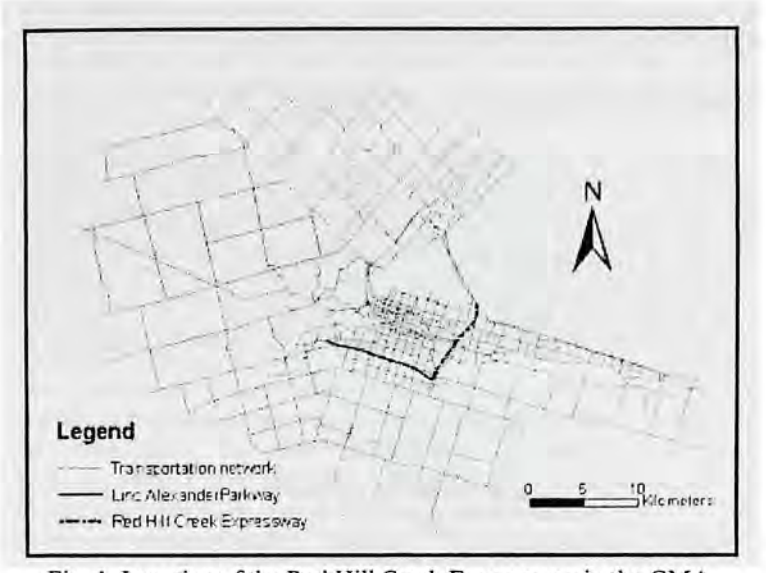


Fig. 1. Location of the Red Hill Creek Expressway in the CMA transportation network

Arguments in support of the expressway have been largely based on its anticipated boost to the regional economy. Opposition to the expressway is mainly based on environmental grounds. Opponents claim that "...to build this highway, approximately 25% of the valley will be cleared for the expressway right-of-way...Therefore, a migration corridor for at least 177 species of birds and 25 mammal species will disappear" (Friends of Red Hill Valley, 2004). More importantly, east Hamilton's air quality will be further degraded along with people's health (Friends of Red Hill Valley, 2004).

With a general goal of solving the above long lasting disputes, the objectives of our research are: 1) to gauge induced travel and diverted travel separately, due to new road infrastructure proposals; 2) to examine the impact of new highways on parallel links or upper/lower (feeder) links flows; and 3) to examine if the road improvements are helpful to ease congestion, and detrimental to surrounding

environment close to these new links. To realize such objectives, we devised three scenarios that span the time periods from 1991 to 2006.

4. Methods and Scenarios

IMULATE is an integrated land-use and transportation model calibrated for the Hamilton CMA, Canada. This model is calibrated for the morning peak traffic period by using data from the 1986 Census of Canada and the 1986 Transportation Tomorrow Survey (Anderson et al., 1994). The basic structure of IMULATE shown in Fig. 2 has four main components (Anderson et al., 1994). A comprehensive discussion can be found in the works of Anderson et al. (1994).

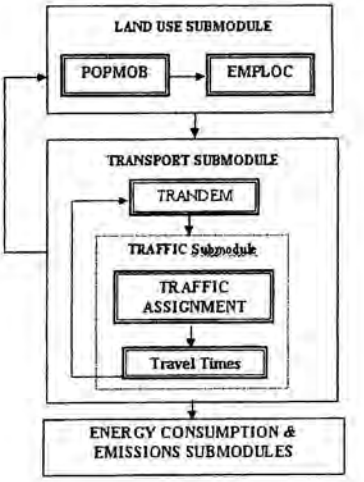


Fig. 2. General structure of IMULATE

The first component is the Land Use submodule, by which both population and workers are assigned to zones by means of a residential choice model system (Kanaroglou, 1999). The second is the Travel Demand submodule. It predicts the number of work,

school, and discretionary trips occurring between each zone pair. Then, trips are broken down into three modes-motorized, public transit, and passenger (Kanaroglou, 1999). The third one is the Network Assignment submodule. In this module, once network equilibrium has been reached, the calculated "inclusive value" is passed to the POPMOB and EMPLOC submodules as a determinant of population and employment mobility. By this way, a bi-directional relationship between the land-use pattern and the transportation system is represented in IMULATE (Anderson et al., 1994). The fourth is the Emissions submodule. It estimates emissions of carbon monoxide (CO), hydrocarbons (HC), and nitrogen oxides (NOx) on each link (Kanaroglou, 1999).

In order to achieve the objectives stated above, three main scenarios are tested spanning four periods: 1986-1991, 1991-1996, 1996-2001, and 2001-2006. The base scenario, which excludes the Linc and RHE, serves as a base for comparison to see the effects of changes in background factors (e.g. population growth, economic activity, and behavioral and/or life-style changes). The second scenario includes the Linc. The third scenario, assuming the existence of the Linc, allows for the implementation of the RHE. The latter two scenarios are called expressway scenarios. Through deductions between different scenarios, it is easy to remove some of the increased traffic due to confounding factors common to both scenarios.

However, IMULATE has some limitations. First, like most other UTMS, this model can only capture trips occurring in the morning peak hour period. Increase in capacity may likely lead to increases in discretionary trips that normally take place in the off-peak period. Therefore, some diverted trips (e.g. temporal shift) may not be captured by this model. The second limitation is related with the prediction span period. In IMULATE, predictions can only be made from 1991 to 2006, which will not be long enough to capture all generated traffic associated with road improvements. This is due to the fact that land-use pattern changes always take a long time to occur.

5. Results and Discussions

5.1 Induced travel and diverted travel (mode shift)

Main trends over the entire study area in both the base and other scenarios runs are summarized in Table 1. Absolute values and changes in key variables between 1991 and 2006 for the base scenario, and between scenarios in 2006 are presented. Here only automobile trips are counted.

Table 1
Estimates for key variables in the morning peak hour period, from simulation scenarios (1991 to 2006).

Scenario	Total Traffic Flows	Vehicle Kilometers Traveled(km)	Average Link Speed (km/hr)	Emissions (kg) HC	CO	NO_x
1991Base	3675457	2402343	53	15449	183844	12285
	5419039	3663491	45	36633	426081	18142
2006 Base	(1743582)**	(1261148)**	(-8)**	(21183)**	(242236)**	(5857)**
2006 Base	5325738	3645990	46	32515	377731	17668
+Linc	(-93301)*	(-17501)*	(1)*	(-4118)*	(-48350)*	(-475)*
2006 Base	5195355	3606574	47	29411	342149	17215
+Linc +RHE	(-223684)*	(-56917)*	(2)*	(-7222)*	(-83931)*	(-927)*

** Numbers are the changes from 1991 to 2006 for the base scenario
*Numbers are the changes from the base scenario to the new scenarios for 2006

An increase in VKT and total traffic flows is observed for the base scenario from 1991 to 2006. This is resulted from both the increasing average length of trips and the increasing number of trips due to more people and more cars. The reduction in average speed reflects more road network congestions, which cause higher emissions of the three pollutants (HC, CO and NOx). After adding the Lincoln Alexander Parkway (see the 2006 Base+Linc), however, both the total traffic

flows and VKT decrease. At the same time, the average link speed increases and thus the overall pollution emissions are reduced. The implementation of the Red Hill Creek Expressway leads to further decreases in traffic flow counts, VKT, and emission rates, but increases in average link speed.

Based on the estimates shown above, we may conclude that no induced travel is caused by road improvements. However, the number of trips estimated by IMULATE tells a converse story. The total number of trips (discretionary, school and work) in 2006 increases from the base scenario (196119) to the other two highway scenarios (196580[1] and 197232[2]), which means that there are new trips stimulated by road improvements.

To distinguish induced travel from diverted travel, the total number of trips by mode is shown in Table 2. It is observed that with the implementation of new highway links, the number of motorized trips and passenger trips increase, while the transit trips decrease.

Table 2
School and work trips by mode for morning rush hour period, estimated from simulation scenarios (1991 to 2006).

School and Work	Motorized	Passenger	Transit	Total
1991Base	116877	13487	7189	137553
	154048	17729	8359	180136
2006 Base	(37171)**	(4242)**	(1170)**	(42583)**
	154193	17751	8261	180205
2006 Base+Linc	(145)*	(22)*	(-98)*	(69)*
	154444	17784	8160	180385
2006 Base+Linc+RHE	(396)*	(55)*	(-199)*	(249)*

** Numbers are the changes from 1991 to 2006 (Base)
*Numbers are the changes from the base scenario to the new scenarios for 2006

[1]-the 2006 Base+Linc scenario
[2]-the2006 base+Linc+RHE scenario

The VKT results in Table 2, which prove there is no induced demand, are apparently at odd with those aggregate studies summarized in section 2. It shows that VKT is probably not a good proxy of induced travel. The reason for this is that road improvements have both a negative and a positive effect on VKT (Kanaroglou, 1999). The negative effect arises from shortening interzonal distances by making the network denser. The positive effect arises because the improved accessibility afforded by the new infrastructure stimulates more discretionary trip making (i.e. up to 1045). Since this simulation is conducted for the morning peak when there is relatively little discretionary trip making, the negative effect dominates the positive. Furthermore, an increase in VKT may be composed of both new and diverted trips (Scott, 2002), which is also not an accurate index in this sense.

5.2 Diverted travel (route shift)

In order to identify the diverted trips switching to the two new highway links (route shift), the output of the traffic flow counts on each transportation link is used. The spatial distribution of the traffic flow changes from the base scenario to the highway scenario (2006 Base+Linc+RHE) is shown in Fig. 3. This map indicates that the flows decrease on those parallel links of the Linc Alexander Parkway and Red Hill Creek Expressway, but increase on their upper or lower links. That means improvements on one link inevitably have impacts on travel demands on other links that feed into it or parallel to it. The results are supported by Cervero (2002), who notes that road facilities are interdependent in a network.

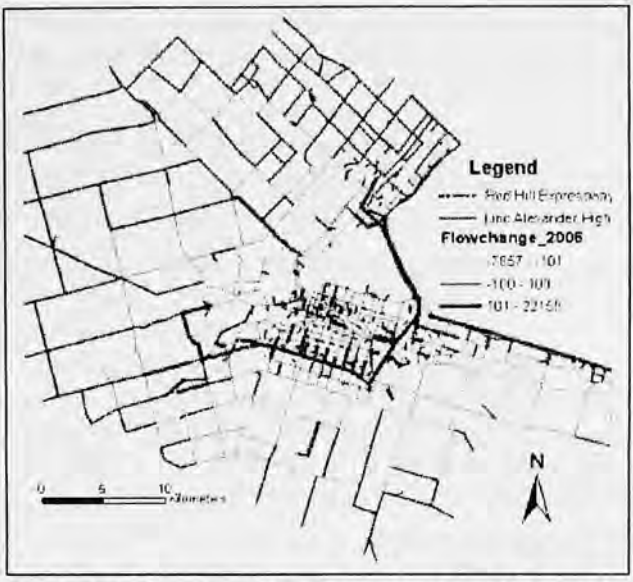


Fig. 3. The link flow change from base scenario to the base+Linc+ RHE scenario (2006)

5.3 Congestion and environmental issues

As shown in Table 1, new transportation infrastructure projects are helpful to relieve congestion and emissions level to the whole transportation network. However, there is appreciable increment of traffic flow counts and pollutant emissions on the new links over time (See Table 3). From 1991 to 2006, the total traffic flow counts increase by 50%, and the growth of HC and CO is more than two times. That means the new highway construction will have a negative impact on the surrounding environment.

Table 3
Changes of the total traffic flows and emissions on the Linc Alexander highway and Red Hill Creek Expressway from 1991 to 2006.

	1991	1996	2001	2006
Total Traffic Flows	277540	353585 (27%)*	383293 (38%)*	407219 (47%)*
HC	692376	1465844 (112%)*	1965491 (184%)*	2505884 (262%)*
CO	7592233	15181529 (100%)*	20555073 (171%)*	26577878 (250%)*
NOx	848501	990313 (17%)*	1088022 (28%)*	1178359 (39%)*

*Numbers are changes from 1991 to the following years (1996/2001/2006)

6. Conclusions

An Integrated Transportation and Land-Use Model is used to demonstrate the generated traffic effects as a result of road improvements. Scenario deduction algorithms are used to control confounding factors. Different from other UTMS, IMULATE is designed to capture land use changes associated with road improvements through the "inclusive value". Results show that the estimated number of total trips increase, which provides evidence of induced travel; at the same time, route switching and mode switching are observed. However, VKT, a popular proxy for induced travel, is proved to be an inappropriate index.

Due to the increase in road capacity, estimates of the congestion level and total emissions decrease from 1991 to 2006. However, on both of the new links, Linc and RHE, emissions and traffic flow counts increase significantly over time. It shows that the surrounding environment of the Red Hill Creek Valley will be degraded by the new highway construction.

References

Anderson, W. P., Kanaroglou, P. S., Miller, E. J., 1994. Integrated land use and transportation model for energy and environmental analysis: a report on design and implementation. McMaster Institute for Energy Studies, Hamilton, Ontario, Canada.

Cervero R., 2002. Induced travel demand: research design, empirical evidence, and normative policies. Journal of Planning Literature 17 (1), 3-20.

—, 2003. Road expansion, urban growth, and induced travel: a path analysis. Journal of the American Planning Association 69 (2), 145-163.

Cervero R., Hansen M., 2002. Induced travel demand and Induced road investment: a simultaneous equation analysis. Journal of Transport Economics and Policy 36 (3), 469-490.

Downs, A., 1992. Stuck in traffic: Coping with peak-hour traffic congestion. Washington, D.C.: The Brookings Institute.

DeCorla-Souza, P., Cohen, H., 1999. Estimating induced travel for evaluation of metropolitan highway expansion. Transportation 26, 249-262.

Friends of Red Hill Valley. Retrieved Dec 20, 2004, from http://www.hwcn.org/link/forhv/expressway/projecthistory.htm#red

Fulton L. M., Noland R. B., Meszler, D. J., Thomas J. V., 2000. A statistical analysis of induced travel effects in the U.S. Mid-Atlantic region. Journal of Transportation and Statistics 3 (1), 1-14.

Hansen, M., Huang, Y., 1997. Road supply and traffic in California urban areas. Transportation Research Part A: Policy and Practice 31 (3), 205-218.

—, 2004. Generated traffic and induced travel: implications for transport planning. Victoria Transport Policy Institute.

Kanaroglou P. S., 1999. New transportation links and urban air quality: the Red Hill Creek expressway, Hamilton, ON. Paper

presented at the 6th Int. Conf. On Environmental Science and technology, Pythagorion, Samos, Greece.

Marshall, N., 2000. Evidence of induced demand in the Texas Transportation Institute's urban roadway congestion study data set. Paper presented at the annual meeting of the Transportation Research Board, January 9-13, Washington, DC.

Noland, R. B., 2001. Relationships between highway capacity and induced vehicle travel. Transportation Research Part A: Policy and Practice 35, 47-72.

Noland, R. B., Cowart, W. A., 2000. Analysis of metropolitan highway capacity and the growth in vehicle miles of travel. Transportation 27, 363-390.

Peace, W. G., Editor, 1998. From mountain to lake: The Red Hill Creek valley, Hamilton, Ontario: W.L. Griffin Printing.

Ruiter, E., Loudon, W., Kern, C., Bell, D., Rothenberg, M., Austin, T., 1980. The vehicle-miles of travel-urban highway supply relationship. In NCHRP Research Results Digest 127. Washington, DC: Transportation Research Board, National Cooperative Highway Research Program.

Scott, D., 2002. Overcoming traffic congestion. A discussion of reduction strategies and behavioral responses from a North-American rerspective. In: Perrels, A., Himanen, V. & Lee-Gosselin, M. (eds.), 2002. Identifying building blocks of sustainable transport – lessons on external effects from both side of the Atlantic. A special issue of European Journal of Transport and Infrastructure Research 2, No ¾, 317-338.

Transportation Research Board, 1995. Expanding metropolitan highways: implications for air quality and energy use, special report 245, National Research Council, National Academy Press, Washington, D.C.

Transportation Research Circular, 1998. Highway capacity expansion and induced travel: evidence and implications. Transportation Research Board, National Research Council, 481.

KEEPING PEEL'S ECONOMY ON THE MOVE
Advancing the Agenda for Goods Movement in Peel

Introduction

Regardless of where it occurs in the world, the process by which people and businesses acquire new physical property or make business transactions involves goods movement and supply chain logistics. While competitive pressures are constantly challenging the freight industry to be more efficient and reduce costs, new technology in freight systems will only advance us part of the way. Canada's large urban regions, including the Region of Peel, have reached a critical point. To advance any further, they must first deal with rapid growth, congested highways and overloaded infrastructure.

The GTA and Peel represent a major goods movement focal point for Canada. The future competitiveness of the GTA area depends on facing up to new freight challenges. The issues are complex, and the need to make the most of opportunities has never been greater.

Other jurisdictions in Canada and around the world are organizing freight councils to bring together practitioners and policy makers in a forum that can generate creative solutions to freight transportation problems. Goods movement, being largely under the purview of the private sector, has generally not been well understood by municipal planners and politicians. Freight policy development and planning have often been weak at the municipal level.

We are trying to change this situation in the Region of Peel. This paper reviews recent work on goods movement in the Region. It discusses 1) the significance of Peel's freight transportation at a local, regional, provincial and national level, 2) the importance of goods movement to the Peel economy, 3) the results of consultations with industry and 4) the Peel study findings and recommendations.

A copy of the "Study of Goods Movement in Peel, Strategic Overview" is posted on the Regional website at the following location: www.peelregion.ca/planning/pubs/transportation.htm.

Murray McLeod

It is hoped that this communication through CTRF will be of value to other municipalities that similarly feel a need to do more comprehensive work in goods movement planning, in order to meet current and future needs. We would like to trigger more dialogue on the subject across all government levels and with the private sector. This study is one of the first municipal goods movement studies of its kind in the province.

Background on the Peel Goods Movement Study

The Peel goods movement study is one of five studies making up the Peel Long Range Transportation Plan. It is being conducted as part of a strategic update to Peel's Regional Official Plan. The existing Regional Official Plan contains relatively little reference to goods movement in relation to its overall importance. The Goods Movement Study has recommended the development of new policies and directions for goods movement. The study had three objectives at the outset: 1) to develop a better understanding of Peel goods movement issues and challenges in consultation with the industry and stakeholders; 2) to develop policies and next steps for the approval of Regional Council; and, 3) to build a foundation of knowledge and partnerships to help implement the policies.

Key Messages from the Peel Study

A number of key messages have come out of the study; 1) Peel Region is a "supernode" for goods movement in Ontario and Canada; 2) Efficient and reliable goods movement is critical to Peel's economic vitality; 3) Peel's freight networks are operating quite well at the present time, but serious threats, such as growing highway and road congestion are looming on the horizon; 4) Better data/information about all of the freight modes in Peel needs to be collected, and 5) Coordination and long term planning are necessary to improve Peel's freight system both now and in the future. Goods

movement is largely a private sector activity with many of the major transportation decisions being taken by the private sector. Strong public private partnerships and alliances will be essential for future success.

Peel is a Major Hub for Freight Transportation

Peel is strategically located in the Golden Horseshoe as shown in **Figure 1**, and is supported by a strong multimodal transportation system of highways, roads, railways, airports, pipelines and marine ports. It is connected to two major NAFTA corridors. Highway 401 operates as the Main Street of Southern Ontario and connects Peel to the US border at Windsor and Detroit, and to the Montreal area in the other direction. The Queen Elizabeth Way connects Peel to the US border at Buffalo and Niagara Falls. Both of these corridors converge within the Region of Peel.

Figure 1:



The expressway network has no less than five 400-series highways: the Queen Elizabeth Way; Hwy 401; Hwy 403; Hwy 407 and Hwy 410. (**Figure 2**) These expressways come together to give Peel the densest freeway grid within the GTA's overall freeway network.

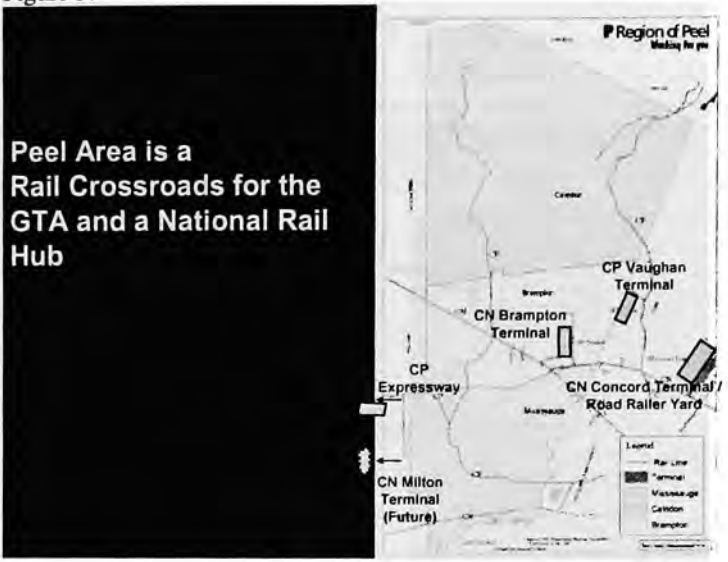
Figure 2:



Peel is located at the crossroads of major railway lines operated by Canadian National, Canadian Pacific and the Orangeville and Brampton short line. (**Figure 3**) There are three major rail intermodal terminals located in Peel or in the immediate vicinity. The CN Brampton Intermodal Terminal is in Bramalea. Just a little to the south of Bolton is the CP Vaughan Intermodal Terminal. Terminals for both the CN and CP piggyback trailer services in the Windsor-Montreal corridor are located nearby. Located just west of Meadowvale in Mississauga is the CP Expressway terminal. Further east is the CN Road Railer terminal in Concord. All of the above terminals, taken together, operate like a rail hub for the eastern part of Canada.

Murray McLeod

Figure 3:



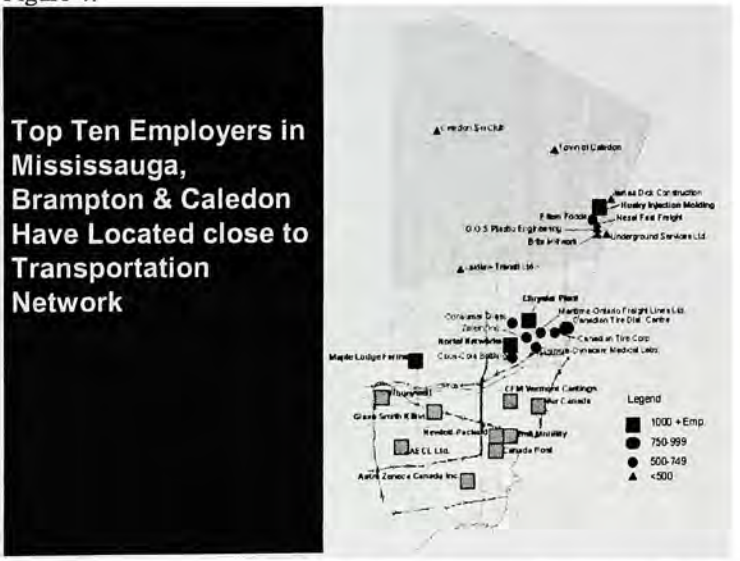
Pearson Airport is a major air gateway for Canada and handles more air cargo than the Vancouver and Montreal airports combined (the next two largest Canadian airports for air cargo). A major factor in Pearson's success as an air cargo hub is the number of direct flights from the Pearson Airport to so many locations around the world.

Finally there are two pipelines crossing Peel and three marine terminals located along the shoreline of Lake Ontario.

Goods Movement is Vital to Peel Region's Economy, Industry and the Competitiveness of its Businesses

Peel's industries have chosen to locate close to the transportation trunk routes and terminals. Commercial and industrial land uses have congregated close to the transportation network. As becomes apparent from a land use map for Peel. As seen in **Figure 4,**

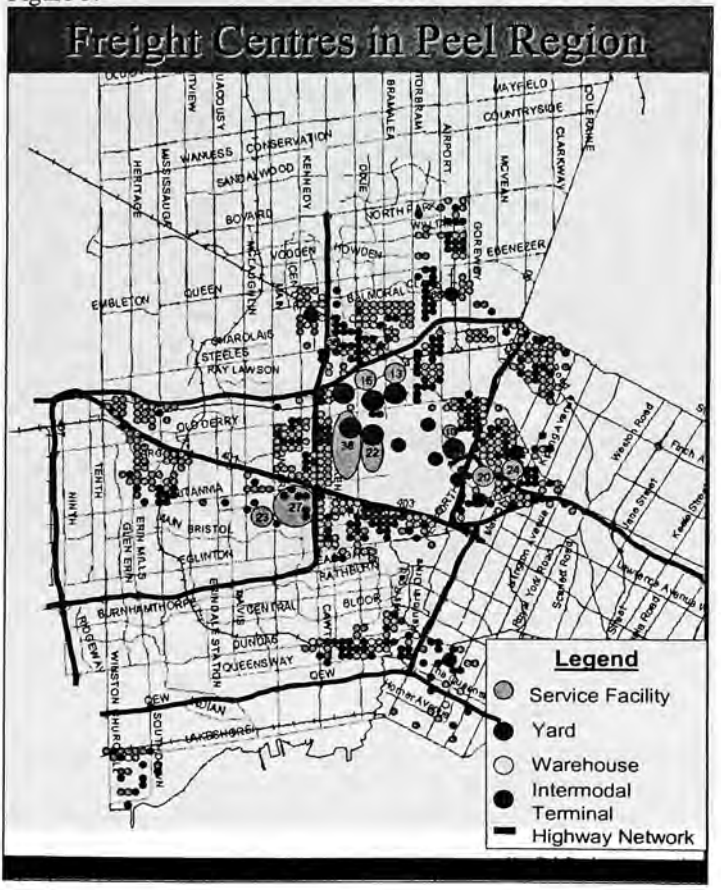
Figure 4:



The top 10 employers, in each of Mississauga, Brampton and Caledon, are located very close to the transportation network, because of their need for good transportation.

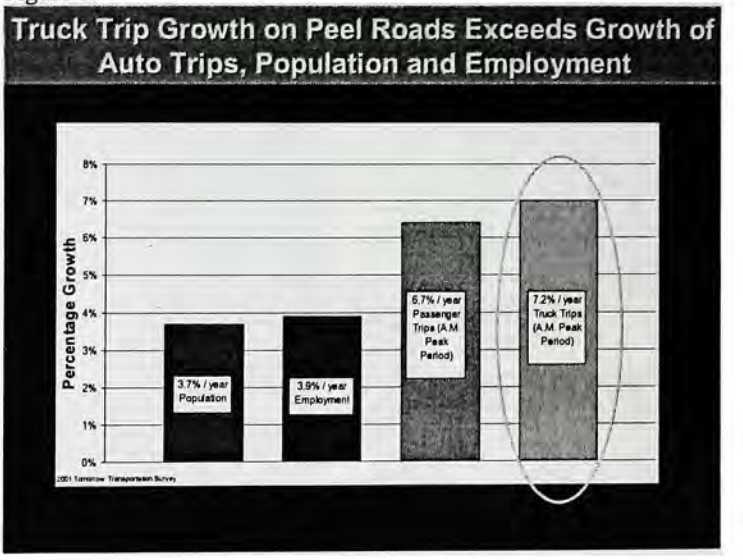
The various types of freight-handling centres, terminals, yards and warehouses are also located close to major transportation facilities as illustrated in **Figure 5**.

Figure 5:



Trucking in Peel Region has been increasing at almost twice the growth rate of population and employment as shown in **Figure 6**. Truck and passenger trips by auto are growing at 6 to 7% a year while population and employment in Peel are growing at just under 4% a year.

Figure 6:



Goods movement is an important employer for Peel as seen in **Figure 7**. Transportation and warehousing together represent about 9% of Peel's total employment. Typically for most other areas of Canada,

and in particular for other metropolitan areas, transportation and warehousing would only represent 4 to 5% of total employment.

Figure 7:



Peel is a Strong Contributor to Ontario's Trade, and Goods Movement Activity Attests to the Success of Ontario's Trade

Peel has a strong economic base, helped by a manufacturing-based economy which produces autos, car parts, durable goods, electronics and food items as some of the major products.

The Peel area accounts for 37% of the Greater Toronto Area's industrial square footage. Peel alone produces about 15% of Ontario's exports.

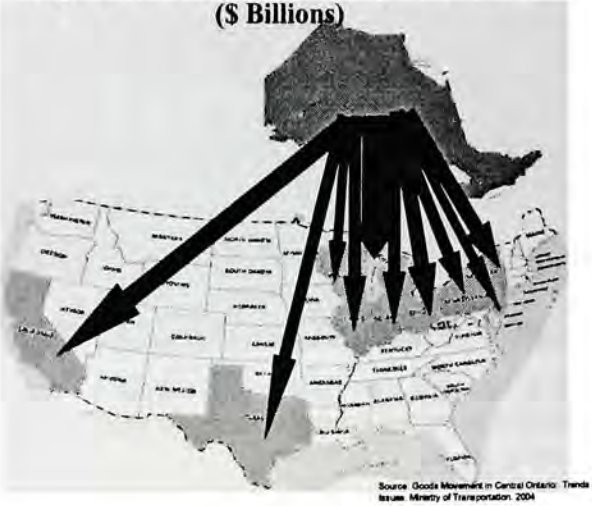
Murray McLeod

Canada's prosperity is heavily dependent on exports. Approximately a third of Canada's GDP is generated by exports. Daily trade between Canada and the United States amounts to $1.6 billion a day and over $1 billion of these goods are moved by truck.

Approximately 84% of Ontario's trade is with the US and this is worth about $360 billion a year. About 94% of Ontario's exports go to the US and the flows of exports to the top 10 states are shown in **Figure 8**

Figure 8:

Top 10 States for Ontario Exports in 2001
($ Billions)



Source: Goods Movement in Central Ontario: Trends and Issues. Ministry of Transportation. 2004

If you look at a map of North America it can be seen that Peel is well-located with respect to US markets. **Figure 9** shows that Chicago, New York, Boston and Washington are generally within a days drive by truck, and Atlanta is within two days drive. Economic data shows that approximately half of the US population and over half of the US economic market is located within a two day drive of Peel Region.

Figure 9:



Peel's multi-modal freight system keeps Peel's businesses competitive. It provides excellent access to business opportunities world wide. It enables Peel to attract high quality businesses and jobs.

A Major Concern Expressed by Peel Industry Representatives was Growing Congestion

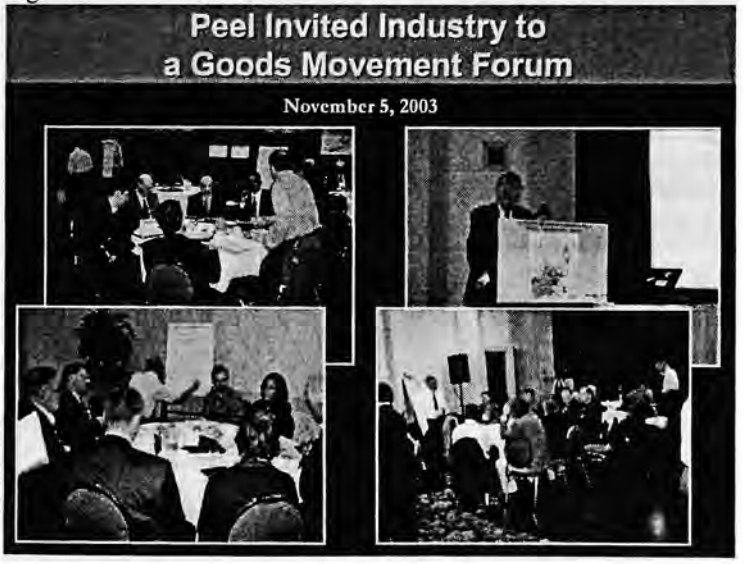
We conducted a telephone survey of industries and transportation companies in Peel. Industry generally gives Peel a good report card because Peel has a good road network, good choice of modes and is close to key markets in the US and Canada. But there are a number of threats on the horizon. Road congestion is growing, especially on highways and major roads. Industry is running out of good locations in Peel and is beginning to spread out. Peel has not been doing systematic freight corridor planning and finally, the collection of additional freight planning data is needed. Even the freight industry acknowledges that data availability is a problem. In Peel we need to focus on understanding our key freight corridors. Congestion is diverting traffic to less suitable routes. The trucking industry has been vocal about the highway 407 tolls. Highway 407 has not been taking its share of truck traffic and there is spillage onto parallel streets such as Steeles and Queen Street. A greater focus on the key freight corridors in Peel can help the transportation system work better.

We have been learning from what other major metropolitan areas are doing in goods movement planning and partnerships with industry. Montreal, for example, has an inter-regional committee for goods movement (CITM) with 70 representatives from the public and private sectors. Orlando Florida has been developing a goods and services strategy plan for helping tourism in the Walt Disney area. Seattle has a freight mobility roundtable that involves strong partnership with the private sector and Philadelphia has a freight planning program that involves an advisory committee and task force.

Industry Was Involved through the Peel Goods Movement Forum, November 2003

A Peel Goods Movement Forum was organized in November of 2003 and leaders from industry, freight companies, associations and all four levels of government were invited. **Figure 10** captures some of the breakout sessions at the Forum. There were over 50 participants. The discussion was enthusiastic and attendees delved into current issues and offered suggestions on how to improve Peel's goods movement system both now and in the future. A full account of the Forum is available in Appendix C of the Peel Goods Movement report which is posted on the Peel Region website: www.peelregion.ca/planning/pubs/transportation.htm.

Figure 10:



The Peel stakeholders told us that new partnerships are needed to improve goods movement in the Peel area. They should be public

private partnerships with strong inter-governmental relationships, and they should provide leadership and seek out funding for needed infrastructure projects.

There was a strong feeling that goods movement facilities, such as the provincial highways, are under-funded and leadership is needed to make the case for the provision of additional transportation capacity. Industry acknowledged that goods movement data needs to be collected to map current goods movement flows to show how the modes are operating and where improvements could be made. They suggested looking for early action improvements. Critical freight corridors should be identified in Peel, corridors that need to work in balance with other corridors such as transit corridors. Land use planning was identified as a key measure to ensure the compatibility of freight activities with neighbouring land uses in the community.

The Goods Movement Study Made Recommendations for Peel in Three Policy Areas

The recommendations identified in the study are based on stakeholder feedback on issues, options and best practices.

The recommendations cover three major policy areas. 1) The building of a new stakeholder partnership to help with Peel's transportation planning process. Peel Regional and Area Municipal staff will continue to work with the other levels of government and the private sector in a new partnership (the Peel goods movement liaison team) to develop an ongoing goods movement agenda and strategy for improving Peel's freight transportation system. 2) The establishment and furthering of Peel's critical goods movement corridors. Staff and the new partnership will develop a multimodal goods movement network across the Region. The critical goods movement corridors and activity centres will be assessed and prioritized for preservation and improvement. 3) The coordination between levels of government of goods movement policy development and funding to support Peel's initiatives. In the area of policy development there needs to be

better synchronization between levels of government to ensure consistency of policy development and to determine the appropriate level of funding that should be directed to goods movement investment in Peel. Policies at all levels need to be more strongly oriented towards the development of freight transportation/land use systems that are sustainable over the long term.

The study recommendations are intended to strengthen the current goods movement policies in the Official Plan and the findings from this study are being used currently in the development of Peel's Long Range Transportation Plan.

The Peel Goods Movement Study sets the stage for the formation of the Peel goods movement liaison team and for further analysis of Peel's freight corridors and activity generators. An initial task for the Peel goods movement liaison team will be to prepare an action plan of needed short term improvements for the goods movement system in Peel and to report back to the Peel Regional Council.

We look forward to dialogue with other municipalities that are embarking on new initiatives in municipal goods movement or contemplating such an activity. We have appreciated the assistance and professional advice of IBI and Wilbur Smith Associates in our initial venture into advancing the agenda

Acknowledgements

The author would like to express appreciation to colleagues who have assisted with the preparation of this paper, including Eric Chan, Reena Sachdeeva and Sara Orchard. The support of Director Tom AppaRao and Commissioner Nick Tunnacliffe is also gratefully acknowledged.

The views expressed in this paper are strictly those of the author, and may not reflect the policies of the Regional Municipality of Peel.

Murray McLeod

Urban Transport Issues and Improvement Options in Bangladesh

M.M. Hoque, B. Khondaker , M.J.B. Alam
Accident Research Centre
Bangladesh University of Engineering & Technology (BUET)
Dhaka-1000, Bangladesh
E-mail: dirarc@arc.buet.ac.bd

ABSTRACT

Bangladesh is a very densely populated and low lying country with around 130 million inhabitants living in an area of 144,000 sq.km i.e. some 900 inhabitants per sq.km. Population growth remains high at around 1.6 percent per annum with nearly half of the total population is under 15 in households of average size of 5.3 people. The current urbanization level at around 25 percent, more crucially for transport, is expected to rise to 30 percent by the year 2010 and to 50 percent by the year 2025. Metropolitan Dhaka accounts for nearly 40 percent of total urban population. The rapid urbanisation process, high vehicular population growth and that of the mobility, inadequate transportation facilities and policies, varied traffic mix with over concentration of non-motorised vehicles, absence of dependable public transport system and inadequate traffic management practices and parking facilities have created a significant worsening of traffic and environmental problems in the major urban centres of Bangladesh. In this paper characteristics of the urban transportation problems and consequent mobility, safety and environmental effects are discussed with particular reference to the situation of metropolitan Dhaka. Key transport issues and possible solution strategies encompassing various opportunities and interventions for alleviating such problems are discussed with particular reference to both existing and future context with emphasis on wider application of traffic management measures. The significance of the recent initiative for the development of a Strategic Transport Plan (STP) in addressing the problems both in the short and long term perspective is also highlighted.

INTRODUCTION

The urban transport issues of mobility, congestion, safety and environmental aspects are becoming increasingly important and critical in Bangladesh. The rapid urbanisation process, high vehicular population growth and that of the mobility, inadequate transportation facilities and policies, varied traffic mix with over concentration of non-motorised vehicles, absence of dependable public transport system and inadequate traffic management practices and parking facilities have created a significant worsening of traffic and environmental problems in the major urban centres, Dhaka in particular. There is now an ever-increasing urgency for mitigating the complex transportation problem. It is high time that we could take some lessons from advanced countries and move in a planned and phased manner for solving the ever increasing urban transportation problems. The purpose of this paper is to briefly discuss some of the key urban transport characteristics and discuss improvement options and opportunities in Bangladesh by highlighting the significance of the very recent initiative of the development of a Strategic Transport Plan (STP) for Metropolitan Dhaka. This paper particularly concentrates on the issues in Metropolitan Dhaka, the capital of Bangladesh and is an extension of the earlier paper by Hoque and Alam (2002). Also discussed here are the possible solution strategies for enhancing mobility, safety and the environment by means of better traffic management measures.

URBANIZATION AND CONGESTION ISSUES

Urbanization Trends

Bangladesh is a very densely populated and low-lying country with about 130 million inhabitants living in an area of 144,000 sq.km i.e. some 900 inhabitants per sq.km. By 2024, the population of Bangladesh is expected to touch 168.7 million with very large percentage residing in urban area. Urbanisation in Bangladesh has been increasing steadily over the last 10 to 20 years with current urbanisation level at around 24 percent. The percentage of people live in urban areas is expected to rise to 30 percent by the year 2010 and to 50 percent by the year 2025. The rate of urbanisation is alarmingly high when compared with other developing countries. The current rate of urban population growth in Bangladesh is found to be the highest in Asia. Rural-

urban migration is considered to have made the major contribution towards the rapid growth of urban population in Bangladesh. A very high proportion of this urban population is poor. The greater challenge thus for transportation professionals is to develop a system of urban transport that meet the basic mobility needs for all urban dwellers at desirable safety and avoiding the unacceptable level of congestion and its consequent overwhelming adverse environmental effects.

Congestion Issues

Conceptually, congestion results from increasing disparities between urban transport demands and the supply of transport services (Bovy, 1990). These disparities are depicted in Figure 1. This situation is particularly prevalent in Bangladesh where transport infrastructural developments are unable to keep up with poor general maintenance of vehicles, faulty traffic planning and disregard for traffic regulations.

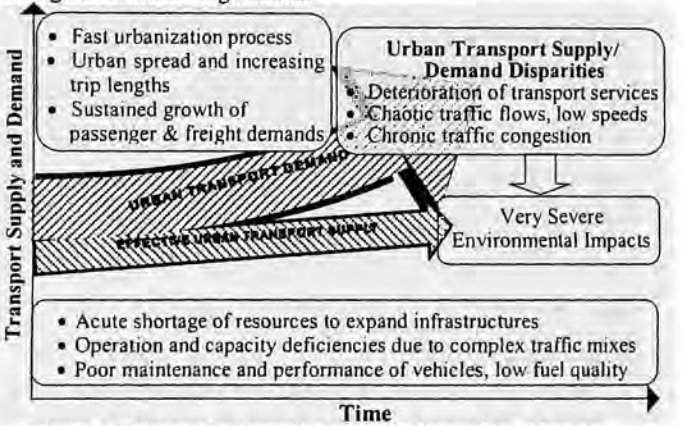


Figure 1: Increasing supply demand disparities of urban transport in cities of developing countries

There are many manifestations of congestion, which may be grouped into two basic classes: point congestion and network congestion. Point congestion is concerned with bottlenecks, i.e. site-specific constrictions on traffic capacity. Bottlenecks may be permanent, or temporary (e.g. at accidents or breakdowns). The effects of bottlenecks are localised, although removal of an individual bottleneck may have important and sometimes severe repercussions across a surrounding network. Network congestion

is an area-wide phenomenon, characterised by unstable flow conditions that may erupt at a variety of sites in an area, either at the link level or at the node level. Network congestion is indicative of a general lack of road capacity in a region, another term for this phenomenon is *gridlock*.

Congestion Factors

It has been noted that the potential causes of ever increasing urban traffic congestion in Bangladesh are mainly attributed to the following:

- The varied traffic mix and heavy concentration of non-motorised vehicles with almost 70 per cent of the available road space is occupied by rickshaws and their dominance is expected continue in the foreseeable future.
- The absence of a dependable public transport system with only about 10-20 per cent of trips catered for by bus and motorised paratransit modes. Non-motorised modes account for 80-90 per cent of the daily total passenger trips.
- Inadequate road infrastructure; road network, intersections and links have not been built on the basis of modern principles of road design. The problems are worse at intersections.
- Inadequate traffic management practices; substantial road space is unavailable to road users, slow and fast moving traffic and pedestrians are unsegregated, traffic signal timings are inappropriate, and there is inadequate enforcement of traffic regulations.
- Undue encroachment onto the available roadspace by vendors and traders. Also the lack of planning control has led to inappropriate roadside developments.
- Poor road user behaviour and education, particularly failure to give way & obey traffic rules.

Mobility of the Urban Poor

The issue of mobility is particularly important for the urban poor in the context of sustainable urban transport in developing countries. Replogle (1993) argued that income plays a significant role influencing transportation choices people have. People with low-income face extremely limited transport choices. Where there is extensive poverty, it is vital to ensure that the modes used by the poor continue to remain available as travel options. Despite rising incomes in many cities across Asia, the distribution of wealth and income remains skewed in much of the

region. Rapid urbanisation and economic growth throughout much of Asia has left behind hundreds of millions of people, who continue to live in desperate poverty. Indeed, two-thirds of the poorest of the poor in the world live in India, Bangladesh Pakistan, and China.

Access to transport is critical for the urban poor, for without it they are literally at risk of physical, economic and social isolation (ADB 2001). The following travel characteristics of urban poor are of particular importance:

- Purpose of trips: the great majority of travel by the urban poor is accessing work and for education.
- Frequency of trips: poor make fewer motorised and shorter trips than average.
- Length of trips: time and effort as well as money limit the range of trips made by the poor.
- Gender differentiated demand: women are much more likely to walk or use cycle-rickshaws and bus.
- Travel mode: mode for the poor is often determined by trip length, for which there may be only one or two possible modes (walking and buses).

STRIKING URBAN TRANSPORT CHARACTERISTICS

Transport Profile of Dhaka – The Mega-City

Dhaka, the capital city of Bangladesh, is the largest and most industrialized city in this riverine country of some 130 million. Dhaka is the administrative, commercial, and culture centre of the country and also continues to serve as the traditional centre of wholesale trade. The population of the Dhaka Metropolitan Area is presently estimated to be 12 million (2004), which is expected to more than double by 2024. The population of expanded Dhaka conurbation is expected to reach 36 million by 2024.

The demographic trends of the last decade that have resulted in rapid population growth are expected to continue in the coming decades. The impact of such rapid growth has major consequences on the ability of the transport sector to provide mobility for all people as they seek to take advantage of employment, health and social opportunities.

The transport sector in Dhaka is comprised of many different modes of travel-both motorized and non-motorized-often using the same road space – resulting in a high level of operational

disorder that significantly diminishes the efficiency and effectiveness of the existing transport uses. Dhaka is perhaps the only city of its size without a well-organized properly scheduled bus system or any type of mass rapid transit system. The deteriorating traffic conditions are causing increasing delays and worsening air pollution, and seriously compromise the ability of the transport sector to serve and sustain economic growth and quality of life. The following are the brief details of Dhaka:

- Nearly 40 percent of country's urban population live in Dhaka.
- Poverty incidence – around 50 percent (highest in Asian cities)
- Dhaka's contribution to country's GDP – nearly 15 percent
- Road network–3000km (with only 450km primary & secondary/ collector roads)
- Road vehicles nearly 250,000 with huge number of rickshaws (500,000)
- Trip percent catered by Mass Transport – 10 percent.
- Cost of congestion and accidents – US$ 520 million per annum
- Air pollution and noise level 4 (in 1-10 scale)
- Urban road accident fatalities – 400 (reported per annum)
- Percent road fatalities who are pedestrians -60 percent
- Non-existent of qualified traffic engineering professionals.

Walking and Rickshaw – Dominant Travel Modes

Pedestrian Share: Analysis of data from home interview and O-D (origin-destination) surveys reveals the predominance of non-motorized transport (NMT) and demonstrates that pedestrian is by far the largest identifiable group (Table 1). It can be seen that over 60 percent of all link trips in Dhaka involve walking alone, followed by rickshaws 13 percent. Equivalent trip studies in medium sized cities showed walking represent 20-40 percent of trips. If person trips by vehicle are considered, modal split reveals that rickshaw comprises the highest with 35 percent. Indeed the extent of walk mode in urban areas of Bangladesh is much more compared with other Asian mega cities (Hoque 2000):

- Tokyo: walk and bicycle 25 percent
- Bangkok: walk 16 percent and other NMT 8 percent
- Jakarta: walk 13 percent and other NMT 12 percent
- Delhi: walk and cycle 44 percent and other NMT 6 percent
- Lahore: walk 12 percent and other NMT 5 percent
- Singapore: walk 12 percent and other NMT 4 percent
- Hong Kong: walk 13 percent

Pedestrians thus clearly form by far the single largest group of road users in terms of total catered number of trips in urban areas of Bangladesh. This is particularly prevalent (about 65 percent) for short trips up to one mile. Walking as primary mode of travel is dominant in all age groups and income classes. It is most prevalent for low income groups, nearly 80 percent.

This characteristic will continue to grow in the future as a large part of the urban population would hardly afford any kind of motorised or non-motorised transport of their own. Indeed walking appears to be a major contributor to a sustainable transport strategy. Yet pedestrian can still claim to be our most forgotten and neglected road user group. It is the motorist not pedestrian who normally receives the attention and greater share of priority. Pedestrians need protection in the form of facilities by ensuring their safety and convenience.

Table 1: Metro – Dhaka: Person Trip by Mode, 1999

Mode	No. of Person Trips		Ave. Trip	Passenger-km	
	000/ day	%	Length (km)	000/ day	%
Car	576	4.0 (10.5)	10.4	5,990	11.9
Bus	1,482	10.2 (27.0)	13.5	20,007	39.8
Auto-rickshaw	845	5.8 (15.2)	12.8	10,816	21.6
Rickshaw	1,927	13.3 (35.0)	4.3	8,286	16.5
Others	675	4.7 (12.3)	7.5	5,066	10.1
Subtotal	5,505	38.0 (100.0)	9.1	50,165	100.0
Walk	9,000	62.0	1.0	9,000	15.0
Total	14,505	100.0	4.1	59,165	100.0

* Figures in bracket are subtotal percentage

Rickshaw Share: In general cycle rickshaws cater quite significant proportion of trips (between 25 and 40 per cent) in cities, about 40 per cent of the metropolitan population are directly or indirectly dependent upon rickshaws. As such rickshaws occupy an important position in urban transport scenario and posses some outstanding advantages as well. It is important to develop and augment NMV facilities both in urban and rural areas, particularly physically separated spaces to protect their legitimacy and safety. Studies are needed to assess the extent to which the level of NMV mode be maintained to achieve a reasonable balance between the question of urban transport demand and supply as well as for comprehensive understanding of the effects of specific policy measures. Most recent research

has demonstrated substantial improvements to the operation of motorised vehicles on the main traffic corridors as a result of banning non-motorised vehicles, some 30 percent reduction in travel time and an increase of passenger movement capacity by nearly 300 percent (Hossain 1996). These results could provide some insights in determining standards and policy framework for mixed traffic in urban areas.

Level of Bus and Minibus use

Although their contribution as travel modes has recently been enhanced, their patronage share remains much low, accounting for around 10 percent. The bus fleet grows at very low growth rate of only 2 percent. The present bus transport mode has not been able to perform satisfactorily and efficiently. Long waiting, delay on plying, overcrowding, lack of comfort and long walking distance from the trip origin are some of the common problems.

Level of Bicycle use

It is highly expected that cycling mode should be able to play its effective part in meeting the transport demands. The most outstanding advantage of the bicycle is its availability for door-to-door trips and flexibility of use. However, the usage of bicycle as travel mode in urban areas is for below the desirable level, accounting for nearly 1 percent in metropolitan Dhaka. The bicycle share is however more in medium sized cities, nearly 5 percent.

Traffic Congestion

As highlighted earlier, in spite of the low level of motorization, Dhaka suffers from serious traffic congestion largely due to the absolute lack of roads, deficient road network configuration and inefficient traffic management (JBIC 2000). Conflict in the use of road space by pedestrians, rickshaw, bus, car auto-rickshaw, and others is visible everywhere in Dhaka and their movements are not properly segregated nor managed (Typical examples are shown in Figures 2 and 3). Poor parking facilities and management, unorganised roadside activities and lack of public transportation and pedestrian facilities among others contribute to the problem. The low awareness level of drivers and pedestrians to traffic rules and regulations, as well as insufficient and poor enforcement capacity of traffic police, further aggravates the situation.



Figure 2: Chaotic Pedestrian Movement in City Road

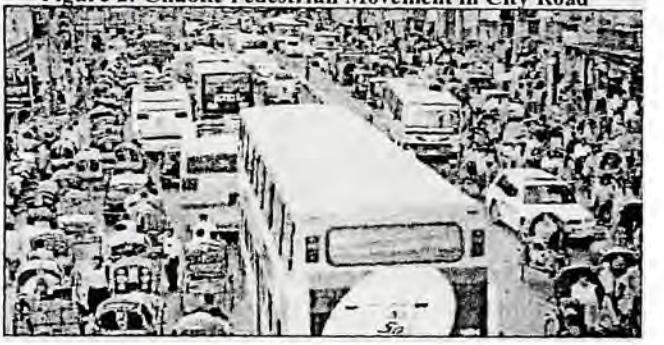


Figure 3: City Street Scene in Dhaka

Urban Road Safety

In urban areas of Bangladesh pedestrians represent often up to 70 percent of the total road accident fatalities. Their involvement varied between 43 and 73 percent of the road accident fatalities in the medium sized cities. Current statistics revealed a deteriorating situation in metropolitan Dhaka. For example, pedestrians as a proportion of road accident deaths increased from 43 percent in 1986-1987 to 67 percent in 1991-1992. In recent years (1996-1998) the number of pedestrian casualties (fatalities and injuries) has increased markedly, from 433 in 1966 to 588 in 1998, an increase of about 29 percent. Pedestrians now make up approximately 63 percent of road accident fatalities and 32 percent of injuries and are involved in about 20 percent of all reported accidents. Indeed, with fatal accidents, 70 percent were pedestrian motorised vehicle collisions. The seriousness of the

pedestrian accident problem of Bangladesh is clearly evident from the comparative data on pedestrian's involvement in fatal and injury accidents for numerous countries in the Asia and Pacific region. Overall, Hong Kong had the highest share of pedestrian deaths (two thirds of all fatal accidents) followed by Bangladesh (63 percent), Pakistan (50 percent), Republic of Korea (48 percent), Sri Lanka (45 percent), Fiji (43 percent) and Papua New Guinea (33 percent).

Traffic Induced Pollution

Dhaka has become one of the most polluted cities in the world. The environment of the city has been deteriorating rapidly during the last couple of years. Now a days, complaints about headache, eye and skin irritation as well as breathing problem are very common among the users of the roadways of the city. The situation is expected to worsen further with increase of population and economic development. The conditions are severely harmful for the children and the elderly people. In the roads of Dhaka city the concentration of Oxides of Nitrogen (NO_x), Oxides of Sulphur (SO_x), Carbon Mono-oxide (CO) and Suspended Particulate Matter (SPM) are 500, 1200, 7500 and 2500 $\mu gm/m^3$ respectively which far exceeds the acceptable limit set by WHO. Car is the principal contributor of SO_x, NO_x and CO emission in the city. The contributions from bus and truck are significant in the case of SO_x and NO_x emission. Auto-rickshaws contribute significantly in the case of SO_x and CO emission. To make the city environmentally sustainable it is required to assess the level of air pollution and develop policies to control the emission of the pollutants as soon as possible.

STRATEGIES FOR ALLEVIATING THE PROBLEMS

In the light of the facts and discussions presented in the previous sections, the following key issues and strategies are outlined for addressing the prevailing and emerging urban transport problem in order to achieve sustainability and safety.

- Economic growth coupled with rising income and urbanisation causing increased transportation demand and no of vehicles.
- Motor vehicle growth rate of 8 to 10% causing considerable congestion and safety problems. Vehicle fleets may be doubled in around 7 years with high number of NMV.
- Well-planned and well co-ordinated approaches are required to cope with future increases.

Directions for the future

Travel options available in developing countries are defined by a combination of technology, form of operation and management measures which can provide priority over other traffic (ADB 2001). These options include walking, non-motorised transport, paratransit, minibus and bus, sub-urban rail and MRT (busway, light-rail transit/tramway, and metros).

- Walking as travel mode would dominate as the largest single group in cities, particularly for the poor.
- The dependence on non-motorised modes will persist in foreseeable future. Bicycles should be treated as an integral part of the urban transport system by providing adequate bicycle facilities (e.g., bicycle paths) with proper design. Policy should favour affordability of bicycles by the poor.
- Measures to separate non-motorised to motorised traffic should be given increased priority.
- Priorities are needed to significantly increase the fleet, frequency and operating conditions of bus transportation.
- Interventions should heavily rely on traffic management and non-capital intensive measures.

Augmentation of mass transit system is an urgent necessity to ensure mobility need, road safety, and liveable urban environment for Dhaka city in future. Rail based rather than road based mass transit system seems to be more appropriate with desired service quality of trip makers of Dhaka. The urgent issues like prevailing socio-economic context, the deteriorating environment of Dhaka, serious urban road accident situation, scarcity of road space, population relocation, and future adaptability should be stressed for the introduction of a dependable mass transit system. Rail based tramways, with its favourable inherent cost and service characteristics can be a potential option for improvements of existing public transport system.

Traffic Management Strategies

Traffic management in urban areas is being used to provide controls on congestion by introducing improved traffic flow and lane occupancy and to restrict access to certain areas at peak periods. An effective management strategy may include: re-routing through traffic, traffic calming measures, improving the geometric layout of junctions, improving public transport infrastructure, measures to improve facilities for pedestrians and

cyclists. When adopted sensibly, traffic management measures could be a useful tool in improving the environment and encouraging walking and cycling. It is clear that environmental issues cannot be addressed in isolation and need to be integrated with various technical, social, political and economic aspects. Efficient transportation systems and control of future urban development form a vital link in improving the urban environment and sustaining development.

There are clear relationships between traffic congestion and environmental pollution. Improvements in traffic systems will result in better environment. There is a wide range of technical solutions available to help reduce pollution from vehicular traffic. In the industrialised countries, many of these solutions are being examined by motor vehicle manufacturers in response to the legislative pressures from national governments, who wish to see environmental protection measures implemented. The available technologies for reducing energy consumption and improving energy efficiency can be conveniently classified into five groups, namely: engine technology, alternative fuels, vehicle design, public transport and traffic management. Reducing environmental impacts through planning policies and traffic management are particularly important and demands greater attention in developing countries like Bangladesh.

Under the prevailing socio-economic context and for mitigating complex urban traffic and transportation problems in Bangladesh traffic management elements requiring urgent attention include pedestrian facilities, NMT, transit system, freight movement etc. Further detail in this regard is available in Hoque et al (2002).

Strategies for Road Safety

The safety situation is very severe by international standards. Pedestrian–vehicle conflicts are clearly the greatest problem with significant involvement of trucks and buses. There is specific need for road environmental improvements aimed at correcting the most common deficiencies. Strategies required preventing pedestrian-vehicle conflicts includes:

- For road environment improvement:
  - Wider application of traffic engineering approaches.
  - Both accident prevention and accident reduction strategies.
  - Systematic implementation of road safety audit.

- Priorities for principles like traffic segregation, channelisation, speed reduction and other self-enforcing measures.
- Implementation required at the places of blackspots identified by systematic accident investigation.
- Strategies should offer cost-effective results.
- Preparation of road traffic safety action plan.

Future actions would require renewed governmental as well as organisational commitments by setting realistic road safety targets. Supports, essentially funding and collaborating, from international agencies and specialised institutes would be particularly important in tackling the problems. With particular regard to the vulnerable road users' safety, it is also important, as Mohan (1991) argued, to consider some fundamental changes in future policies: increasing public transport, promoting safer bicycling, reducing motorised two-wheeler production, urban planning which reduces trip lengths, rationalisation of vehicle speeds to reduce variation and lower average speeds etc.

STRATEGIC TRANSPORT PLAN (STP) FOR DHAKA

Recently the Government of the People's Republic of Bangladesh (GOB) has committed to providing better living conditions for the populace and one of the main sectors in which action will be taken is in urban transport. Of most recent achievement has been the development of Strategic Transport Plan (STP) for greater Dhaka, the capital of Bangladesh to look ahead with a wider vision and to plan for the future with a co-ordinated approach to transport and land uses so as to maximise resources and minimise the adverse impacts on the people and environment. It is expected that STP will provide a sense of direction towards achieving this objective (STP 2004).

Scope of the STP

The primary objectives of the Strategic Transport Plan Study are as follows:

- Develop a coherent long-term Strategic Transport Plan (2004-2024), following and updating the earlier Dhaka Integrated Transport Study (DITS) and other transport related studies;
- Formulate an Urban Transport Policy document for governmental discussion and approval that would guide urban transport development, operations and management in Dhaka Metropolitan Area (DMA);

- Identify institutional weakness of several organizations and prepare a specific plan for their institutional strengthening and capacity building in the area of urban transport strategic planning; and
- Finalize the scope and Terms of Reference for the priority multimode transportation improvement projects identified in the Strategic Transport Plan for the first five years.

Key Policy Issues

The STP project has identified fifteen key emerging issues which is hoped to influence the policy decisions. These emerging issues are:

- Safety of all road users and adequate pedestrian facilities
- Dependable public transport system Mass rapid transit system
- Control on the numbers and operations of NMT
- Travel demand management measures
- Urban freight transport
- System integration for increased efficiency and patronage
- Traffic management measures and parking supply and demand
- Environmental issues and Social and political aspects
- Proper integration between transport and land use planning
- Institutional development
- Privatisation, deregulation and subsidies

Of necessity, the STP has investigated a wide-ranging series of issues. Detailed descriptions of these issues are well beyond the scope of the paper and can be seen in the Strategic Transport Plan (STP, 2004). Starting from a consideration of population and economic growth patterns, the study proposed a series of possible visions for the likely growth and settlement patterns for the 20 year planning horizon up to 2024. Throughout the course of the STP study, a large number of issues have been raised; covering a broad spectrum of multi-modal transport systems, and numerous recommendations has been proposed. A set of seven recommendations has been labelled as 'primary recommendations' and those are: traffic management, bus consolidation, integrated mass rapid transit system, selected highway projects, safety improvements, pedestrian facilities and railway resolution each of which is summarized in Strategic Transport Plan (STP, 2004). Critical examination and increased consideration of the sustainability of these policy issues and recommendations are particularly important in alleviating the problems.

CONCLUDING REMARKS

The rapid rise in population along with increased and versatile urban land use patterns have generated considerable travel demand as well as numerous transport problems in major cities in Bangladesh. The congestion, safety and environmental problems are rapidly growing due to the combined effects of rapid motorization and urbanization particularly in Dhaka, the capital of Bangladesh. This paper has briefly reviewed the striking urban transport problem characteristics and the factors inhibiting mobility and the development of a safe and sustainable transportation system with particular reference to the situation in Dhaka. Strategies encompassing various opportunities and interventions for alleviating such problems are discussed with particular reference to both existing and future context.

REFERENCES

Asian Development Bank (ADB), Urban Transport Policy Paper, Finalise Draft, ADB, Manila, June 2001.

Bovy, Ph.H. (1990), Transport and the environment; a tentative overview of issues in cities of developed and developing countries, CODAU-V, Sao Paulo, September, 1990.

Hoque, M.M. (2000), Road planning and engineering for promoting pedestrian safety in Bangladesh, 10th REAAA Conference, Tokyo, September 2000.

Hoque, M.M., McDonald, M. & Hall, R.D. (2001), Road safety improvements in developing countries: priority issues and options, ARRB Transport Research, Melbourne. March 2001.

Hoque, M.M., Ahsan, H.M & Alam, J.B (2002), Traffic congestion and safety in Dhaka city: some contemporary issues and options, ASCE Local Chapter (Bangladesh), Dhaka.

Hoque, M. M., Hossain, T. (2004), Augmentation of Mass Transit Mode in Dhaka, Bangladesh, CODATU XI.

Hossain, M. (1996), Modelling of traffic operations in urban networks in developing countries, Ph.D. Thesis, Dept of Civil & Env. Engg. University of Southampton, August 1996.

Japan Bank for International Co-operation (JBIC) (2000), Study of the Improvement of Transportation and Environment in Dhaka, Dhaka.

Mohan, D. (1991), Safety of Vulnerable Road Users, Proc. of the XIXth PIARC World Congress, Marrakech, 1991.

Replogle, M. (1993), Bicycle and cycle rickshaws in Asian cities, 6th conference on urban transport in developing countries, CODATU-VI, Tunis, 1993.

Strategic Transport Plan for Dhaka, The Louis Berger Group, Inc, Bangladesh Consultants Ltd, Final Draft-31 October 2004.

Aircraft Noise: A Survey of the Regulatory Framework and Enforcement Mechanisms in Canada

Michael T. Kariya[1]

Air travel is a relatively recent phenomenon. While air transportation's infant years conjure images of rural grass landing strips and biplanes, the state of air transportation today is characterized by a much more complex and massive operation. To accommodate the needs of modern society, most large cities are serviced by airports that are in close proximity to, if not surrounded by, the urban landscape. When one considers that the large urban airports are the operation hubs of large jetliners, it is not difficult to see why the issue of noise is of serious concern to neighbouring property owners.

In Canada, Transport Canada and local Airport Authorities are responsible for balancing the needs and concerns of affected landowners on the one hand, and the air travel business, and industry it supports, on the other. This paper considers the issue of noise associated with air transportation by examining the administrative regimes applicable to airports and aircraft. This discussion is followed by a survey of various regulatory and legal enforcement mechanisms.

Part I – Regulatory Framework

Aeronautics as a Federal Entity

In 1932, the Privy Council held that the field of aeronautics was a federal undertaking supported by s. 132 of the *Constitution Act, 1867*.[2] The Privy Council also suggested that aeronautics could be supported by the federal peace, order and good government (POGG)

---

[1] B.A. (U.B.C.), LL.B. anticipated 2005 (Calgary).

[2] *Aerial Navigation A.-G. Can. v. A.-G. Ont.* [1932] 1 D.L.R. 58.

power. Indeed in *Johannesson v. West St. Paul*,[3] the Supreme Court of Canada held that aeronautics was properly sustained by the POGG power, and that any new matter concerning aeronautics that was not already assigned to the federal government, could fall within the catchall POGG power. The test may be stated as follows: Does the impugned legislation go to the essence or root of the federal undertaking; if it does, the law will be found to be *ultra vires* the subordinate level of government. Therefore, in *Mullaney v. Red Deer*,[4] it was held that laws which are in pith and substance land use, will be *intra vires* the province; however, if the law is fundamentally targeted at aeronautics, it will be *ultra vires* the province.

The problem for provinces and municipalities is that noise is a natural byproduct of aircraft operations, and thus it is difficult to imagine any legislation or bylaw which could effectively regulate noise at the source without also directly regulating aircraft operations. As a result, provinces that wish to address the issue of aircraft noise in their communities must legislate carefully, so that the focus of the legislation is in pith and substance a matter falling under section 92 of the *Constitution Act*, 1867. For example, provincial initiatives may focus on land use planning, or building codes which stipulate special noise insulating requirements.[5] However, it should be noted that provincial land use initiatives are always subject to federal authority to engage in zoning regulation of areas surrounding airports. While federal regulations cannot address how buildings around airports should be constructed, they may address matters such as building height restrictions and prohibited land uses which may attract wildlife.

The Aeronautics Act and Regulations

In Canada, the *Aeronautics Act* gives the Minister of Transport responsibility to enact regulations in relation to aircraft and airport

---

[3] [1952] 1 S.C.R. 292.

[4] [1999] A.J. No. 648.

[5] For e.g., *Calgary Airport Vicinity Protection Area Regulation*, Alta. Reg. 318/1979, specifies that land uses within a certain radius of the airport comply with Part 11 of the Alberta Building Code, 1990.

noise.[6] Pursuant to this authority, the federal government has enacted regulations in three main areas: (1) the use of noise restricted runways,[7] (2) aircraft operating restrictions,[8] and (3) noise abatement at the source.[9]

Since the first and second categories are primarily concerned with day-to-day operations, these regulations are specifically targeted at pilots and air traffic control personnel. Transport Canada communicates these regulations via applicable noise abatement procedures and noise control requirements, which are published in the *Canada Air Pilot*, or the *Canada Flight Supplement*. These publications are amended on a 56 day cycle and are produced according to region/province. The noise abatement procedures addressed in these publications are listed as follows, as they appear in the *Canadian Aviation Regulations*: (a) preferential runways; (b) minimum noise routes; (c) hours when aircraft operations are prohibited or restricted; (d) arrival procedures; (e) departure procedures; and (j) the minimum altitude for the operation of aircraft in the vicinity of the aerodrome.[10]

With regard to noise restricted runways, section 602.106 of the *Canadian Aviation Regulations* stipulates that the restrictions apply to aircraft exceeding a take-off weight of more than 34 000 kg, with a broad exemption afforded to aircraft that meet prescribed noise emissions standards. Restricted runways are specified in the *Canada Flight Supplement* and are meant to prevent noisier aircraft from assuming a flight path over heavily populated areas.

While certain airports have implemented nightly curfews, or hours when aircraft operations are prohibited, pursuant to section 602.105(c) of the *Canadian Aviation Regulations*, such measures have not been uniformly adopted. Instead, many airports rely solely on the use of "preferential runways", taking account of such factors as

---

[6] *Aeronautics Act*, R.S.C. 1985, c. A-2, s. 4.9(f).
[7] *Canadian Aviation Regulations*, SOR/1996-433, s. 602.106.
[8] *Ibid.*, s. 602.105.
[9] *Ibid.*, s. 602.150 – 602.162 (Transition to Chapter 3 Airplanes).
[10] SOR/1996-433.

an aircraft's noise emissions and time of day, to determine the preferred runway.

Departure and arrival procedures enacted under sections 602.105(d) and (e), usually integrate sections 602.105(a), (b) and (j) of the *Canadian Aviation Regulations*. Hence, departing aircraft may be required to use certain runways based on an aircraft's noise emissions rating. Additionally, under these regulations, aircraft are often required to maintain a straight course out until a particular altitude is achieved. This prevents aircraft from flying over high density areas at low altitude; it also has the effect of channeling aircraft through particular corridors, thereby limiting the most severe impacts of aircraft noise to these corridors. Arrival procedures are usually similar to the departure procedures. These procedures normally mandate that an aircraft intercepts a course that lines up with the runway's centre line before descending below a certain altitude.

The third category, noise abatement at the source, primarily addresses Canada's international treaty obligations under Annex 16 of the *Chicago Convention*[11] to phase-out older and noisier Chapter 2 aircraft. "Chapter 2" refers to the chapter in Annex 16 of the *Chicago Convention* that relates to aircraft prototypes certified prior to 6 October 1977. Moreover, Canada has agreed to comply with future aircraft phase-outs to meet increasingly stringent noise emission standards.[12] Based in Montreal, the International Civil Aviation Organization (ICAO) is the United Nations body responsible for the administration of the *Chicago Convention*. ICAO oversaw the phase-out of all Chapter 2 aircraft by 1 April 2002 among signatory states.

Land Use Planning – Alberta's Municipal Government Act and Airport Vicinity Protection Area Regulations (AVPA)

Since the provinces cannot regulate in relation to aircraft or airport operations, they have turned to land use planning as an alternative to minimizing aircraft noise impact on surrounding

[11] *Convention on International Civil Aviation*, 7 December 1944 (*Chicago Convention*).

[12] The "Chapter 3" standard is the current standard, and applies to all prototypes accepted since 6 October 1977 and before 1 January 2006.

communities. The ability to regulate land use of areas adjacent to major airports makes it possible to prevent development which is incompatible with the elevated levels of noise. Section 693(1) of the Alberta's *Municipal Government Act*[13] gives the Lieutenant Governor in Council the power to make regulations with respect to land use in the vicinity of the Calgary and Edmonton International Airports. Any regulation made under section 693(1) supersedes "any statutory plan, land use bylaw or other regulations" under Part 17 (Planning and Development) of the *Municipal Government Act*, and is binding on "any subdivision authority, development authority and subdivision and development appeal board and the Municipal Government Board".[14] Furthermore, section 694(6)(b) allows the Lieutenant Governor in Council to make regulations "prescribing how municipalities are to manage the use and development of land in the vicinity of an airport".[15]

Pursuant to the *Municipal Government Act*, two Airport Vicinity Protection Area (AVPA) Regulations were created, the *Calgary International AVPA Regulation* and the *Edmonton International AVPA Regulation*.[16] Schedule A of these regulations establishes the Protection Area and gives a legal description of the relevant lands. Lands contained within the AVPA are subject to special regulations contained in Schedule C. Table 1 of Schedule C lists a variety of land uses, which according to the Noise Exposure Forecast (NEF), are permissible (P), prohibited (N/A), or have conditions attached (C followed by a footnote corresponding to the condition). Table 2 of Schedule C stipulates what conditions apply with regard to those uses which have been made subject to conditions. Of particular interest are conditions C1 and C2, which require that a building meet a certain standard of acoustic insulation, pursuant to Part 11 of the *Alberta Building Code, 1990*. The NEF model takes into consideration inputs

---

[13] *Municipal Government Act*, R.S.A. 2000, c. M-26.
[14] *Ibid.*, ss. 693(2)(a) and (b) respectively.
[15] *Ibid.*, s. 694(6)(b).
[16] *Calgary International AVPA Regulation*, Alta. Reg. 318/1979; *Edmonton International AVPA Regulation*, Alta. Reg. 63/1981.

such as existing and proposed runways, aircraft operations and volume, noise certification levels, fleet mixes, and combines this with results from studies done in the United States on community response to noise. The result is an NEF contour that predicts the accumulated perceived noise levels that would be experienced in the vicinity of the airport during an average busy day. Theoretically, higher contour numbers result in greater objection to the noise.

The Local Airport Authority and Development of Noise Abatement Procedures

In 1994, the federal government implemented the National Airports Policy (NAP). This initiative involved the transfer of operatorship of Canada's 26 most active airports from the federal government to local arms length operators (Airport Authorities). Under this arrangement, the federal government remains the owner of the transferred airports and continues to maintain its regulatory role via Transport Canada; the Airport Authorities now assume responsibility for financial and operational management.

Since the Airport Authority oversees day-to-day operations, it is logical that the federal government has a process in place for amending existing noise abatement procedures, based on Airport Authority recommendations. The comprehensive process is outlined in Aerodrome Safety Circular 2002-018 (Appendix A – Noise Abatement Procedures and Restrictions Implementation Process). This paper simplifies the process and focuses on how it has been implemented by the Calgary Airport Authority.[17]

The first stage is comprised of describing any existing problems, outlining the rationale behind suggested changes, and reporting the results of any trials or experimentation. Proposed solutions are tabled along with a comprehensive survey of the following, where

[17] A more comprehensive version of the process is located at Transport Canada's website: <www.tc.gc.ca/civilaviation/aerodrome/safetycirculars/2002018appendixa.htm>. Information specific to the Calgary Airport Authority can be found at: <http://www.calgaryairport.com/document.cfm?DID=149>.

applicable: alternatives, associated costs, noise impacts, effects on aircraft emissions, effects on current and future airport capacity, implications of not proceeding with the proposal, implementation issues, impact on aviation system, safety implications, air traffic management, and fleet impact.

Second, the Airport Authority is required to consult with stakeholders. These interests are served by holding various stakeholder meetings. The Calgary Airport Authority discharges its duty to consult by holding Airport Community Noise Consultative Committee Meetings (ACNCC). Stakeholders that must be consulted include: community associations,[18] airlines, and other agencies such as city planning departments, Transport Canada, and Nav Canada. Additionally, the Calgary Airport Authority has a Noise Abatement Procedure Working Group (NAPWG) which discusses concerns expressed at ACNCC meetings, and makes any revisions to the proposed procedures. These concerns may or may not be addressed at future ACNCC meetings. The final step is submission to Transport Canada. If the changes are approved, they are then inserted into the *Canada Air Pilot* or the *Canada Flight Supplement*. In contrast to the elaborate consultation process that is required of the Airport Authority, Transport Canada can unilaterally change noise abatement procedures without having to consult.

Part II – Monitoring Compliance and Enforcement

***The Airport Authority*[19]**

At the Calgary International Airport, the Calgary Airport Authority monitors all departures and arrivals using a flight tracking system which monitors and records information about an aircraft's flight path, altitude, speed, and position. Each plane is identifiable by

---

[18] Community eligibility to sit on the ACNCC is location within the NEF 25 Contour (or greater). See Calgary Airport Authority website: <www.calgaryairport.com/document.cfm?DID=150>.

[19] Information specific to Calgary Airport Authority enforcement practices can be found at their website: <http://www.calgaryairport.com/document.cfm?did=24>.

its own unique transponder signal and code. When a noise complaint comes in to the Airport Authority from a surrounding community, the Airport Authority can retrieve the information on the relevant flight. First, the Airport Authority will determine whether the aircraft in question is civilian or military. The distinction is important because military operations are exempt from compliance with certain operating restrictions. If the aircraft is determined to be a civil aircraft, the next step is to ascertain whether an infraction of a noise operating procedure was committed with reference to information recorded by the flight tracking system. If a *prima facie* infraction has been committed, the Airport Authority will first consider whether there were any extenuating circumstances that may have forced the pilot to deviate from noise operating procedures for safety reasons. If a safety reason is adduced, such a reason may form a defence to the infraction. Conversely, if an infraction of noise operating procedure is ascertained, the Airport Authority will usually notify the offending airline prior to notifying Transport Canada. Should the airline's response be unsatisfactory, or the airline has a high incidence of repeat infractions, the matter is referred to Transport Canada.

The Minister of Transport

In the event that soft enforcement measures fail, or the seriousness of the contravention mandates a stiffer response, section 7.7 of the *Aeronautics Act* authorizes the Minister to levy penalties in accordance with section 7.6 of the *Act*. Section 7.6 of the *Aeronautics Act* allows the Governor in Council, to designate certain regulations and orders as "designated provisions", which carry maximum fines of $5000 for an individual and $25 000 for a corporation. Section 103.08 of the *Canadian Aviation Regulations* addresses contraventions for which sections 7.7 to 8.2 of the *Aeronautics Act* apply. Section 602.105 ("Noise Operating Criteria") of the *Canadian Aviation Regulations* is among the "designated provisions" covered by section 103.08 of the *Regulations*. Under section 7.7 of the *Aeronautics Act*, the Minister is required to provide notice of the alleged contravention, as well as the amount of the penalty that has been levied and time for payment. If the penalty is paid, section 7.8(1) of the *Act* states that no further proceedings under the Part will be

taken. Where the individual or corporation fails to pay within the specified time, the Minister is required to forward the matter to the Civil Aviation Tribunal, pursuant to section 7.8(2) of the *Act*.

The Canadian Civil Aviation Tribunal

Part IV (section 29) of the *Aeronautics Act* establishes the Civil Aviation Tribunal. Provisions specific to the Tribunal are found at sections 29 to 37 of the *Act*. These provisions primarily address the composition and powers of the Tribunal. Additionally, the Tribunal is governed by the *Civil Aviation Tribunal Rules*[20] regulation which addresses hearing procedure and the contents of procedural fairness. The federal government describes the Tribunal as follows:

> The Civil Aviation Tribunal, an independent quasi-judicial body possessing aeronautics expertise, is integral to the enforcement of aviation safety, airworthiness, and aviation security measures in Canada. It fulfills the essential role of providing an independent review of ministerial enforcement and licensing actions taken against holders of Canadian Aviation Documents under the *Aeronautics Act*.[21]

Hence, Ministerial enforcement and licensing decisions are the primary source of appeals to the Tribunal; the sole function of the Tribunal is the adjudication of these decisions.

Since the *Canadian Aviation Regulations* came into force in 1996, replacing the *Air Regulations*, *Air Regulation Series*, and *Air Navigation Orders*, there have been 3 reported cases brought before the Tribunal involving section 602.105 of the *Canadian Aviation Regulations*. These cases were all appeals of penalties levied by the Minister pursuant to contraventions of section 602.105(c) ("hours when aircraft operations are prohibited or restricted"), and all of which occurred at Montreal's Dorval airport. Dorval is rather unique as it is one of the few International Airports in Canada which has implemented a nightly curfew on all departing and arriving aircraft. The curfew currently in place at Dorval is published in the *Canada Air Pilot*, and is effective every day of the week from 2300hrs to

---

[20] *Civil Aviation Tribunal Rules*, SOR/1986-594.

[21] *Civil Aviation Tribunal of Canada Performance Report*, Treasury Board of Canada Secretariat (2001), at 8.

0700hrs for Chapter 2 aircraft and 0100hrs to 0700hrs for Chapter 3 aircraft. No flights are to arrive or depart during the curfew period unless prior permission has been granted by the airport duty manager.

In the first case, *Canada (Minister of Transport) v. Air Transat*,[22] the operator airline landed one of its aircraft outside of aircraft operating hours. Air Transat was fined $5000 by the Minister, for having contravened section 602.105(c) of the *Canadian Aviation Regulations*. However, the most important feature of this case is the Tribunal's discussion of the relevant mitigating and aggravating factors to be considered when reviewing the appropriateness of a monetary penalty assessed by the Minister. These factors are laid out at paragraph 13 of the decision. Aggravating factors include: infractions involving dishonesty; planned breaches; premeditated breaches; extent of harm to victims of the offence; past record of similar offences; and prevalence of the offence. Mitigating factors include: no previous offences; time since last offence; degree of remorse; whether or not an admission of the offence; degree of cooperation with authorities; delay between the commission of the offence and the time of the sentence; conduct (involvement) of any "victims"; restitution; type of operation (commercial or private flight); impact on aviation community; special factual circumstances; relevance of enforcement manual recommendations; effect of a monetary v. suspension penalty on individual; occurrence impact on aviation safety; and manner of proceeding by authorities.[23] The Tribunal noted that the list of criteria is not closed.

The second case, *Canada (Minister of Transport) v. Air Canada*,[24] involved two Air Canada flights (AC 196 and AC 788) that arrived past Dorval's curfew, thereby contravening section 602.105(c) of the *Canadian Aviation Regulations*. AC 196 departed from Toronto with Ottawa as its final destination, but due to bad weather diverted to Dorval without prior permission. AC 788 departed from Chicago's O'Hare airport and was scheduled to land at Dorval; however, it was

---

[22] [1999] C.A.T.D. No. 5.

[23] *Ibid.*

[24] [2001] C.A.T.D. No. 67.

delayed from departing for more than two hours due to bad weather in the Chicago area. AC 788 landed despite the fact that a request for an exemption from the Dorval's duty manager was denied. For both flights, it was held that Montreal's Mirabel airport was an alternative to Dorval, because of its proximity and because it was not subject to the same operating curfew. Air Canada submitted that Dorval was more convenient for the airline and its passengers, as it would position the aircraft at one of its main hubs. The Tribunal rejected this argument on the basis that Air Canada had not received exemptions for either flight, yet chose Dorval over Mirabel for its own economic reasons. On the economics of Dorval over Mirabel, the Tribunal suggested that this line of argument was not material to the case at hand and that it was not for the Tribunal to question the wisdom of a validly implemented policy. In the result, the Tribunal upheld the two fines of $12 500 each for flights AC 196 and AC 788.

Finally, the case of *Canada (Minister of Transport) v. Commercial Aviation Enterprises Inc.*[25] involved a commercial jet that departed Dorval within the operations curfew period. The respondent company submitted that the flight was for the medical purposes of transporting a patient. No exemption from the airport duty manager was ever sought. As a result, Transport Canada requested additional information to substantiate the necessity of departing without an exemption; however, the company's response still yielded nothing to support that the flight was an emergency. Consequently, the Tribunal upheld the finding of a contravention of section 602.105(c) of the *Canadian Aviation Regulations*.

In summary, these cases offer important insight into the scope of the Civil Aviation Tribunal's powers. The *Air Transat* case demonstrates the Tribunal's development of precedent on when it will be appropriate to reduce or increase the amount of a fine assessed by the Minister, by laying out mitigating and aggravating criterion. The *Air Canada* case suggests the Tribunal will be weary of business economics arguments in the face of contraventions where the airline had clear notice of a noise operating procedure policy. Finally, the

[25] [2001] C.A.T.D. No. 48.

Commercial Aviation Enterprises case further highlights the Tribunal's (and the Minister's) receptiveness to emergency situations when they can be substantiated.

Private Enforcement

As an alternative to regulatory enforcement via the Minister and the Civil Aviation Tribunal, the use of the courts offers certain advantages to the private citizen. First, the private citizen has more control over the proceedings than would be the case where the citizen is reporting the matter to the Airport Authority or the Minister; and second, the civil action allows a successful plaintiff to receive a wider range of remedies. For the most part, civil actions have been pursued by private property owners who claim, among other things, a loss of use and enjoyment of their property, property value depreciation, and health-related problems associated with aircraft noise. While there has been some case law on actions framed in trespass,[26] actions in private nuisance have been more common.

The main issue in an action of private nuisance is whether there is an unreasonable interference with one's use and enjoyment of their property. In *Tock v. St. John's Metropolitan Area Board*,[27] La Forest J. suggested that what is "unreasonable" will require examination on a case-by-case basis, in light of all the circumstances, including: "severity of the harm, the character of the neighbourhood, the utility of the defendant's conduct, and whether the plaintiff displayed abnormal sensitivity".[28] Hence, in *Lewis v. St. Stephen (Town)*,[29] a case involving a 15 year old girl who had an aircraft phobia, the court rejected the nuisance action against the local airport on the basis of unusual sensitivity. The court held that such an action was unreasonable in light of the girl's sensitivity and noted that there was no evidence that others in the area were similarly affected. In

---

[26] See e.g. *Lacroix v. The Queen*, [1954] 4 D.L.R. 470; *Harcourt v. Canada (Minister of Transport)*, [1973] F.C. 1181.

[27] 64 D.L.R. (4th) 620.

[28] *Ibid.* at 639-640.

[29] (1981), 34 N.B.R. (2d) 508.

Acciaroli v. Canada,[30] the Federal Court considered the nuisance actions of two property owners who lived in close proximity to the Hamilton Airport. The plaintiffs claimed that noise levels had markedly increased following expansion of the airport's facilities and boundaries. Nevertheless, the court dismissed the claims for lack of actionable nuisance, citing noise measurements showing that post-expansion noise levels were not significantly different from pre-expansion noise levels.

While the above cases demonstrate some of the hurdles that must be crossed in order to make out a successful nuisance claim, even if a plaintiff is successful at making out the nuisance, it is still open for the government to raise the defence of statutory authority. In *Ryan v. Victoria*,[31] the Supreme Court held that liability will "not be imposed if an activity is authorized by statute and the defendant proves that the nuisance is the 'inevitable result' or consequence of exercising that authority".[32] *Sutherland v. Canada*,[33] represents the most recent case law on aircraft and airport noise nuisance actions. Moreover, it illustrates a situation where the plaintiffs successfully made out actions in private nuisance, but were barred from receiving damages due to the defence of statutory authority. The plaintiffs in this case were property owners living in homes within close proximity to Vancouver International Airport and under the flight path to the airport's new North Runway. The plaintiffs claimed they were subjected to unreasonable levels of noise and vibrations due to aircraft operations over the new flight path. At trial, the British Columbia Supreme Court held that the actions in private nuisance had been made out and that there was no statutory authority for the nuisance.

---

[30] (1988), 17 F.T.R. 216.
[31] [1999] 1 S.C.R. 201.
[32] *Ibid.* at para. 54.
[33] [2002] 10 W.W.R. 1, leave to appeal to S.C.C. refused, [2002] S.C.C.A. No. 385.

However, on appeal the British Columbia Court of Appeal[34] overturned the trial judgment, on the basis that while the trial judge did not err in holding that the aircraft noise constituted a private nuisance, he erred in holding that the defence of statutory authority was not established. The Court of Appeal held that (1) statutory authority existed, and (2) nuisance was the inevitable result. On the first criterion, the Court noted that statutory authority did not need to be express, but could be implicit. Hence, authority for construction of the new North Runway could be found in an Order-in-Council, which was authorized by the *Public Lands Grants Act*, as well as section 4.2(o) of the *Aeronautics Act*. Additionally, the precise location and configuration of the new runway was held to be connected to flight and glide paths, which are governed by the *Canadian Aviation Regulations*. On the second criterion, the Court held that the nuisance was the inevitable result of the exercise of statutory authority. Whereas the trial judge suggested that the runway could have been built in a different orientation (thereby altering the flight path), the Court of Appeal held that it was an error to engage in this line of inquiry, and that "the inevitable result test should have been applied to the newly constructed North Runway in its authorized location, and where it was in fact built".[35] Therefore, the *Sutherland* case appears to be clear authority for the following propositions: First, as long as an aeronautical activity is authorized pursuant to statutory authority, whether express or implicit, and nuisance is the inevitable result, an affected property owner will be barred from remedy. Second, the case suggests that the duly diligent purchaser of a home, without notice of future aircraft operations, will have little recourse if he/she is subsequently affected by aircraft noise.

Conclusion

Aircraft noise is a serious environmental issue that has widespread impact on the surrounding community. However, community interests must be tempered by the commercial reality of air transportation as a lucrative industry that has many beneficial contributions to make to

[34] *Ibid.*

[35] *Ibid.* at para. 104.

the larger community. Consequently, a careful balancing of interests must occur; the commercial reality is that aircraft operations are here to stay, though the adverse impacts of aircraft noise can be mitigated. Balancing is accomplished through regulation at the federal level, though subordinate levels of government play an important role in land use planning. Enforcement of federal regulations are the responsibilities of the Minister of Transport and the Civil Aviation Tribunal. Finally, while civilians have recourse to pursue private actions in the courts, current case law suggests that it will be difficult to succeed in an action in tort.

An Urban Governance Model to Achieve Sustainable Urban Transportation and Development: A Case Study of Vancouver, BC

Jesse W.R. Coleman
Department of Civil Engineering, University of Toronto

1 INTRODUCTION

There is a wistful myth that if only we had enough money to spend – the figure is usually put at a hundred billion dollars – we could wipe out all our slums in ten years, reverse decay in the great, dull, gray belts that were yesterday's and day-before-yesterday's suburbs, anchor the wandering middle class and its wandering tax money, and perhaps even solve the traffic problem.
But look what we have built with the first several billions...

Jane Jacobs (1961)

This quote reflects a common misconception about the current state of urban affairs. There are many who believe that most urban problems are a direct result of a lack of funding. These same people then believe that if somehow funding could be significantly increased, cities' problems would simply disappear. While there may be some truth to these ideas, the reality is that the in most cases current governance would not be able to effectively use this money to cure the ills that are afflicting today's cities.

1.1 Purpose and Objectives

Most cities today are not developing in a sustainable manner. The definition of sustainability varies from source to source. The World Bank (2000) defines sustainability as "Development that meets the needs of the people today without compromising the ability of future generations to meet their own needs." Richardson (1999) provides a more specific definition of sustainable transportation, defining it as a system "...in which fuel consumption, vehicle emissions, safety, congestion, and social and economic access are of such levels that

they can be sustained into the indefinite future without causing great or irreparable harm to future generations of people throughout the world."

This study is interested in evaluating the role of urban governance in achieving sustainability. The study ties together findings from many diverse fields to create a framework for an ideal urban governance structure. Once this framework has been explained, it is used to evaluate the case study of Vancouver BC. This paper is structured into 4 main sections. The first section presents the pyramid evaluation framework. The second section describes the Vancouver governance case study. The third section uses the pyramid framework to evaluate Vancouver's governance model, and the final sections ties together the paper's findings for some final conclusions.

2 EVALUATION FRAMEWORK

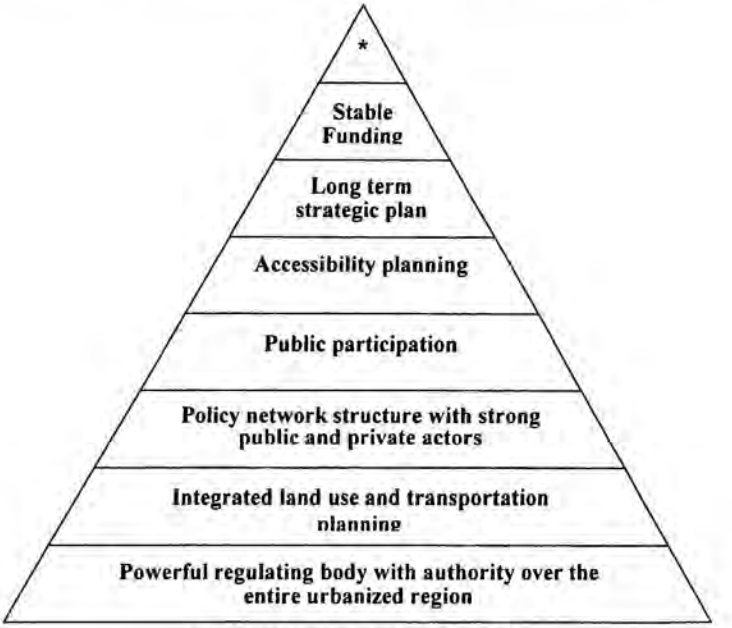
Four major criteria for sustainable urban governance were established and discussed in the study by Coleman (2004). These criteria are governance structure, regulation philosophy, public participation, and sustainable funding. Governance structure deals with how governments should be structured (i.e. horizontal vs. vertical structures), and uses policy networks to study relationships between important actors. The regulation philosophy discusses the need for a focus on accessibility planning, and the need to integrate transportation and land use planning activities. The public participation criteria argue that a high level of public participation is necessary, but that this participation shouldn't be taken so far as to overrule important expert viewpoints. The final criterion is the need for stable funding sources while exploring the role of the private sector in funding transportation. (Coleman 2004)

The focus of this section of the paper is to establish how each of these elements and criteria are related to each other and how they must be implemented to achieve the ultimate goal of sustainable transportation. A general representation of how the criteria are related is presented, titled the pyramid of urban governance for sustainability. This pyramid framework will then be used to evaluate the case of Greater Vancouver later in this study.

2.1 *Pyramid of Urban Governance for Sustainability*

A basic representation of how the different aspects of an ideal urban governance system interact is depicted in the *pyramid of urban governance for sustainability* in Figure 1. The pyramid representation is used as it not only shows what elements are necessary but what is the logical order needed to build an ideal governance structure. The elements below any given level can be though of as the building blocks needed to achieve that given level. Each level of the pyramid will be described in detail.

Figure 1 - Pyramid of urban governance for sustainability



1. Powerful regulating body with authority over the entire urbanized region: The basis of the ideal governance model is the regulating body that oversees the decision making process. This organization is where all decisions begin and end. For this reason, the base of the pyramid is a powerful regulating body with authority over the entire urbanized region. This first basic level is important because if the regulating body only has power over a part of the urbanized region, the structure will essentially be inadequate right from the start. The body must be powerful, meaning that it must have the authority and means to implement decisions concerning all modes of travel.

2. Integrated land use and transportation planning: The second level of the pyramid deals with the integration of land use and transportation planning. Once the regional body has been established, it must be given the mandate to plan both land use and transportation functions across its jurisdiction. This integration is key in making effective transportation and land use decisions. It is simply not enough to have a minimal degree of coordination within two separate regulating bodies, (a transportation body and a land use planning body). In reality, transportation and land use are two intrinsically interdependent functions, and the governance structure should be organized to reflect this fact

3. Policy network structure with strong public and private actors: The next step in building the governance structure is to establish the relationships between governmental actors, private sector actors and the public. These relationships take the form of a policy network that governs how decisions are made. The important characteristics of this policy network are that the governmental actors must have a strong role to be able to dictate a mandate of sustainability, and that the network must incorporate private sector and community actors.

4. Public participation: Once the regulating body and policy network decision making structure are in place, the public participation programs can be established so that the public can have en effective voice within the decision network. These programs must ensure that the public voice is heard and that the process does not end in tokenism. This successful result is guaranteed partly by the structure of the policy network in the precious step.

5. ***Accessibility planning:*** At this point in the pyramid, the governance structure has essentially been determined. What remains is using the organizational and decision making structure as a way of making and implementing smart planning decisions. For the goal of sustainability, it is important to focus planning efforts on providing accessibility and not mobility. This focus will lead to initiatives such as travel demand management and other effective planning initiatives.

6. ***Long term strategic plan:*** The second last level in the pyramid is the creation of a strategic long term plan. This plan outlines the type of development that needs to take place and the transportation strategies that will be used. This plan can now be developed since the regulating body is in place, the decision making structure is functioning properly, the public has been involved in the process and the correct planning philosophy has been adopted.

7. ***Stable funding:*** With the governance structure in place and with a plan outlining how to reach the ultimate goal of sustainability, all that is needed is the funding to implement all of the desired strategies. This funding strategy not only provides the means to build physical infrastructure, but also aids to ensure that the entire population will eventually have equal accessibility for all travel within the urban region. This fact that funding finds itself at the top of the pyramid ties back into the Jane Jacobs quote at the beginning of the paper.

3 GOVERNANCE MODEL: VANCOUVER, BC

This section of the study describes the current urban governance model of the Greater Vancouver Area. The case study is chosen with a goal to understand how the governance model functions and to identify the major problems and successes of urban governance in Vancouver.

For the purposes of this case study, Vancouver will be defined as the Greater Vancouver Regional District (GVRD), which is a federation, comprised of twenty municipalities and two electoral areas. The GVRD has a population of approximately 2.2 million and is one of the fastest growing metropolitan areas in North America with growth of over 400,000 people over the past decade (GVTA 2003). The transportation and land use planning responsibilities have been split into two distinct organizations: the GVRD, and the Greater

Vancouver Transportation Authority (GVTA), also known as TransLink. The GVRD provides region-wide functions such as strategic land-use planning, water and sewage, housing, garbage and recycling, and regional parks. The GVTA's mandate is to plan, develop, and finance transportation facilities and services for daily travel within and between municipalities in the GVRD.

3.1.1 Overview of Regional Governance Structure

There are six major groups of actors that have a role in transportation governance in the Greater Vancouver Area. These actors are the provincial government, the municipalities, the GVRD, the GVTA, the private sector, and the public. The major roles of each actor group are summarized below. The GVTA will be discussed in detail following this section.

- **B.C. Provincial Government**: The BC government has traditionally played a very active role in the provision of public transit in Vancouver. Until the formation of the GVTA in 1999, all transit service was provided by BC Transit, a provincial crown corporation. Since 1999, the BC government's role in public transit has diminished. There are three provincially appointed members on the GVTA board. The province still pays a large share of most capital investments in public transit. Additionally, the BC Provincial Government is still responsible for all provincial highways and most major bridges.
- **Greater Vancouver Regional District**: The GVRD is made up of 35 directors representing the member municipalities. These directors are municipal mayors and city councilors that have been appointed to the GVRD board by their respective councils (GVRD 2004). The GVRD has a strong impact on transportation governance through its land use planning functions. The *Livable Regions Strategic Plan* (LRSP), a planning document produced by the GVRD, is legislated as the directive for all transportation decisions taken by the GVTA.
- **Municipalities**: The primary role of the municipalities in the transportation governance process is to nominate members of city council to the GVRD. Once this responsibility is fulfilled, the municipalities' role is limited. The municipalities are

involved to a minor degree in the creation of official plans, both by the GVRD and the GVTA.

- **Public**: The citizens of Vancouver have limited input into the transportation governance process. Their input is limited to electing the municipal and provincial politicians who will eventually make their way onto the boards of the GVRD and the GVTA. The public has no say into who is appointed to these organizations, though a low level of accountability is maintained by ensuring that the members of the GVRD and GVTA boards are elected officials. Another avenue where the public has input is in the creation of the GVRD and GVTA strategic plans. The public is consulted extensively during the formation of these plans. The main issues is that there is no form of continuous public participation in Vancouver, and as a result the public only has a voice approximately every four years when new politicians are elected, and new strategic plans are formed.
- **Private Sector**: The private sector has no clearly defined role in the transportation governance process. There are no formally defined communication channels between private actors and either the GVRD or the GVTA. More informal relationships, however, are present, and the GVTA has been looking increasingly towards private partnerships to fund capital improvements and to operate transit services

3.1.2 Greater Vancouver Transportation Authority

The GVTA was created in 1999 with five key policy responsibilities (Meligrana 1999):

1. Develop and finance regional transportation in the GVRD
2. Coordinate and deliver transit services in the GVRD
3. Support the Major Road Network (MRN)
4. Manage and develop new Transportation Demand Management (TDM) programs
5. Administer AirCare within the GVRD

As is shown by the scope of the GVTA's mandate, the GVTA has comprehensive authority over the transportation system and has the capability to manage both supply and demand. Prior to the formation of the GVTA, there was very little coordination of transportation

planning activities with the GVRD. Most transportation decision making authority was in the hands of BC Transit, a provincial body. In addition, municipalities often had very little interest in owning and operating their own transit systems. As a result of these issues Vancouver had never previously had an accountable, effective and responsive regional transit agency (Meligrana 1999).

Currently, the GVTA functions independently from the province and from the GVRD, but the programs and policies that the GVTA implements must conform to the GVRD Liveable Region Strategic Plan (LRSP). This relationship with the GVRD can be seen in Figure 2. The GVTA is an entity comprised of 15 representatives, 12 of whom are mayors or GVRD directors, while the final three are provincially appointed Member of the Legislative Assembly (MLAs) (Meligrana 1999).

One of the most important components of the GVTA's overall function is the transit system. The GVTA manages the entire transit system ranging from commuter rail service to local bus systems. The GVTA is structured to separate policy and planning from the operation and transit service delivery. The GVTA (or TransLink) essentially purchases the services, while subsidiaries provide the service. The GVTA has been looking increasingly into partnerships with the private sector to operate services and to fund the expansion of the transit system.

A final important element that should be considered is the GVTA's authority to implement needed strategies. The GVTA does in fact have the legislative authority to implement strategies that conform to the GVTA Strategic Transportation Plan. The authority's powers include the ability to generate and manage funds, as well as the ability to implement policies and plans without public consultation. As a result, the GVTA has a wide range of dedicated funding sources, as well as the ability to pursue new sources of funding. These funding sources are outlined in Table 1. Although these funding sources seem quite comprehensive, the lone funding source that is the most unpredictable also one of the largest sources: grants from the provincial government. There have been examples in the past of political meddling on the part of the provincial government to

influence transportation decisions. An example of this effect is the choice of technologies for the extension of the SkyTrain line. The decision to implement the existing technology was made by the province with no approval by the GVTA (Kar 2001).

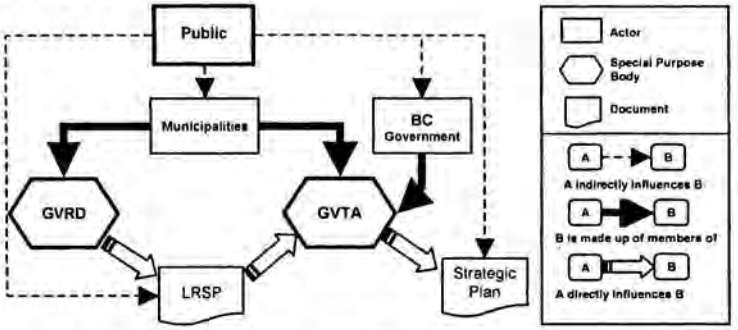
Table 1 - GVTA Funding Sources, adapted from (Kar 2001)

Current GVTA Funding Sources	*Possible Additional Sources*
▪ Transit fares ▪ Gasoline tax (8 cents/liter, increasing to 10 cents/liter in 2005) ▪ BC Hydro levy ▪ Property Taxes (Residential, commercial and industrial) ▪ Parking sales tax (A transfer of the provincial sales tax on parking, paid as a fixed value of $10 million) ▪ AirCare test fees (User fees from auto emissions test program)	▪ Cost-recovery tolls on improvements to the major road network ▪ Vehicle registration charges ($75/vehicle) ▪ Additional local tax on parking lots ▪ Benefiting property tax (A tax on properties that benefit from rapid transit)

3.1.3 Observations and Conclusions

Figure 2 shows a representation of the basic transportation governance relationships in the Greater Vancouver area. Upon study of the roles of each of the actors, there are two key observations that emerge. The first is that the general public is largely excluded from the transportation governance process. The public can influence the process indirectly at best, and is generally separated from the actual decision making processes (GVRD and GVTA) by at least two degrees. The second important observation is that there is an organizational disconnect that exists between the land use planning functions and the transportation planning functions in Vancouver. This can be seen in the diagram as there is no direct link between the GVTA and GVRD, a fundamental weakness of the structure.

Figure 2 - Greater Vancouver Area: Current relationships in transportation governance



There are five important conclusions that can be drawn from the Vancouver case study:

1. Vancouver is widely regarded as an excellent governance model of region-wide planning. The fact that all planning activities are made for the whole urbanized area immediately sets the GVRD apart from many other urban areas
2. The GVTA funding system is quite strong, and should result in a stable flow of funds. This funding will make it possible to implement needed system improvements to move towards greater sustainability.
3. One glaring deficiency implicit in the Greater Vancouver governance structure is the lack of opportunity for public participation. Both the GVRD and the GVTA operate several steps removed from the public with little decision making accountability. This is one of the largest weaknesses of the Vancouver model.
4. A second deficiency is the lack of coordination between land use planning and transportation planning. The division of functions between the two entities leads to the possibility of contradictory or uncoordinated policy-making. This disconnect between the two bodies is only worsened by a lack of communication lines between the organizations. Transportation decisions should be made in a coordinated fashion. Instead, the GVRD takes the lead

in decision-making and is then followed by the GVTA. The GVRD does not necessarily have all the expertise necessary to ensure that the correct plans are being implemented as a large part of the transportation expertise lies within the GVTA.

5. A final concern is the lack of regulation defining what roles the private sector can and cannot play in the current system. Meligrana (1999) argues that this lack of regulation could result in the GVTA's regional role in transit being eroded through the ability of the GVTA to approve independent transit services.

4 EVALUATION OF VANCOUVER'S GOVERNANCE

Now that the desired urban governance framework for sustainability has been established, and the Vancouver Case study has been outlined, the pyramid framework will be used to analyze the case of Vancouver's urban governance system.

4.1.1 Pyramid Evaluation

1. Powerful regulating body with authority over the entire urbanized region: Vancouver has established the GVRD to manage land use and other services across the jurisdiction. The GVRD meets the requirements of the first step as it has been granted sufficient decision making power and the jurisdiction does cover the entire urbanized area. The GVTA similarly has authority over the whole urbanized area.

2. Integrated land use and transportation planning: This area is a weakness for Vancouver. Instead of integrating transportation functions within the GVRD, a separate body was created responsible for all transportation functions. To make matters worse, there is very poor coordination between the two groups. The only coordination is from the GVTA following the mandate set by the GVRD's Livable Region Strategic Plan (LRSP).

3. Policy network structure with strong public and private actors: It is unclear to what extent this type of decision making model is followed in the Vancouver case study. The one obvious deficiency is the private sector involvement in the decision making process. This relationship with the private sector is still in the formative stages, and is poorly defined. The governmental actors do currently have the power to enforce a mandate of sustainability.

*4. **Public participation:*** Public participation is another of the largest weaknesses in Vancouver. The voice given to the public is minimal at best. The problem lies in the separation between the public and the decision making bodies. As it currently stands, the public can only influence decision making through elected officials and through inputs into the creation of official strategic plans for both the GVRD and the GVTA.

*5. **Accessibility planning:*** Both the GVRD and GVTA do have some limited focus on accessibility planning (GVRD 2004). Part of the mandate of the GVTA is to provide a Transportation Demand Management (TDM) to encourage more sustainable travel patterns.

*6. **Long term strategic plan:*** Both the GVTA and GVRD produce long term strategic plans as the basis of their decision making. These plans are generally carried out on ten or twenty year horizons (GVRD 2004, TransLink 2004)

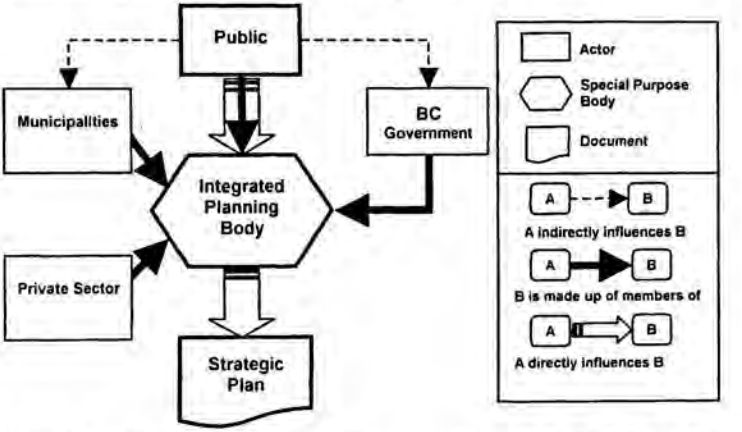
*7. **Stable funding:*** One weakness of both the GVRD and the GVTA is that they both lack complete autonomy from the provincial government and depend on the BC government for subsidies. Due to this fact, decision making is often prone to political meddling from upper levels of government. Other than this problem, Vancouver does have a stable funding system, with revenue sources including the general tax base, fuel taxes, parking fees, and transit fares. Though the funding is substantial and stable, it would still not be considered equitable as it heavily favours automobile travel.

4.1.2 Recommendations

Performing a pyramid evaluation of the Vancouver governance structure illustrates several interesting points. While the current structure does not conform to many of the requirements of the pyramid of urban governance, Vancouver is actually in good shape when compared to many other urban regions. This state of affairs is due mostly to the fact that Vancouver does have a regional body with a significant amount of autonomy and authority. There are problems with the integration of land use and transportation planning as well as problems with the public participation process, but the reality is that the foundation is in place to piece together the rest of the requirements to form the ideal governance structure. Figure 3 presents an illustration of the ideal governance structure for the

Vancouver region. Contrasting this diagram with the existing relationships depicted in Figure 2 shows much greater simplicity and more clearly defined channels. Where before the public was several degrees removed from the decision making process, the public now directly influences the decision making process and is given seats on the integrated planning body council. In general, all actors have become more directly involved in the governance process, leading to much more efficient governance.

Figure 3 - Greater Vancouver Area: Ideal governance relationships for sustainable development



The following five recommendations can be made based on the pyramid analysis to further improve transportation governance:

1. Integrate the GVRD and GVTA into one integrated planning body. This would not be a trivial process, so if necessary in the interim a suitable compromise would be to significantly improve the communication channels between the organizations and to promote more coordinated decision making.
2. Provide an explicit role for private sector actors, while involving these actors into the decision making process
3. A restructuring of governance is necessary to give the public a more direct voice in the governance process. Programs should

be implemented to insure that participation is effective and that transit users can have a direct impact in the decision making process.

4. Implement a more explicit focus on accessibility planning
5. Introduce new funding mechanisms to charge the actual marginal costs for transportation service. This will improve the overall equity of the transportation system, while reducing the dependence on central government subsidies.

5 CONCLUSIONS

Through a literature review of diverse sources and through a study of the Greater Vancouver Area, the following conclusions have been established. This study argues that governance structures are weakened by the long-standing separation of transportation and land-use planning and by the lack of regional coordination of transportation and land use activities. This weakness becomes exacerbated by the growing size of metropolitan areas and by the complex relationships with higher levels of government. Reforms to these structures are needed to permit the coordination of transportation and land-use planning and to pave the way to sustainable development. In addition, governance must provide a means to involve and to draw upon the knowledge and expertise of the community, other government agencies and the private sector. Although some systems of governance have incorporated private sector actors into the negotiation process, few have paid sufficient attention to the users in the general public. Planners need to draw on the knowledge and expertise of the transportation system users and to involve them in policy-making. Such a step would not only lead to better policy, but also increase the legitimacy and support of the general public for sustainable policies.

This paper opens the door to many possible avenues for future research. One such possibility would be to continue with the pyramid framework and to attempt to apply it to other cities. In addition, the recommendations for Vancouver can be looked at in more detail in an attempt to establish the more detailed steps that are needed to implement the pyramid based governance structure.

6 REFERENCES

Coleman, Jesse W.R. 2004. An Urban Governance Model to Achieve Sustainable Urban Transportation and Development. BASc Thesis, University of Toronto

Greater Vancouver Regional District. 2004. Greater Vancouver Regional District homepage. 3 Mar. 2004 <http://www.gvrd.bc.ca>.

Greater Vancouver Transportation Authority. 2003. 2005-2007 Three-Year Plan & 10-Year Outlook: Strategic transportation plan amendment. GVTA.

Jacobs, Jane. 1961. The death and life of great American cities. Toronto: Random House of Canada Ltd.

Kar, Dennis A. 2001. Metropolitan transportation agencies: Three case studies of multi-jurisdictional transportation planning in Canada. Paper presented at the annual Bus and Paratransit Conference, May 5-10, in Calgary, Canada.

Meligrana, John F. 1999. Toward regional transportation governance: A case study of Greater Vancouver. Transportation. 26: 359-380.

Richardson, Barbara. 1999. Towards a policy on a sustainable transportation system. Transportation Research Record 1670, 27-34.

TransLink. 2004. Greater Vancouver Transportation Authority (TransLink) homepage. 3 Mar. 2004 <http://www.translink.bc.ca>.

World Bank, Development and Education Program. 2000. Glossary. Washington, D.C <http://www.worldbank.org/html/schools/glossary.htm> (accessed April 6, 2004).

AUTHOR INDEX

Alam, M.J.M., 748
Allaby, Peter, 611
Anderson, Darryl, 417

Bahramian, Hamidreza, 673
Baldwin, Gordon, 44
Bazant, Lubor, 115
Bedi, Nilam, 158
Bigras, Yvon, 55
Boardman, A. E., 356
Boucher, Michel, 495
Buliung, Ron N., 555
Byelyayev, Oleg, 341

Cairns, Malcolm, 323
Coleman, Jesse, W.R., 778
Corredor, Alfonso, 130

Davies, Philip, 29
Dawson, Keith, 400
Debrezion, Ghebreegziabiher, 449
Dubé, Vincent, 70

Fan, L., 597
Ferdous, Farhana, 14

Gillen, David, 356
Groskopf, Karl, 130

Habib, Muhammad Ahsanul, 14
Haider, Murtaza, 571
Haughton, Michael A., 585
Hellinga, Bruce, 385
Hickson, Allister, 662

Hoque, Md. Mazharul, 748

Kanaroglou, Pavlos S., 1, 201, 209, 224, 239, 647, 718
Kang, Hejun, 718
Karim, Dewan Masud, 100
Kariya, Michael T, 763
Khondaker, Bidoura, 748
Krause, David, 253, 270
Kube, Douglas, 34
Kubursi, Atif, 158

LeBlanc, Louis A., 172
Lynn, Jill, 634

Madill, Ron, 130
Maniruzzaman, K. M., 14
Maoh, Hanna, 1, 239, 718
McLeod, Murray, 733
Mercado, Ruben, 209
Mocek, Hyneck, 115
Monteiro, Joseph, 253, 270, 400, 432

Newbold, K. Bruce, 201, 209, 224, 239
Nix, Sean, 625
Noh, Jaehyun, 385
Normandeau, Guy, 55

Ogwude, Innocent C., 703

Páez, Antonio, 201, 209, 224, 239
Park, Peter (Young-Jin), 85
Pels, Eric, 449
Pilon, Denis, 70
Pooryari, Maghsood, 673
Potoglou, Dimitris, 224, 647
Prokop, Darren, 286

Prentice, Barry, 480, 597

Raicu, Raluca, 465
Rezaee-Arjroody, Abdolreza, 687
Rietveld, Piet, 449
Robertson, Gerald, 400

Sabourin, Marcel, 540
Saccomanno, Frank, 385
Saffrzadech, Mahmoud, 673, 687
Scott, Darren M., 201, 209, 224, 239, 718
Scrafton, Derek, 465
Sokol, Christy, 662
Spurr, Timothy, 571
Stephens, Mobolaji S., 703

Taylor, Michael, 465
Trudel, Mathieu, 144

von Wartburg, Markus, 371

Waters II, W.G., 356, 371
Wright, Christopher, 186

Yevdokimov, Yuri V., 341
Yuen, Andrew, 516

Zhang, Anming, 356, 516

APPENDIX I

STATEMENT OF PURPOSE

The CTRF is
a broadly based network of transportation professionals with an interest in research and commentary on public policy issues. Membership is open to anyone in any aspect of transportation.

The CTRF's objectives are
- to encourage transportation research in those disciplines which address themselves to the economic, managerial, technical, political and social aspects of transportation
- to provide a common meeting ground to exchange ideas and research results
- to provide a forum for presenting and discussion transportation issues
- to bring together persons involved with Canada's transportation systems

The CTRF's program consists of
- a national conference with research papers, panel discussions, and invited speakers held annually in a different Canadian location
- published proceedings of the national conference
- a one-day semi-annual conference on a topic of national interest
- regional seminars promoted and co-sponsored with other organizations
- an annual CTRF student paper competition
- an annual CTRF scholarship competition
- a newsletter published five times a year

Le GRTC est ...
un groupe de professionnels(les) en transport intéressés(ees) à la recherche et aux questions de politiques publiques. Etre membre du GRTC convient à quiconque s'intéresse aux transports.

Les objectifs de GRTC sont ...
- d'encourager la recherche dans les domaines suivants liés aux transports: économie, gestion, technologie, politique et aspect social
- de fournir un lieu de rencontre pour échanger idées et résultats de recherche
- de créer un forum pour présenter et discuter de questions de transport
- de réunir les personnes intéressées au système canadien de transport.

Au programme du GRTC ...
- une conférence nationale annuelle avec site variable au Canada, qui consiste en une présentation de textes, tables rondes, conférenciers invités
- la publication des textes présentés à la conférence nationale
- une conférence semi-annuelle d'une journée sur un juget d'interet national
- la promotion et le parrainage de colloques régionauz avec d'autres organismes
- un concours annuel de textes d'étudiants(es)
- un fonds permettant de décerner des bourses
- un bulletin de nouvelles publié cinq fois l'an

HISTORY OF CTRF

The desire to form a transportation research group began in the early 1960s. In 1965, the possibility of forming such an organization met with significant positive response. John Baldwin, then Deputy Minister of Transport, indicated the Department would provide financial assistance to back the CTRF formation and the CTRF was born in 1965.

The first meeting of CTRF was held at Manoir St. Castain in Lac Beauport, Quebec on September 9, 1965. The CTRF began with a membership of 120. Throughout its history, CTRF has published a newsletter issued five times a year today. In 1983, the CTRF began printing its papers and publishing them in a bound volume.

In 1974, the annual Paper Competition was initiated for students attending Canadian universities in three categories. In 1976, CTRF awarded an additional prize for the best overall paper under the auspices of the Jim Davey Memorial Trust Award. In 1988, CTRF established a Scholarship Fund to encourage university students to pursue studies leading to careers in transportation.

CANADIAN TRANSPORTATION RESEARCH FORUM
LE GROUPE DE RECHERCHES SUR LES TRANSPORTS AU CANADA

EXECUTIVE-EXECUTIF
2004-2005

President
Gordon Baldwin
Statistics Canada

Executive Vice President
James Nolan
University of Saskatchewan

Secretary
Gail Sparks
Sparks Enterprises Ltd.

Vice President Finance & Treasurer
Gerry Kolatis
Via Rail Canada

Vice President External
Gordon Tufts
Manitoba Transportation and Gov. Services

Vice President Meetings
Michael Kieran
IBI Group

Vice President Programs
Bill Raney
Ministry of Transportation Ontario

Vice President Awards
Stephen Ellsworth
Transport Canada

Immediate Past President
Clarence Woudsma
University of Waterloo

VP Organization and Development
Gerry Kolatis
Via Rail Canada

Vice President Sponsorships

Councillors:

Lloyd Ash
Mount Royal College

Malcolm Cairns
Canadian Pacific Railway

Gregg Corcoran
Canadian National Railways

Paul Earl
University of Manitoba

Murtaza Haider
McGill University

Doug Johnson
PD & IB, NB Transportation

Joseph Monteiro
Industry Canada

Garland Chow
Centre of Transportation Studies
UBC

Paul Ouimet
Intervistas Consulting Inc.

Paul Larouche
Bombardier Transportation

Rachid Raffa
Quebec Ministry of Transport

CTRF PAST PRESIDENTS / PRÉSIDENTS ANTERIEURS DU GRTC

2003-04 Clarence Woudsma
2002-03 Gerry Kolatis
2001-02 Eric Hilderbrand
2000-01 Gordon Tufts
1999-00 Lloyd Ash
1998-99 Janet Weiss
1997-98 Barry Prentice
1996-97 George Clark
1995-96 Doug Duncan
1994-95 Roger Roy
1993-94 Bill Raney
1992-93 Terry Phillips
1991-92 Andrew Elliott
1990-91 Mary Brooks
1989-90 Eugene Ludwick
1988-89 William Waters
1987-88 Craig Dickson
1986-87 Rod Taylor
1985-86 Frank Trotter
1984-85 Karen Hecks
1983-84 Ray Rohr
1982-83 Neal Irwin
1981-82 Jim Telford
1980-81 Albert Stevens
1979-80 Merrienne Duncan
1978-79 Al Bach (deceased)
1977-78 Vic Stechishin (deceased)
1976-77 Gordon Wales
1975-76 Bill Hughes
1974-75 Norman Lee
1973-74 Michel Chevalier
1972-73 Howard Mann
1971-72 John Gratwick
1970-71 Richard Soberman
1969-70 John Eyre
1968-69 Ron Wilson
1967-68 Harvey Romoff
1966-67 Pierre Camu
1965-66 Ray Cope

CTRF LIFETIME MEMBERS / MEMBRES À VIE DU GRTC

R. L. Banks (deceased)
Frank Collins
Ray Cope
Craig Dickson
John Gratwick
John Heads
W.B. Hunter (deceased)
Joe Howard
Neal A. Irwin
Donald MacKinnon
Alan Dick Mcleod
Nick Mulder
H. M. Romoff
Richard Soberman
Albert Stevens
William Waters

2005 ANNUAL CONFERENCE SPONSORS / COMMANDITAIRES DE LA CONFÉRENCE 2005

Bison Transport
Bombardier Transportation
CAD Railway Services Inc.
Canadian Freightways Limited
Canadian Transportation Agency / Office des Transports du Canada
Canadian Pacific Railway / Canadien Pacifique
City of Hamilton
Hamilton Airport Authority
Hamilton Street Railway
IBI Group
Industry Canada / Industrie Canada
LDK Engineering Inc.
Logistics Solution Builders Inc.
McMaster University, Centre for Spatial Analysis
Ministry of Transportation Ontario
Statistics Canada / Statistique Canada
Taylor Raynauld Amar & Associates Inc.
Transport Canada / Transports Canada
TTX Company
VIA Rail

This list was updated April 1, 2005. Additional sponsors will be acknowledged in the Forumation, on the CTRF Website, and at the conference.

2004 SEMI-ANNUAL CONFERENCE SPONSORS / COMMANDITAIRES DE LA CONFÉRENCE 2004

Transport Institute, University of Manitoba

SCHOLARSHIP PROGRAM AND AWARDS
DONORS TO THE CTRF SCHOLARSHIP PROGRAM
PARRAINS AU PROGRAMME DE BOURSES DU GRTC
2005-2006

Bombardier
Canadian National Railway
Canadian Pacific Railway
Canadian Transportation Research Forum
Government of Québec, Ministry of Transport
Transport Canada
Individual Donors (numerous)

SCHOLARSHIP WINNERS 2004-2005

Bombardier Scholarship
Marc Fournier, Ph.D. Engineering, l'INSA, Lyon and École de technologies supeireure (ETS)

Canadian Pacific Railway Scholarship
Wesley Andreas, MA Geography, University of Calgary

CN Scholarship
Leanne Whiteley, MASc. Civil Engineering, University of Waterloo

Transportation Industry Scholarship
Oleg Byelyayev, Msc.E. Transportation Group, University of New Brunswick

Transport Canada Scholarship in Sustainable Transportation
Simon Véronneau, Ph.D. Production and Operations Management, HEC Montreal

Transport Canada Scholarship in Safety and Security
Kris Thiele, Ph.D. Clinical Psychology, Queen's University

Transport Canada Scholarship in Integration and Intermodalism
Jinha Jeong, MS Transportation and Logistics, University of B.C.

2003-2004 STUDENT PAPER COMPETITION DONORS
PARRAINS DU CONCOURS DE ÉTUDIANTS EN 2003-2004

Albert Stevens Fund
Jim Davey Trust Fund

STUDENT PAPER WINNERS 2003-2004

Undergraduate Category
First Prize - Past Presidents Award ($1,000)
Michael Kariya, Faculty of Law, University of Calgary
"Aircraft noise: A survey of the regulatory framework and enforcement mechanisms in Canada"

Second Prize ($500)
Dan Freeman, Arts & Science, McMaster University
"The promise of the hydrogen fuel cell vehicle"

Third Prize ($250)
Jesse W.R. Coleman, Department of Civil Engineering, University of Toronto
"An urban governance model to achieve sustainable urban transportation and development"

Masters Category
First Prize - Albert Stevens Award ($1,000)
Timothy Spurr, Faculty of Civil Engineering and Applied Mechanics, McGill University
"The development of accessibility indicators for the greater Montreal region"

Doctoral Category
First Prize - Jim Davey Award($1,000)
Dimitris Potoglou, School of Geography and Geology, McMaster University
"Predictive mapping of vehicle-emitted carbon monoxide in urban areas - The case of Hamilton Census Metropolitan Area, Canada"

CANADIAN TRANSPORTATION RESEARCH FORUM
209 - 15 Innovation Blvd. Saskatoon, Saskatchewan
S7N 2X8
www.ctrf.ca
Telephone: (306) 242-6199 Fax: (306) 242-6199

INVOICE FOR 2005 CTRF MEMBERSHIP

CTRF Membership for Calendar Year 2005 ($75.00)	$ _____
GST at 7% ($5.25)	$ _____
Student Membership for Calendar Year 2005 ($20.00)	$ _____
GST @ 7% ($1.40)	$ _____
TOTAL	$ _____

GST Registration No. R126274562 (If GST exempt, please indicate)

Please make cheque payable to Canadian Transportation Research Forum and mail to the address noted above. **PLEASE RETURN A COPY OF THIS INVOICE WITH PAYMENT IN ORDER THAT OUR DATABASE WILL CONTAIN ACCURATE INFORMATION RELATING TO YOUR MEMBERSHIP.**
VISA and MasterCard are also acceptable.

___ **VISA**

___ **MasterCard**

Card Number _____
Expiry Date _____
Name on Card _____

Please complete the following to ensure accurate membership data. The information below will appear in the next published CTRF/TRF Directory.

Name:_____ Title:_____

Company:_____

Address:_____

_____

_____

Telephone: _____ Fax:_____

e.mail:_____

Please indicate your language preference: ☐ English ☐ French

LE GROUPE DE RECHERCHES SUR LES TRANSPORTS AU CANADA
209 - 15 Innovation Blvd. Saskatoon, Saskatchewan
S7N 2X8
www.ctrf.ca
Téléphone: (306) 242-6199 Télécopieur: (306) 242-6199

FACTURE 2005

Membre du GRTC pour l'année 2005 (75.00 $)	$ _____
TPS (5.25$)	$ _____
Membre étudiant du GRTC pour l'année 2005 (20.00$)	$ _____
TPS (1.40$)	$ _____
TOTAL	$ _____

Numéro d'enregistrement pour la TPS R126274562. (Si vous êtes exempté de la TPS, veuillez l'indiquer.)

Veuillez libeller votre chèque au nom du Groupe de Recherches sur les Transports au Canada et poster le tout a l'adresse ci-haut. **SVP RETOURNEZ UNE COPIE DE CETTE FACTURE AVEC VOTRE PAIEMENT AFIN QUE NOTRE BASE DE DONNÉES CONTIENNE UNE INFORMATION EXACTE SUR VOUS COMME MEMBRE DE L'ASSOCIATION.**
VISA et MasterCard sont acceptes.

____VISA
____ MasterCard
Numéro de Carte _____
Date d'Expiration _____

Nom Figurant sur la Carte _____

fournir les renseignements demandée ci-dessous afin de nous permettre d'avoir une information complète sur nos membres. Cette information va apparaitre dans le prochain répertoire du GRTC.

Nom: _____ Titre: _____
Compagnie: _____
Adresse: _____
_____
_____ Téléphone: _____ Télécopieur: _____
Courrier électronique _____
Veuillez indiquer la langue de correspondance de votre choix.
☐ anglais ☐ français

ANNOUNCEMENT / INFORMATION

- 2005 CTRF SEMI-ANNUAL Dates and Location to be Announced

- 2006 CTRF ANNUAL CONFERENCE TO BE HELD IN **QUEBEC CITY, QUEBEC**.

WATCH FOR CALL FOR PAPERS COMING IN SEPTEMBER 2005

For further information, contact:

Gail Sparks
(306) 242-6199
e-mail: gail.sparks@ctrf.ca

CTRF website: www.ctrf.ca

ISSN #1183-2770
ISBN 0-9737440-0-6

Cover images: Copyright © 2005 Canon Inc.
http://web.canon.jp/Imaging/eosdigital2/eosdigital2_sample-e.html

Printing Services • 966-6639
University of Saskatchewan