CTRF

Transportation: Emerging Realities



Les Transports: realités en puissance

VOLUME 2

Canadian Transportation Research Forum
Proceedings of the 32nd Annual Conference
Toronto, Ontario May 25-28, 1997

Le Groupe de Recherches sur les Transports au Canada
Actes de la 32ième conférence annuelle
Toronto, Ontario 25 au 28 mai, 1997

VOLUME II

Volume I

Opening and Plenary Session/Séance Plénièré d'ouverture

Session/Séance 102: *Keynote Speaker*

- *Demographic Change in Canada: Implications for Transportation.* Thomas W. M°Cormack..... 1

Session/Séance 141: *Urban I*

- *How Might Canadian Experience Contribute to the Transition of the Russian Urban Transportation?* Elizaveta B. Bedniakova..... 17
- *Synopsis - Sustainability in Urban Transportation.* Richard M. Soberman..... 41
- *Transportation Control Measures and the Urban Environment: A Comparative Review.* William P. Anderson and Clarence Woudsma. 43

Session/Séance 143: *Passenger Issues I - VIA Rail (translation available)*

- *La Nouvelle Loi sur les Transports: Ses Conséquences sur les chemins de fer d'intérnêt local et sur la société VIA Rail.* Louis Doyle..... 58
- *Revenue Management at VIA Rail Canada.* James Berwick and Carole Therrien..... 70
- *Two-dimensionally Nested Inventory.* Peter Borowiak..... 85

Session/Séance 151: *Urban II (Local Experience)*

- *Focusing Public Consultation in Urban Transportation Planning: Winnipeg's Transplan 2010 Process.* David Hicks and Richard S. Tebinka and Dr. Barry E. Prentice..... 102

- *The Calgary Diesel LRT/Commuter Rail Experiment.* Kells Boland. . 119

Session/Séance 152: *Highway I*

- *The Changing Relative Important of International and Interprovincial Merchandise Trade in Canada with Particular Reference to Ontario.* Frank Sinanan and Kevin Junor. 129

- *Truck Traffic Between Canada and the United States.* Fred P. Nix and John R. Billing and Michèle Delaquis. 150

Session/Séance 153: *Air I (translation available)*

- *Mesure de Prix et de Production dans L'industrie du Transport Aérien au Canada.* Raynald Ouellet. 165

- *The Effect of Liberalized Air Transport Bilaterals on Foreign Visitor Volume and Traffic Diversion: The Case of Canada.* Martin Dressner and Tae Hoon Oum. 180

- *The Toronto and Montréal Aviation Markets - A Tale of Two Cities.* Gordon G. Baldwin. 211

Session/Séance 211: *Regional I*

- *Evaluation of the New Jersey I-80 HOV Lane.* Nicholas C. Antoniou and Hong Lin and Louis J. Pignataro. 222

- *Emergence of the Modern Roundabout as a Reality in Vermont and its Relation to Vermont Urban Design and Development.* Tony Redington. 237

Session/Séance 212: *Highway II*

- *A Rough Sets Determination of Financial Variables Most Influential in Predicting Motor Carrier Failures.* Ray R. Hashemi and Louis A. Le Blanc and William Hinson and Walid Al-Masyabi and Garland Chow. 252

- *Disaggregate Analysis of Freight Transportation Characteristics.* Edward S.K. Fekpe. 267

Session/Séance 221: *Regional II*

- *Modelling a Trip Interactive Planner (TIP) to Manage and Monitor Tourist Traffic in a Regionally Based GIS.* Stefania Bertazzon and Nigel Waters. 281

- *Performance Measurement in the Public Sector.* Bill Raney and Mayank Rastogi and Mark Cowley. 294

- *Recent Estimates of Truck Travel in Ontario.* Joseph Jones and Greg Little. 309

Session/Séance 222: *Highway III*

- *An Industry With Many Faces: Balance Sheets Analysis of For-hire Trucking Companies.* Robert Masse. 326

- *Moral Aspects of Road Pricing Policy.* Dr. Mark Bunting. 340

- *Private Fleet Competitiveness: Is Backhauling an Opportunity?* Andrew Mardon. 355

Session/Séance 223: *Marine*

- *Shipment of Grain to Mexico: Sea Routes vs. The Mid-Continent Corridor.* Dr. Barry E. Prentice. 370

- *Shipping and Government - A Difficult Marriage.* John L. Eyre. . . . 384

VOLUME II

Session/Séance 241: *Rail*

- *Analysis and Transfer of Commodity-Specific Shortlines in Western Canada - Case Study: CN Avonlea Subdivision.* Stephen D. Granata and K.L. McGown. 401

- *What If? / Why Not? A Railway Tridea.* F.H. Howard. 418

Session/Séance 242: *Highway IV*

- *1995 CCMTA National Roadside Survey Project.* Bill Harbour. 432

- *An Experience in Modelling Commercial Vehicle Surveys.* Robert J. Tardif and Don L. McLeish. 445

- *Operating Costs of Trucks in Canada - 1996: A Synopsis.* Lloyd L. Ash and Barry Gardner. 461

Session/Séance 243: *Passenger Issues II*

- *Focusing on Growth Through High Quality Service and Innovative Products at Low Cost.* Gerry Kolaitis. 476

- *Logistics in Passenger Transportation.* Fredrick Petrie. 490

Session/Séance 252: *Highway V*

- *Manitoba-Related Trucking in the Mid-Country Corridor.* Jeannette Montúfar and Alan Clayton. 510

- *The "Grain Train" in Washington State: An Evaluation of Internal and External Benefits.* Kenneth L. Casavant and Richard Mack and Gary Richardson. 525

- *The State of the Canadian Intercity Highway System, 1986-1993.* Bob Leore..... 540

- *The Welfare Grain From Trucking Deregulation: Transborder Trucking and Cabotage.* Darren Prokop..... 555

Session/Séance 253: *Policy Experiences*

- *Turning the New Reality Into Reality: Lessons From Europe.* Mario M. Bozzo and Lee S. Sims..... 574

- *What Is Canada's National Transportation Policy?* Javier Caceres and Darrell Richards..... 586

Session/Séance 311: *Energy*

- *Emissions From Light-Duty Vehicles in Hamilton-Wentworth Under Policy Scenarios: Applications of an Integrated Model.* Pavlos S. Kanaroglou and William P. Anderson and Robert South..... 600

- *The Relationship of Fuel Consumption and Road Surface Quality.* Jerry Fruin and Dan Halbach..... 614

Session/Séance 312: *Safety*

- *Overblown: Examining the Costs and Benefits of Automotive Air Bags.* W.G. Waters II..... 625

- *The Safety of Sea Transport: A Crew Injury Perspective.* Wayne K. Talley..... 644

Session/Séance 313: *Highway Infrastructure*

- *Pavement Deterioration and Truck Loads in Winnipeg.* Mohammed M. Alam and Alan M. Clayton and Fred P. Nix..... 661

- *The Economy-Transportation Linkage in Ontario: A Sectoral Perspective.* N. Bedi and A.A. Kubursi. 675

Session/Séance 321: *Modelling*

- *The Saskatchewan Transportation Model and Its Use in Regional Transportation Planning.* Doug Neis and Paul Rachar and Daryl Nixon. 691

Session/Séance 322: *Logistics*

- *Performance Evaluation by North American Carriers: Preliminary Thoughts.* Mary R. Brooks. 708

- *Re-engineering the Logistics of Canadian Grain Handling or "Back to the Sacks?"* Dr. Barry E. Prentice and Jake Kosior and Jim Thomson. 722

Student Paper Winners:

- *Non-Stop Appeal: Viability of Long-Range Low-Density Air Routes.* Terence P. Fan. 738

- *Virtual Transportation as a Solution to urban Transportation Problems: A Critical Analysis of the Potential of Telecommuting as a Work Trip Reductions Strategy.* James A. Buczynski. 753

- *Power Failure - An Examination Into the Political Economy of the Electric Vehicle.* Daniel Schwartz. 768

- *Snake: A New Information Fusion Approach to Augment GPS Using Map-Derived Information.* Chuang Tao. 785

- *A Proposed Framework for Optimized Road Maintenance.* Curtis F. Berthlot. 799

- *An Analysis of Canadian Air Transport Policy.* Mark Grist. 812

ANALYSIS AND TRANSFER OF COMMODITY-SPECIFIC SHORTLINES IN WESTERN CANADA - CASE STUDY: CN AVONLEA SUBDIVISION

Stephen D. Granata, M.Eng., Eng., Project Officer, Canadian National Railways
K.L. McGown, PhD., Director, Transportation Management Center, Concordia University

ABSTRACT

A brief overview of the transfer vehicles available to Class 1 railways is first presented. This is followed by a description of the financial and related qualitative factors considered by Canadian National Railways (CN) during the course of a potential shortline transfer. The method by which CN evaluates potential shortline partners is then presented. Conditions specific to the analysis of the Avonlea Subdivision are discussed.

Part 1: HISTORICAL

1. Brief history of shortlining

Shortlines may be described as generally short segments of lighter density railway whose primary function is to act as a feeder to main line, higher density carriers.[1] American Class 1 railroads began network rationalization efforts approximately 15 years ago. Today there are about 500 shortlines operating on 20 percent of the freight track in the United States. Successful shortlines are notable for their greater total assets, higher operating revenues, higher traffic densities and greater length of track operated.[2]

Several regional and terminal railroads have operated in Canada for many years, spread out nationwide and carrying a wide variety of commodities including grain, chemicals, coal, forest products, intermodal and manufactured products.[3] However the creation of newer shortlines in Canada had advanced at a relatively slow pace since the introduction of the National Transportation Act (NTA) in 1987. This may possibly be due

to the fact that until passage of the Canada Transportation Act (CTA) on July 1st, 1996, the process whereby a Class 1 carrier could divest itself of lighter density line was cumbersome at best, involving the introduction of a private members bill in the provincial legislature and/or obtaining a federal certificate of fitness.

Several environmental factors are now in place to ensure that shortlines assume a role of greater importance in the Canadian freight railroad industry, including the passage of the CTA, increased Class 1 focus on mainline traffic and the existence of several established American and Canadian shortline railroad companies anxious to maintain growth rates through further acquisitions.

Thanks to years of overbuilding and subsequent government restrictions on network rationalization, the two major Canadian Class 1 railroads currently possess a sprawling prairie rail network. Western Canada is therefore a major focus for network rationalization efforts. Grain-dependent lines would intuitively imply a larger collection area to avoid dependency on a single commodity, assuming that each line is considered independently. However the shortline industry has evolved to the point where larger firms are seeking to benefit from geographic and commodity diversification and as such may be less concerned about the risk level inherent to any given shortline. Of greater concern is how the new acquisition will complement the mix of lines already owned by the firm.

2. Overview of selected legislation and their effects upon the creation and operation of shortlines

Canada Transportation Act (CTA)

Implementation of the CTA has simplified the process by which Canadian Class 1 railroads may divest themselves of unprofitable lines. It established three categories of Class 1 rail line in Canada:

Retain: The line remains in the Class 1 rail network
Transfer: The line is offered for sale to shortline operators, but will continue to be operated by the Class 1 railroad in the interim
Discontinue: The line is offered for sale, but will be abandoned if nobody makes an acceptable offer to lease or purchase

Under terms of the act surplus track may be sold to another economic entity or abandoned, according to the following methodology:

Railroads must make public a rolling plan detailing which lines they foresee transferring or discontinuing over the next 3 years. Canadian National Railways (CN) currently updates and makes public its plan on a quarterly basis. Once a plan is announced the railroad must wait for at least 60 days before proceeding with plans to abandon any given line. Once a line is advertised for discontinuance potential buyers have a 60 day period in which to approach the railroad and express interest in the line. If a buyer expresses interest, the railroad and the buyer have 120 days to conclude an agreement. If no agreement is reached the railroad is free to proceed with the abandonment process. The line is next offered to different levels of government, who each have 30 days to accept an offer to purchase at a maximum of net salvage value.

Gone are the mediation and public interest provisions contained in the NTA, which tended to hinder the Canadian Class 1 railroads significantly in their attempts to discontinue or transfer unprofitable lines.

Successor Rights

The issue of successor rights, in which the new owner of an economic entity is bound by labour agreements in place at the time of purchase, is provincially mandated following a test case in 1990 before the Supreme Court of Canada. Among the provinces British Columbia and Saskatchewan have enacted legislation providing for successor rights when an enterprise changes from federal to provincial control. Reaction to the existence of this legislation amongst shortline operators has been mixed. Some firms believe that such laws will effectively negate shortlining opportunities in the provinces in question, while others acknowledge that, despite the presence of such laws, unions and government appear prepared to demonstrate a significant degree of flexibility in their positions given the alternative with which they are faced: the loss of a significant number of lighter density railway lines. For the unions this represents the loss of a significant number of jobs, while the provincial governments in question, and specifically the Government of Saskatchewan, are considering the effects of a loss of rail lines on the communities to which they currently provide service, as well as the increased maintenance cost of their provincial roadway system.

Part 2: CURRENT PRACTICE

1. Rail Line Transfer Vehicles

Internal Shortlines

Internal shortlines appear to be a distinctly Canadian phenomena in North American rail operations. While CN appears to have coined the term Internal Shortline, Canadian Pacific Rail (CP) initiated the first practical application of this concept in the early 1980's when the Canada Atlantic Railway (CAR) was established to assume operations of CP's Atlantic Provinces rail network.

At CN an internal shortline is defined as a semi-autonomous organization within corporate CN possessing its own management, operations, and field marketing/sales functions. The internal shortline is envisioned as having modified labour agreements, work practices, and management techniques. However, it relies on the central organization to supply car equipment, motive power, and specialized resources such as technical and environmental expertise, program engineering maintenance, accident recovery equipment, and so on.

The benefits associated with an internal shortline centers on heightened innovation and cooperation by both management and labour, as well as a more local focus and increased responsiveness to customer service issues and opportunities. Also, modified collective agreements can result in a reduced cost base via combining job positions (composite employees) and simplifying and reducing the wage rate schedules.

Finally, the risk associated with a potential loss of control of the customer base on a branch line(s), when compared to an external shortline sale, are negated. Also, provisions of the CTA which impact on running rights and joint access can be mitigated by establishing an internal entity versus an external line sale.

Liabilities confronting internal shortlines are primarily related to the fact that in the final analysis we are still dealing with the same people, the same labour organizations, and the same labour agreements (albeit that the latter may have been modified). In addition, the internal shortline must also compete for capital and operating funds within the corporate organizational framework.

External Sale

Several different forms of shortline external sale transactions are possible. These include the Conrail Express shortline network whereby the independent shortline becomes, in essence, a franchisee and operating arm of the Class 1 railroad. Under this scenario, Conrail performs all the billing, claims, and other customer service relations, and also provides car supply. The shortline in not mentioned on any documentation such as the waybill or bill of lading.

At the other end of the spectrum is a totally stand alone partner who is identified as an independent, full service entity by the customer and, by agreement, deals with the Class 1 railroad in an arms length fashion.

Oftentimes the Class 1 railroad offers financing to the potential shortline partner. The Class 1 railroad has a greater degree of flexibility in setting the capitalization rate, which would be established in the context of the overall economic benefit of the project to the railroad, compared to the financing rates charged by third parties which are established in the context of the perceived risk/return profile of the investment. As financial institutions become more familiar with the shortline railroad business, the risk premium to be paid on debt will likely decrease. This bodes well for future shortline startups.

Lease

A fundamental difference between a sale and a lease is that with a lease the Class 1 railroad retains title to the asset and accordingly is afforded a greater degree of control over the future of the shortline. The financial distinction between a sale and a lease can be one of cash flow. We would intuitively state that a lease agreement tends to minimize the startup capital requirements of the shortline operator, whereas a sale transaction tends to increase capital requirements. However, given the various methods available to structure a lease or sale agreement, this distinction has become less clear. For example and as mentioned earlier, oftentimes the Class 1 railroad will offer financing to the short line partner, and as such may defer receipt of the purchase price to a future period. Another factor influencing the cash flow requirements of the purchaser involves the negotiation of the car division, the amount paid by the Class 1 railroad to the shortline partner for delivery of traffic to the former. Various methods exist for determining this level, which may be paid on a per car, per tonne or on a distance related scale. However this value and the

transfer price or lease rate are closely related, such that the purchase price of a given line, for example, may effectively be driven down to zero if the proper car division is agreed upon.

One drawback related to the lease option is that the original owner remains at least partially liable for any future costs associated with the rail line, for example environmental remediation projects or the removal of infrastructure should the rail line eventually be abandoned. However, to what degree the Class 1 railroad distances itself from this problem through an external sale is not completely clear.

Other drawbacks include the fact that the lessee has less of a stake in the asset, and hence possibly less motivation. Another is the possibility that the line will remain under federal jurisdiction, and the application of successor rights in certain provinces will remain an issue. However this may be circumvented through the implementation of a capital lease. Finally, there remains the possibility that the lessee will "bleed" the railroad through systematic avoidance of capital investment in the infrastructure and eventual abandonment.

Whichever degree of interrelationship is adopted, it is of primary importance that the shortline partner shares CN's view and direction regarding their business relationship. The degree of success of such compatibility ranges from the general (i.e. overall corporate philosophy) to the specific (i.e. data system, billing/claims processes, physical operating interface, marketing).

2. Rail Line Analysis and Transfer Procedures

A. Rail Line Analysis

Financial Analysis

The process whereby lines are scheduled for transfer or discontinuance stems from work performed by the Costing and Network Rationalization Departments at CN. Economic assessments of lines, segments or groupings are reviewed by the Network Rationalization Committee, which meets on a quarterly basis and consists of representatives from Costing, Network Rationalization, Marketing, Transportation, Business Planning, Law, Real Estate, Engineering and Public Affairs. The method consists of using historical costs and revenues to identify underperforming

assets and to determine which networks should be abandoned and which are potential shortline candidates.

CN has compiled data on its subdivisions in a database called NRS (Network Rationalization Studies). The database provides three years of historical information such as originating and terminating traffic (by point of origin/destination, commodity), revenue, and long term volume variable costs such as line operating costs, other on-line costs and off-line costs. This data is used to provide a preliminary indication as to which lines should be considered for discontinuance or transfer. Those lines identified as possible transfer candidates are subjected to a more detailed economic analysis in order to ascertain the terms necessary to negotiate an agreement on behalf of CN.

The detailed economic study typically takes the form of a net present value analysis discounted using the CN internal cost of capital and a time span which varies depending upon the life expectancy of the line in question. Several scenarios are examined, including status quo, discontinuance and several transfer scenarios measuring sensitivity of the result to a variety of factors such as variance of future revenue, sale price and rate division, physical point of transfer and realized savings through discontinuance of our operations over the lines in question. All calculations based on NRS data are then reexamined using direct or budgetary costs, given the sometimes significant differences between attributed, long term volume variable costs and direct costs.

The comparison of internally attributed (long term volume variable) costs vs. direct costs is valuable. Costs attributed to a given line for managerial accounting purposes do not necessarily reflect the actual cost associated with operating that line. For example, how will the sale of a given line actually effect the operating budget of the Engineering or Transportation departments? Will the transfer result in actual operational budgetary savings or are the associated costs largely fixed?

In the case of lines designated as grain dependent (those whose originating or terminating traffic is greater than 60% grain) an additional consideration is the WGTA rate scenario, in which the cost of owning and maintaining the line in question is removed from the cost base should the line be salvaged, thereby possibly reducing the rate that can be collected

by the Class 1 railroad. This effect must be considered in the discontinuance scenario.

A fact particular to the Canadian prairie provinces is that many of the grain lines were rehabilitated in the 1980's under the Branch Line Rehabilitation Program and as such tend to have higher salvage values. Canadian Class 1 railroads have tended in past to insist on a transfer price equal to or greater than salvage value. However given the above fact and lower traffic volumes typically associated with grain line feeder networks, the purchase price for these lines could become prohibitive. One possible solution is for the Class 1 railroad to retain the right to buy back the line for the original sale price should service on the line be discontinued. The Class 1 could then salvage the line and benefit from its full liquidation value. Note that this methodology was instituted by Union Pacific Railroad and RailTex several years ago in the United States. Under this scenario it is in the best interest of the Class 1 railroad to ensure that adequate maintenance of the physical plant is performed by the shortline partner. Another potential problem presented by the higher salvage values of the prairie grain lines is that the resulting higher asset levels carried by the shortlines could cause them to lose their small business status and associated tax savings.

In performing a financial analysis of a potential transfer line CN must consider the net economic benefit to be derived by CN, as well as the potential worth of the given asset to the shortline partner. For this reason a thorough analysis of the estimated financial position of the shortline entity must be undertaken in order to negotiate the best possible agreement for CN. Factors to consider include the potential for traffic growth on the line, as well as renegotiated labour rates to be instituted post line transfer. CN has developed an internal model to undertake this type of analysis.

Historically Class 1 railroads tended to consider cost minimization or traffic retention as the principal drivers in their network rationalization efforts. However as efforts progress they must begin to consider total traffic and contribution maximization as the ultimate goals of their rationalization efforts. It is recognized within CN that the contribution of currently profitable lines could also be improved upon by utilizing the inherent competitive advantages enjoyed by shortline railroad operators.

Other Considerations

To what extent can the past predict the future? It is obviously not correct to assume that because a Class 1 railroad has spent x dollars per year maintaining a line over the past 5 years, that the same annual sum will be required in future. Factors affecting annual expenditures include level of available capital, projected traffic volumes, loading capacity and possible rationalization efforts. The parties must therefore attempt to determine the future capital requirements associated with line maintenance and other factors not adequately measured using historical data, in order to correctly assess the viability of the shortline.

Another factor to consider is the potential savings due to improved car cycles on the line in question. It is difficult to quantify with any degree of precision savings associated with reduced car cycle times. However it must be recognized that shortline operators in most instances achieve significant reductions in car cycle times on the lines over which they operate.

How should the Class 1 railroad compensate the shortline partner for traffic delivery? By car, by weight, by division based on miles hauled? This will depend in part upon the nature of the particular line under consideration. If the line is to operate over a limited distance, then a distance-based division is probably not equitable given the high handling time to hauling time involved. CN prefers the method whereby the shortline partner is compensated on a per car basis, as it has proven to be cost-effective to administer.

Other intangible factors include location of the line. Is the line located in a highly sensitive area where competition is high? Does CN want to maintain a presence in the area even if it is not cost effective to do so in the short term? Does selling the line provide the shortline partner access to competitors, thereby allowing him the possibility of siphoning traffic to a competing railroad?

A wide range of incentives can be utilized by a Class 1 railroad to influence the routing of traffic. Likely the most stringent incentive occurs when the Class 1 railroad provides financing for a portion of the initial purchase price. Allowances may be linked to the delivery of specified volumes of traffic to the Class 1 connection. If these targets are not met or if traffic is diverted to a competing carrier, allowances would be waived.

A variation on this approach is to establish a link between traffic delivered and a credit being applied toward the purchase price/mortgage.

Another type of incentive involves CN making accessible its various specialized services, expertise and its procurement process to the shortline partner in return for traffic commitments. Access to areas such as locomotive fuel, maintenance of way material, and so on can positively impact on the cost base of the shortline partner.

Specific, defined clauses forming part of the contractual agreement between the Class 1 railroad and the shortline partner can also be used to encourage the shortline to deliver traffic exclusively to the Class 1 railroad. Aspects of the Competition Act may influence the degree of control the Class 1 railroad can exert in specifically directing traffic to itself. Existing traffic traveling on historic routings does not appear to infringe upon anti-competitive considerations as contained in the Competition Act. However, new traffic volumes developed after the shortline acquisition are less well defined with regard to a restrictive routing covenant. Competitive considerations under CTA legislation may have an impact on such restrictive clauses. This area is currently untested in law.

In establishing shortline entities, several U.S. Class 1 railroads have successfully restricted access of the smaller carriers to competing Class 1 railroads by designating the official interchange location in an isolated or at a "safe" point which physically separates the shortline from other Class 1 carriers. To reach an operationally suitable exchange location, the shortline is granted operating rights over the intervening section of Class 1 railroad. While this concept has been successful in the U.S., aspects of the CTA legislation regarding competitive line access will have to be tested before it can be conclusively stated that this restriction is valid in Canada.

B. Transfer Procedures

Identification and Evaluation of Potential Partners

The main driver behind CN's selection of a shortline partner is to ensure good rail service for customers on a line slated for transfer. It is therefore crucial that the correct partner be selected.

A full range of criteria are reviewed before a short list of candidates and a finalist are selected. It is the intent of CN to form long term relationships with our shortline partners. CN must evaluate all likely shortline partners using common, all encompassing selection criteria in order to arrive at the optimum choice for each shortline transaction.

Likely the most important criteria in evaluating a shortline candidate is the financial resources available to the partner to successfully negotiate, implement and operate the line/network being transferred by CN. Financing is the primary hurdle to new entrants to the shortline railroad industry.

CN's evaluation criteria must thoroughly measure the ability of the potential partner to source funding for the various phases of the project's life cycle, including acquisition, start up and operations. While the most obvious financial requirement is the candidate's ability to finance any lump sum payment in the initial acquisition phase of the project, it is also imperative to measure the party's ability to weather a prolonged negotiation process (due, for example, to legal challenges, government approvals, etc.) leading to the conclusion of a transfer agreement.

Prior to revenue flows becoming available to the shortline, operating start up costs must be addressed out of initial financing. These costs generally include the acquisition of locomotives, establishment of maintenance and office facilities, recruitment and training expenses, supplies and materials, and so on. Collectively, these expenses may range into the hundreds of thousands of dollars for even the most elementary of shortline operations.

In addition to normal operating flexibility, the potential partner must also be measured on his ability to bridge seasonal or market related fluctuations in revenue. Is the proposed shortline a single unit operation or a stand alone component of a holding company? Are financial resources available from other rail or non-rail components of the partner's enterprise to bridge such fluctuations on the specific line segment? These questions are doubly relevant in western Canada, where several lines to be transferred are single commodity dependent (particularly grain products), and the ability to weather market variability becomes a major concern for long term viability. Larger shortline companies appear to have as one of their corporate objectives the acquisition of properties that may be single commodity dominant but, in the overall corporate framework, produces a

complimentary and diversified commodity mix possessing geographic diversification.

A second important indicator used to gauge a shortline partner is their experience level. It is necessary to measure the operational, marketing and managerial experience available to the potential partner, as well as any experience directly related to operating on the territory in question (national, provincial). Of particular importance to CN is the past safety record of the shortline operator.

Receipt and Evaluation of Business Plans

Following the signing of a confidentiality agreement, CN provides potential partners with a standard package containing information generally required to develop a business plan.

Upon receipt, submitted plans are evaluated for validity given the detailed knowledge available to CN regarding the current operating and maintenance expenditures as well as the marketing prospects on the line(s) in question. Major elements for review include the validity of the initial and ongoing financial aspects of the business plan.

Regarding the operating component of the projected plan, do train and switching operations appear to be adequate to properly service the existing and anticipated customer base? Do these operations link effectively with the connecting CN operations and network? Does the maintenance of way, structures, and facilities schedule meet the agreed upon sustainable objectives?

Given the unique nature of CN's western candidate lines, which are largely one commodity traffic dependent, the potential partner's plan must include provisions to account for fluctuations in traffic volume due to seasonality and/or market shifts. CN's review must also examine whether the potential partner has assessed the local and macro economic environment as well as the impact of competition (rail, road, water) on their projected operation.

In conclusion, a well researched, detailed business plan developed by the potential partner, which can be thoroughly defended and supported, will increase the likelihood of success regarding the successful completion of

the transaction and the ongoing business relationship between CN and the shortline partner.

Shortlisting of Potential Partners

Once the short list phase is reached, additional evaluation of the potential partner will shed light on the suitability of the candidate. How thorough and with what degree of understanding does the candidate evaluate the physical plant offered? Does he appear to have a complete understanding of the assets to be included in the transaction (physical plant, motive power, car equipment, maintenance of way equipment, buildings, and so on)?

The customer base on the territory affected by the sale can provide valuable insight into the potential partner's suitability. The candidate's understanding of the customers to be served on the shortline and his willingness to provide conventional and innovative service will provide further feedback regarding his marketing abilities. In addition, customer feedback stemming from primary and follow up due diligence contact with the candidate will aid in determining his "fit" in this transaction.

Part 3: CN AVONLEA SUBDIVISION (SASKATCHEWAN)

1. Brief description

The Avonlea Subdivision is currently designated as a transfer candidate by CN. It extends southward from Moose Jaw, Saskatchewan (mile 88.3) to Avonlea, Saskatchewan (mile 51.7), connecting with the CN Central Butte Subdivision in the north (Moose Jaw Jct., mile 38.9 of the Central Butte Subdivision) and the Southern Rail Cooperative Railway (SRC), a shortline established in 1989 when CN divested itself of the southern half of the Avonlea Subdivision, in the south. The right of way encompasses 455 acres of land which includes 36.6 miles of main track and 6.28 miles of siding track.

Total carload traffic originating/terminating on the Avonlea Subdivision during the past three years are as follows: 1472 in 1995, 2520 in 1994 and 2689 in 1993. The line is classified as grain-dependent, with a minor amount of terminating traffic consisting of machinery, fertilizer, lumber and iron and steel products. Major customers include Saskatchewan Wheat Pool, Agpro Grain Inc. and Parrish & Heimbecker Ltd.

Service is currently provided once weekly by dedicated wayfreight originating at Regina Yard. Cars are spotted at two CN stations and for SRC on Thursday and the engines are left at Avonlea station. The crew is deadheaded back to Regina on Thursday evening and another crew is deadheaded to Avonlea on Friday morning for return service. Moose Jaw customers are serviced thrice weekly by separate wayfreight via a CN/CP interchange located on the Avonlea.

The Avonlea Subdivision was upgraded between the years 1979 and 1990 under the auspices of the Branch Line Rehabilitation Program. Rehabilitation works included 100 lb. rail relay, ties, ballast, and surfacing. Net salvage value of the line has been established by CN to be in the order of $1,000,000.

Another detail relevant to this study is the fact that a high capacity grain elevator is being constructed on the Central Butte Subdivision, approximately 5 miles to the west of the Central Butte-Avonlea junction.

2. Analysis

Four scenarios were considered in attempting to determine whether to transfer the Avonlea Subdivision:

- Status Quo: CN to retain the line
- Discontinue: CN to salvage the line
- Transfer to a shortline operator south of the CN/CP interchange
- Transfer entire subdivision to a shortline operator, including CN/CP interchange

As the future of the Avonlea elevators is uncertain in the medium term, the net present value analysis was discounted over a relatively short time span (10 years) in order to assure a quick pay back for both CN and any eventual shortline partner. Parallel NPV calculations were performed using both long term volume variable and direct costs. As preparations for the possible transfer of the Avonlea were ongoing as of the writing of this paper, the following discussion will provide an overview of financial results, more emphasis being placed upon the relevant factors affecting the analysis.

Although relatively small compared to other CN shortline offerings, the Avonlea provides a relatively low cost means for larger shortline firms

new to single-commodity railroading in western Canada to gain valuable experience and contacts. Its purchase would certainly assist them in preparing for the acquisition of larger offerings in this area, such as the networks currently scheduled for transfer in northwestern Saskatchewan.

Analysis of the various scenarios was complicated by the construction of a high capacity grain elevator on the Central Butte Subdivision, a few miles to the west of the Avonlea/Central Butte junction. Associated effects upon service design were such that not even the status quo scenario was straightforward to quantify. Service design was in the process of being studied for eventual modification as the analysis of the various scenarios for transfer of the Avonlea proceeded. Savings associated with the cancellation of the dedicated wayfreight on the Avonlea were relatively simple to quantify. It was determined that existing train service on the Central Butte, currently underutilized, could also incorporate additional traffic feeding off the Avonlea. As such the Avonlea traffic could be transported to Regina at little incremental cost to CN. Assumptions had to be made regarding eventual service design to the newly-constructed grain elevator in order to assess the possibility of incorporating the Avonlea traffic into this mix.

Another factor affecting the analysis was the possible transfer of the CN/CP interchange. Discussions with departmental representatives regarding the interchange indicated that the Marketing Department preferred to retain it, primarily for competitive reasons, while the Transportation Department preferred to transfer it due to investment requirements and operational difficulties associated with the interchange. It was felt that CN could adequately protect the Avonlea traffic by contractual means using one of several techniques as described in earlier sections of this paper. However transfer of the interchange would also imply the presence of a second intermediate operator in servicing CN's customers in Moose Jaw, leading to possible service delays.

Characteristics of the Avonlea physical plant were examined to determine where traffic could be interchanged between the shortline partner and CN. Given existing space limitations, would additional trackage be required at the Central Butte yard (located at the junction between the Central Butte and Avonlea Subdivisions), or could traffic be left on the wye connecting the two subdivisions? The answer to this question, as well as the associated capital expenditures required in the event that space was not

sufficient, was dependent in large measure upon the as yet undetermined level of train service frequency required at the new grain elevator.

With regard to the issue of car division and sale price, there are advantages to each of the two available approaches. By insisting upon a larger up-front payment from the buyer CN mitigates the risk associated with recovering our investment over the longer term due to the possibility of shortline failure. However CN does recognize that the debt burden placed upon the shortline will directly affect his probability for success on the line in question. In order to reduce CN's operating ratio it is more expedient to accept a lower cash payment in return for a lower car division and hence lower annual operating expenses.

Having completed the financial analysis for each of the scenarios using both direct and long term volume variable costs, as well as an analysis of the estimated financial position of the shortline assuming ownership of the Avonlea Subdivision (using a 20% before tax Return on Investment (ROI)), it was determined that a mutually beneficial agreement could be reached whereby CN could improve its current financial position and the shortline could earn a sufficient return on equity to make the shortline a viable venture. However the window of mutual benefit was determined to be quite small, especially when considering the direct cost NPV analysis.

It is our belief that the transfer of the Avonlea Subdivision is too small a project on which to realize significant economies of scale. Several smaller scale lines would also have to be transferred on the district in order to impact local operating costs. Given the fact that the effort required to divest a small line is quite comparable to that required to divest a larger network of lines, this scenario would seem unlikely in the near term.

CONCLUSION

Environmental factors are now in place to ensure the future of a strong Canadian shortline industry. The ensuing network will feed Class 1 railroads increasingly focused upon long-haul traffic traveling along their core lines. As network rationalization efforts progress, the Class 1 railroads must begin to emphasize traffic and contribution maximization as main drivers in the identification of additional shortline candidates, recognizing and taking advantage of the inherent competitive advantages of shortlines. This will necessitate a change in mentality within certain areas of the Class 1 railroads that must be driven by top management, in

order to overcome a historical bias which emphasizes strict revenue retention. As it has been determined that the effort involved in transferring a smaller line is comparable to that of transferring a larger network, emphasis will be placed on the transfer of larger networks in the near term.

Bibliography

---

[1] J.M. Bowland, R.G. Friend. "Opportunities for Short Lines in Canada." KPMG, p 1. March 29, 1995.
[2] J.M. Bowland, R.G. Friend. "Opportunities for Short Lines in Canada." KPMG, p 11. March 29, 1995.
[3] J.M. Bowland, R.G. Friend. "Opportunities for Short Lines in Canada." KPMG, p 5. March 29, 1995.

WHAT IF? / WHY NOT?

A Railway Tridea

F.H.Howard P Eng Richmond B C

1. RUBBER AND RAIL

The ability of a locomotive to exert tractive effort or drawbar pull - the measure of what it could lift vertically, say over the edge of a cliff - and so, once the train's resistance has been determined, what tonnage of train it can start, is restricted by its wheels' adhesion to the rails, normally about ¼ of its weight. wheel slip control (replacing sanding) has now raised this to about 1/3. Too much power, and its wheels will slip. Starting a train is adhesion-limited; running it is horsepower-limited.

When inverted, this fraction becomes "Factor of Adhesion" and refers to steel wheels gripping - or slipping - on steel rails. Some locomotives are ballasted to achieve adhesion, which affects braking too..

Heavier trailing tonnages can be moved if a much lower Factor of Adhesion can be developed. Such a low factor is indeed developed by rubber on paving. Assuming the road is dry and the tires sound, a rubber-tired highway tractor can lift half its own weight, with the corresponding capacity to pull a heavy trailing load, usually a semi-trailer, also on rubber tires.

A number of rail vehicles use a Hi-Rail device, hydraulically-lowered sets of rail wheels, commonly attached to rubber-tired track inspection and maintenance equipment, especially automobiles, pickup trucks or vans; sometimes little cranes. This makes them amphibious, so to speak, at home on road and rail. When finished with rail duties, they can get off the track and go away.

This feature has been used in a most sophisticated way in Saskatchewan. Originally conceived by the Highways Department as a solution to the projected abandonment of some railway branches and consequent overloading of thinly-paved roads, the Brandt Road Rail unit is a unique heavy-duty bi-medium quasi-locomotive. Using this basic Hi-Rail principle to guide it while in the rail mode, it weighs up to 25 tons (depending on auxiliary attachments) of which 14 tons rest on the driving wheels, and exerts 30,000 lbs of drawbar pull to start and 425 HP to run.

This ingenious machine depends for traction on its tires adhering to steel rails at a Factor of 2, clearly superior to steel on steel, and can pull its own steel-borne weight. The original purchaser uses it to alternate between two isolated short railway lines. Another promising employment is in areas where unconnected branches need irregular service; a single power unit can visit them on different days. At about $500,000, it costs considerably less than a new diesel switcher.

Fig 1 shows this versatile creation



Fig 1

This paper proposes the further development of that marriage of steel on rubber. Instead of adding Hi-Rail flexibility to the truck/locomotive, let it remain solely rubber-tired, but pave the railway. Not pave over the tracks, but up to each side, and let the trailing trains steel wheels remain on steel rails as always, with their low rolling resistance*. The towing truck to do what it does best, provide high rubber-on-asphalt adhesion (but no longer pulling a high-resistance trailer). The paving need be to no particular degree of smoothness and can extend the length of a branch or just a siding, wherever this low-cost system can be exploited. In fact an entire small freight yard could be paved right across, and say a heavy-duty ballasted 4x4 truck be substituted for a diesel switcher at half the capital cost. It could then dart around the yard from track to track without shunting in out, back and forth, through rail-bound switches. Automatic transmission is mandatory to protect clutches against their destruction

*Train resistance is an arcane science, involving a complex calculation known as the Davis Formula. While starting resistance is an indistinct number, rolling resistance, varying with speed and inversely with weight per axle, comes out at about 5 lbs per ton at medium speed

Care must be taken to prevent the train exercising its resistance by trying to overturn the pulling vehicle, as farm implements used to overturn tractors by coupling them too high. The standard railroad coupler height is 34½", and no drawbar should be mounted lower.

Now a mythical 17-ton 4x4 truck on steel wheels could exert 8,000 lbs drawbar pull. The same truck, its rubber tires rolling on asphalt, can exert 17,000 lbs, about as much as a young locomotive. The Brandt truck-motive develops 30,000 lbs, more than many small steam locomotives of 65 tons used to.

It has hauled 2,000 tons with its wheels adhering to rails. When they are on paving, with the train still on rails, what tonnage can be started? Let us speculate: 5,000 tons?

Any road vehicle converted to rail operation must have standard air brakes, which Brandt has, and the authorities will almost certainly ordain headlight, ditch lights, a horn and bell. The rear view must be of the best; Brandt uses a TV monitor.

Caterpillar offers a variation on this theme: a massive 4x4 log skidder with big wheels, suitably modified by replacing the logging and bulldozing apparatus with couplers, air compressor, reservoir and controls. The model examined generates a starting drawbar pull of 44,000 lbs, which, in common with all constant-horsepower locomotives, declines with speed in a hyperbolic curve.

This is shown in Fig 2

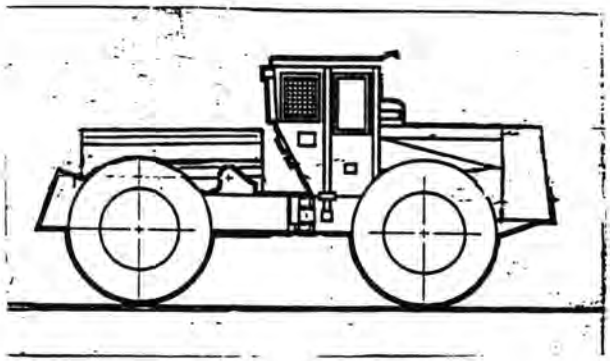


Fig 2

Horsepower in this case is markedly lower than Brandt's which will affect its top speed. The factory weight is about 16 tons, and in pursuit of effective traction, ballast is needed, along with the fattest of bald tires. Its drawbar pull of 10,000 lbs at 6 mph should move along a 10,000 ton. train. It will be priced at about $450,000±.

A one-man engine crew can work the yard, and for the road, negotiations will determine its size. In this regard, a Short Line may be able to economize on engine manning.

2. TRACKS UP - MAINTENANCE DOWN

This is the revival of an old idea and there may not be many more lines of railway laid, since the industry is mature and if anything is diminishing in size if not in all species of traffic. Short Lining is not going to reverse this trend.

But when a new railroad is to be built, such as for a High Speed Train, or an existing line relocated or massively improved, or a heavily-damaged or washed-out line repaired, this paper explores the practicability and poses the alternative of the trackage no longer laid on the ground but on an elevated structure instead.

It is acknowledged without argument or reservation that the greatest obstacle to lowest-cost railroad operation is gradient. To quote the late John Barriger,"....whenever a train ascends 8 feet, irrespective of the rate of grade, the locomotive does as much work in overcoming the force of gravity as is performed in hauling the train one mile on level tangent track.* An elevated structure can provide an exactly level track by varying the heights of its supports.

An elevated main line railway, not to be confused with the New York or Chicago elevated transit lines, eliminates the grading, ballast and ties required with conventional trackwork, replacing them with a level structure of rails fastened to a concrete slab. Such a form of construction eliminates the maintenance consisting of repetitive ballast cleaning, tie replacement, the constant ditching for the critical drainage and the every-spring recovery from frost, to say nothing of the burden of highway grade crossings and keeping culverts clear.

* John W Barriger, "Super Railroads" Simmons-Boardman 1955

In short, along with grade crossings and blocked culverts, these are all the impedimenta of railroad Maintenance of Way Departments, reporting to the Civil Engineers. This profession also designs bridges, and this proposal is for a continuous bridge.

The Skytrain system of British Columbia Rapid Transit Inc provides a modern example of cross-country elevated railroad construction. Although moving small trains of very light weight at medium speed. the design can be upgraded so as to accomodate long trains of very heavy tonnage or short trains at higher speeds.

This railway was built for Expo 86 as an example of future technology in tune with a World Exposition. Originally joining two areas of the Fair, there are now about 17 km. of elevated high-density operation, with plans for more - when funds are available. Double trackways are supported on poured-in-place reinforced concrete T columns. They consist of pre-stressed post-tensioned concrete slabs about 100'long, with rails fastened to the concrete with conventional fasteners. The slabs are drained, and in this case incorporate low coamings along each side for protection against derailments.

Such a concept forms one basis for main line railroad construction, although there could be others, one of which is a steel structure with heavy bridge-type spans on steel columns. Steel's advantage is its superior weight-strength ratio. Weathering steel at higher cost will eliminate frequent painting; concrete of course is practically weatherproof. Steel bridges would be of the deck girder configuration for better snow-clearing. Wooden trestles can also produce a level elevated line without ballast, but with significant maintenance and with labour-intensive erection.

The structure should properly be no less than 10' off the ground, to allow motor vehicles, persons and farm animals to pass under; there may be local by-laws or zoning restrictions on this, but always off the ground.

For most of the distance it will be possible to maintain precisely level elevation. Traditional ground-laid track often has ups and downs, hills and dales, perhaps no great fuel handicap as they may not violate the Barriger 8'rule and grades are often momentum grades, but productive of slack run-in and out, demanding expert train handling, of throttle and dynamic brake. Broken coupler knuckles and pulled drawbars attest to long train slack action gone awry. Such track demands fills to cross more than slight depressions; elevated track simply requires higher pedestals. For major valley crossings, all types of course have to have bridges built.

Whereas ground-laid track needs wide rights of way, with attendant land purchases or expropriation, an elevated line takes about the area of a high-tension utility line.

Final examination of this proposal will match the respective capital expenditures per mile against the respective yearly maintenance costs, including snow clearing. Conventional location and construction, say in Eastern Ontario or on the Prairies run not less than $1 million per mile (excluding rail), more in the Shield, with maintenance at $1-2 per thousand gross ton-miles per year. The price of an elevated line will vary widely with the quality of the ground under the pedestals, but its will be a fraction of the alternative.

Shown below are scenes from coventional trackage dirty ballast, rotting ties, plugged spike holes, frost-heaved alignment, blocked culverts and weed-filled ditches, with grade crossings to be looked after and brush to be cut back.









One experience consultant, when retained to appraise track for a prospective buyer, says he looks for three things, "drainage, drainage and drainage"*

* Charlie Banks of Robert L Banks & Associates Inc of Washington, at the Second Canadian Short Line Conference, Toronto, 1996

With elevated track, no such anxiety, in fact, he can almost run along the line, barely looking at it. Fig 3 and 4 show the pier and line arrangement (in this case double track) of an elevated railway with none of these aggravations.

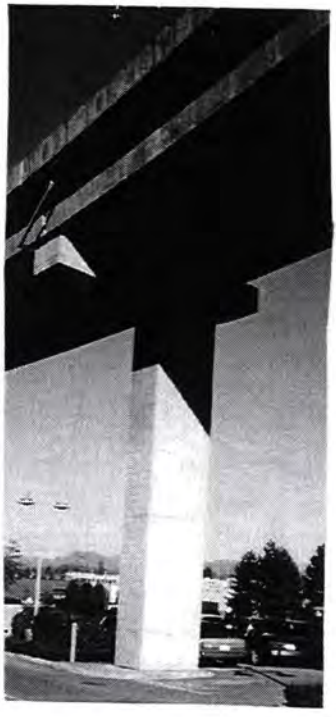


Fig 3



Fig 4

3. A FLATTENED HUMP

While much of modern railway freight operation consists of unit trains of great tonnage, loading and unloading wiuthout stopping, or of block trains of a single commodity destined to a few consignees (perhaps only one) there remains a significant number of trains of mixed car types, loaded and empty, gathered at a number of origins, that have to be sorted out into new trains, distributed among many destinations.

This classification has to be carried out at various nodes along the system, spaced according to operating vicissitudes, usually at Divisional terminals.

Very large classification yards are to be found at particular strategic points, long distances apart, often but not necessarily in big cities. An outstanding example of this is the Union Pacific's enormous facility at North Platte Neb, which classifies incoming long trains into a score or more of outgoing long trains at the rate of several thousands of cars a day over 24 hours. Similar Canadian facilities are at Toronto: Agincourt for CP, MacMillan for CN. In order to classify and reroute incoming trains in such principal yards, they are pushed up one side of a carefully-designed little hill. At the peak, one or more cars are uncoupled by a "pin-puller" (a reminder of the days of link and pin couplers) and gravity takes them down the other side, guided into the appropriate tracks by a controller operating powered switches; cars are decelerated by him through pneumatic retarders pressing on wheels as they go by. The "hump" is at the "throat" of a fan arrangement of many departure tracks. All this goes on without stopping, with the pusher locomotive set at a continuous humping speed. This serves the major freight yards effectively.

However, there are a great many intermediate yards that must perform these same car separations and classifications without the benefit of this high-class mechanism. Nearby examples are London for CP and Belleville for CN.

In these yards unclassified groups are separated by yard engines pushing them forward at (hopefully) moderate speeds until, just short of the throat where the destination tracks diverge, the locomotive is stopped abruptly, its cars are uncoupled, one or more, and continue to roll on into the track of choice, usually to join onto cars already there, thereby creating new trains. or at least new strings of cars.

But the remainder of the train having stopped, it must now be reversed, taken back and started forward again, once more to stop and provide momentum for further car sorting. This kicking procedure is time-consuming, inefficient and somewhat labour-intensive. The back and forth movement has been called "drilling", may take an hour or more with minimal control of car speeds, occasionally and inadvertantly resulting in rough switching, frowned upon by management.

This paper describes a method of introducing the efficiency and control of the high-volume classification process into lower-volume flat yards without humps.

It is clear that, in order to separate a car from its followers, it must be accelerated away forward, so as to allow it to be directed into one of the variety of new tracks. This the hump performs at high rates. The same rate can be attained at the throat of a flat yard by means of a speed-controlled hydraulic ram positioned firmly beside the track on a solid base. It could be relocated with little effort. The ram to have an arm fastened to it that is rotated down through 90^{0} and introduced behind the coupler pocket of the car to be separated, which is immediately and smoothly accelerated away from its follower. After separation, the arm is rotated back up to its vertical position as the ram retracts, ready to be rotated down again to begin another stroke. The controller is stationed at the ram and chooses the point of separation.

All cars are uncoupled by a pin-puller working from a switch list, just as at a hump. As with the hump, the pushing yard engine never stops, with coupler slack bunched to allow uncoupling. The controller will activate the ram/arm consecutively, automatically sequenced; retraction is also automatic, as soon as the car has reached the chosen speed. The controller selects the radio-controlled "downstream" switch positions. These can be modified easily so as to operate like this (as long as not snow-blocked.

The 90° rotation of the cross arm from vertical and clear to horizontal and pushing, demands precision of position and of timing, so an optical apparatus must be mounted across the track so the arm is lowered only when light can pass between the cars. This prevents fouling a car by premature lowering.

Fig 5 shows the ram in the pushing position behind the car being separated.

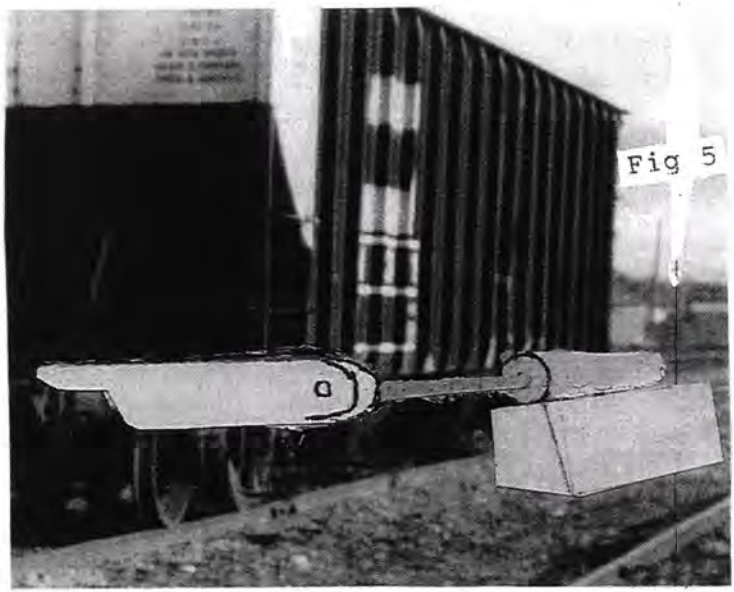
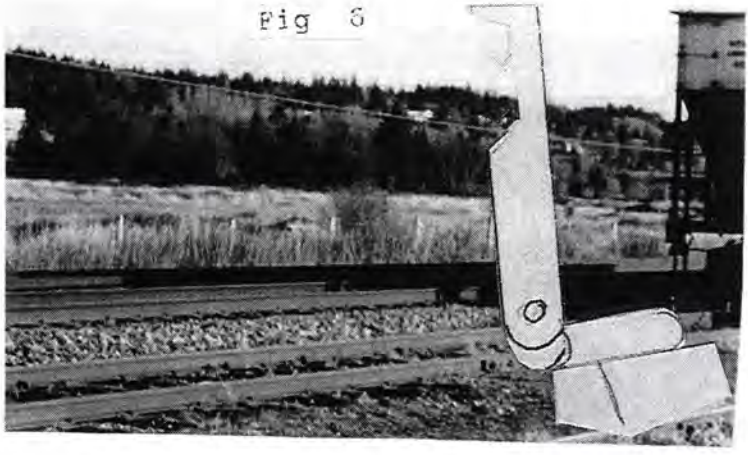


Fig 6 shows the ram fully withdrawn with the arm rotated back up, having separated the first car. It is ready to be advanced up to position for the next car.

Fig 6



In this way can classification be carried on in smaller flat yards with the same productivity and care as in large yards and with the same manpower. The number of cars to be coupled up, extracted and shunted is limited only by the yard engines tractive effort and the length of the lead track.

At major coal dumping wharves, where cars are brought up a steep incline and turned over, then rolled back down empty, a somewhat similar device is used. This is a "barney" running beside the track on a track of its own and pushing the car up. However, its return is not as smooth and precise as with this device.

The author acknowledges the assistance of Gordon Boal, William Woodcock P Eng, Brandt Road Rail Inc, Richard Lake P Eng, Finning Ltd, John Edsforth, BC Rail Rapid Transit Inc, L J Howard and N D Lea Consultants

1995 CCMTA National Roadside Survey Project

by Bill Harbour[1]

1. Introduction

In 1995 the Canadian Council of Motor Transport Administrators (**CCMTA**) sponsored a roadside survey of truck traffic. It was undertaken to produce a profile of the volume and characteristics of truck activity on Canada's highways, information needed for policy development and highway planning.

This paper presents a sample of the survey results (using **preliminary** data), and discusses aspects of the survey experience.

2. Background

The survey was carried out by officials of the federal, provincial and territorial governments under the direction of a Task Force established by the Standing Committee on Compliance and Regulatory Affairs of the CCMTA. Preparations for the survey began in late 1994. The data was collected in the summer and fall of 1995 and analyzed during 1996. The output of the survey was finalized in the spring of 1997.

---

[1] *Mr. Harbour is an analyst in the Motor Carrier Branch of Transport Canada.*

The views presented in this paper are those of the author.

The author would like to thank fellow members of the team processing the data collected in the survey: John Lawson, Jim Young, Glenn Robbins and Mark Smith of Transport Canada; and Abdellatif Demnati and Hew Gough of Statistics Canada.

The 1995 survey was the second CCMTA national roadside survey. Methodologically it was an improvement on the previous survey, which took place in 1991. In particular it:

- collected more characteristics of the traffic;
- produced an estimate of trip distance;
- was based on an improved sampling design; and
- used hand held computers for direct data entry.

In developing the CCMTA survey, officials were able to draw on the experience of the Ministry of Transportation of Ontario (**MTO**). MTO's 1995 Commercial Vehicle Survey (**CVS**) pioneered the use of hand held computers in a roadside survey. The CCMTA survey used the same type of computer equipment as selected by MTO, used software derived from the software developed for MTO, and used surveyor instructional material adapted from material prepared by MTO.[2]

3. Scope of the Survey

The output of the survey is an estimate of the truck traffic:

- traveling on the survey highway system;
- during a week (7 day total) in the summer/fall of 1995;
- for each of the 12 jurisdictions[3] and for Canada as a whole.

The survey highway system (**SHS**) consists of all roads that are part of the National Highway System[4] plus several other roads which are of significance to truck traffic in individual jurisdictions. The SHS, which is depicted on two maps in the annex of this paper, covers almost 26,000 kilometres of road.

---

[2] *The hand held computers were described in a paper titled "The Use of Hand-Held Computers in Roadside Surveys" that was presented at the 1996 CTRF conference by Greg Little and Doug Smith of MTO.*

[3] *A jurisdiction being a province or territory.*

[4] *The National Highway System is defined in the report National Highway Policy Study for Canada, Steering Committee Report on Phase 2, National Highway Policy Steering Committee, November 1989.*

The survey questionnaire collected information on 69 characteristics of the truck, driver, carrier, cargo and trip which are listed in the annex of this paper.

4. Survey Design

In designing the sampling of the population, a key objective was to maximize the number of survey sites and disperse the sites as widely as possible throughout the SHS. Location was expected to be the greatest source of variation in traffic characteristics. Constraining the number of sites that could feasibly be selected were the logistics and expense of moving survey crews and equipment.

Under the solution adopted, the SHS was divided into links, a link being the segment of highway joining two intersections of the SHS. Further, each traffic direction on a segment was treated separately, doubling the number of links. A survey site would be located on each link.

The survey collected data at 148 sites[5], the locations of which are depicted on the maps of the SHS in the annex of this paper.

The survey collected two types of information at each site. All trucks passing the site during a 7 day period were counted, with each truck categorized by day, time and configuration. A random sample of the trucks were interviewed to obtain truck, driver, carrier, cargo and trip characteristics.[6]

Ideally the sampling would occur continuously throughout the 7 day period. In practice this would not be feasible. For the survey the target

---

[5] *A survey site is defined by two features, its physical location (e.g., 2 miles west of Moose Jaw on Highway 1) and the traffic direction (e.g., trucks eastbound on the highway). If the survey covers both traffic directions at the same physical location, then there are two separate sites (e.g., Moose Jaw eastbound and Moose Jaw westbound).*

[6] *Randomness was approximated by pulling the first truck from the traffic stream that appeared after a surveyor became available to conduct the next interview.*

was to interview a random sample of trucks during a period covering at least 24 hours during the 7 day period.

The sample taken at each site was expanded to the total traffic count at the site, with individual interviews weighted according to their probability of being selected. When the expanded samples from the 148 individual sites were combined to form the national data file, the weight of each interview was adjusted to allow for the probability that the interview could have been observed at multiple sites.

Sampling errors were estimated following the design by Statistics Canada methodologists for the replication process of WesVarPC,[7] a computer program for survey analysis. For truck traffic estimates produced from the survey data, the sampling error is expressed as a conventional 95% confidence intervals (±1.96 standard errors).

5. Data Processing

5.1 Editing

The answers in each interview were examined for internal consistency. Inconsistent answers were corrected if possible, and marked as unusable if not correctable. Custom written computer programs identified the inconsistencies. The programs would automatically mark inconsistent answers as unusable. Corrections to answers, on the other hand, were defined on a case by case basis by an analyst.

Surveyors maintained a log book describing their operation (location, date, start and finish times, traffic direction, problems). These proved invaluable for editing the location, date, time and traffic direction answers in the interview, critical items in the calculation of the sample weight.

Standards for empty truck weight emerged as a key element in data editing. They were used when assessing whether a truck had cargo on-board, when assessing the reasonableness of the interview answers on

[7] *Produced by Westat, Inc.*

truck weight and on cargo weight, and when assigning a truck weight (assigned if the answer was not provided or unreasonable). Standards by configuration/body style/axle number category were derived from analysis of the 7,578 empty trucks captured in the interviews.

5.2 Route Analysis

In the route analysis, a commercial computer program called PC*MILER[8] was used to project the roads traveled by each truck interviewed. The location information recorded in the interview provided control points (trip origin, trip destination, survey site, and any international border crossings). PC*MILER used its built-in highway network information and route assignment algorithms to select a path between these locations.

Considerable value was added to the survey by the route analysis.

- It provided an estimate of trip length (categorized by jurisdiction of travel).
- The survey site traffic direction reported on the interview could be confirmed or corrected.
- It could be determined which of the 148 survey sites would be passed during the trip.

The sites passed information facilitated the merging of the expanded samples of the 148 individual sites into a single, national data file. Since the possibility of multiple counting of each trip was explicitly identified, the weight calculated for the truck sample could be adjusted to eliminate multiple counting.

The national data file increased the sample size for each jurisdiction. The effective sample consisted of the interviews actually collected in the jurisdiction plus the interviews collected elsewhere but inferred to travel in the jurisdiction. The impact is depicted below.

---

[8] *Produced by ALK Associates, Inc.. The program calculates route, distance, time and cost information, and is widely used in the trucking industry.*

Jurisdiction	Number of Survey Sites	Number of Interviews			Inferred as a Percentage of Collected
		Collected	Inferred	Effective (Total)	
Newfoundland	2	306	56	362	18.30
Prince Edward Island	4	466	75	541	16.09
Nova Scotia	6	332	614	946	184.94
New Brunswick	10	554	1,358	1,912	245.13
Quebec	21	3,604	5,538	9,142	153.66
Ontario	45	28,727	889	29,616	3.09
Manitoba	9	314	1,731	2,045	551.27
Saskatchewan	17	732	1,035	1,767	141.39
Alberta	14	469	1,435	1,904	305.97
British Columbia	14	292	727	1,019	248.97
Yukon	4	270	8	278	2.96
Northwest Territories	2	176	6	182	3.41
Total	148	36,242			

5.3 Ontario Data

The data collected by MTO in its 1995 CVS was used as the Ontario component of the CCMTA data. The CVS methodology is virtually identical to that of the CCMTA survey. The main difference is that several questions asked in the CCMTA questionnaire do not appear in the CVS questionnaire.

6. Data Analysis Package

The wide range of information collected by the survey makes it impractical to prepare a single report or even a series of reports that cover all potentially useful analyses. The national data file consisting of the 36,242 weighted interviews is the important product of the survey. With the file officials can perform the analyses needed for a specific policy development or highway planning exercise.

To that end a *Data Analysis Package* has been prepared. It consists of the national data file and a survey analysis program called WesVarPC.

With this package[9] a user can produce an estimate of truck traffic on the SHS within a particular jurisdiction or throughout Canada. Traffic:

- can be measured as: number of trips; distance traveled; cargo tonnage carried; tonne-kilometres of cargo moved; gross truck weight; or tonne-kilometres of truck moved; and
- can be tabulated by one or more of the 70 truck, driver, carrier, cargo and trip characteristics.[10]

For each traffic estimate the package produces an estimate of the sampling error, from which can be calculated a confidence interval for the traffic estimate.

7. Survey Results

Presented in this section are a sample of the results that can be generated with the *Data Analysis Package*. The results are based on the data file as of February 1997. Final editing of several characteristics and fine tuning of the sample weights had not been completed. Estimates produced from the final data file are not expected to differ significantly from estimates based on the **preliminary** data file, though the possibility can not be ruled out.

A total of 707.6 thousand truck trips (±2.6%)[11] occurred on the SHS during a summer/fall week in 1995. These trips traveled a total distance of 163.7 million kilometres (±5.5%) on the SHS, and moved 7.2 million tonnes of cargo (±5.4%). The transportation output on the SHS was 2.0 billion tonne-kilometres of cargo hauled (±4.7%).

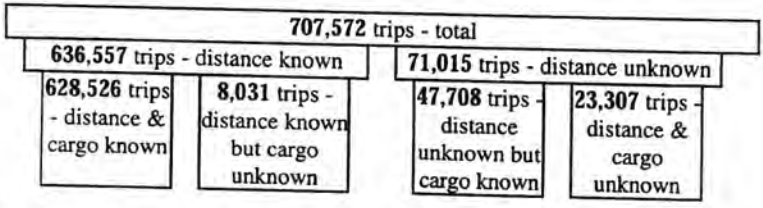
---

[9] *The package runs on an IBM compatible personal computer that has an Intel 80386 or better central processing unit with a math co-processor and a Windows 3.1 or better operating environment. The computer should have a minimum of 16 megabytes of memory, and due to the large size of the data file a hard disk with at least 125 megabytes of free space is required.*

[10] *To the 69 characteristics collected in the interview a 70th was added during the data processing, trip length.*

[11] *The percentage is the 95% confidence interval.*

Cargo tonnage, travel distance and transportation output are somewhat understated because answers to the cargo weight and trip distance questions were not available for all interviews. The dimensions of the problem are illustrated below. It shows, for example, that both trip distance and cargo weight are known for most of the expanded sample (88.8%).



Truck traffic was concentrated in Ontario, Quebec and New Brunswick according to the estimates for individual jurisdictions (see Table 1). When the length of highways covered in the survey is considered, New Brunswick had the highest intensity of truck activity (see Table 2).

Table 1 -- Truck Traffic Estimates

	Trips (thousands)	Distance (million kilometres)	Cargo (thousand tonnes)	Output (million tonne-kilometres)
Nfld.	2.9 ± 23.4 %	1.3 ± 33.0 %	36.2 ± 13.3 %	19.2 ± 35.6 %
PEI	7.3 ± 22.6 %	0.2 ± 35.3 %	78.0 ± 16.2 %	2.5 ± 53.0 %
NS	34.8 ± 18.3 %	5.2 ± 18.6 %	420.3 ± 9.7 %	63.1 ± 21.2 %
NB	120.2 ± 6.1 %	20.7 ± 8.1 %	1,361.8 ± 7.5 %	252.7 ± 16.5 %
Que.	160.1 ± 4.2 %	26.1 ± 5.6 %	1,580.9 ± 3.7 %	316.2 ± 9.0 %
Ont.	348.1 ± 1.8 %	65.2 ± 3.2 %	3,354.7 ± 1.9 %	778.1 ± 4.4 %
Mani.	25.9 ± 16.2 %	6.2 ± 18.1 %	315.3 ± 10.3 %	79.7 ± 18.7 %
Sask.	38.4 ± 19.1 %	10.8 ± 19.5 %	461.5 ± 11.3 %	144.7 ± 21.7 %
Alta.	52.6 ± 9.0 %	14.4 ± 13.6 %	696.9 ± 4.8 %	200.4 ± 10.5 %
BC	54.4 ± 20.5 %	13.3 ± 35.8 %	785.4 ± 16.3 %	171.1 ± 16.1 %
Yuk.	1.2 ± 17.6 %	0.2 ± 19.5 %	17.2 ± 11.2 %	3.6 ± 27.3 %
NWT	0.5 ± 19.6 %	0.1 ± 34.9 %	7.4 ± 11.6 %	1.4 ± 34.9 %
Can.	707.6 ± 2.6 %	163.7 ± 5.5 %	7,239.4 ± 2.7 %	2,033.5 ± 4.7 %

Table 2 – Truck Traffic per Kilometre of Survey Highways					
	SHS Length (kilometres	Trips (per SHS kilometre	Distance (truck-kilometres per SHS kilometre)	Cargo (tonnes per SHS kilometre)	Output (tonne-kilometres per SHS kilometre)
Nfld.	995	2.9	1,347.8	36.4	19,265.8
PEI	115	63.6	1,881.7	678.0	21,993.1
NS	1,062	32.8	4,875.1	395.7	59,392.5
NB	1,405	85.6	14,703.6	969.6	179,905.7
Que.	2,989	53.6	8,717.2	529.0	105,812.5
Ont.	5,323	65.4	12,255.5	630.2	146,173.0
Mani.	880	29.4	7,061.0	358.2	90,507.8
Sask.	2,208	17.4	4,888.6	209.0	65,559.1
Alta.	3,489	15.1	4,132.7	199.7	57,442.3
BC	5,621	9.7	2,357.5	139.7	30,450.3
Yuk.	1,008	1.2	242.4	17.0	3,598.1
NWT	564	0.8	169.8	13.1	2,533.4
Can.	25,658	27.6	6,381.8	282.1	79,253.1

The majority of the truck traffic was domestic, at least 71.4% (see Table 3). Transborder trips to/from the U.S. made up the rest of the identifiable trips. Transborder trips to/from Mexico were statistically insignificant. Of the Canadian "domestic" traffic, 1 in 83 trucks completed the trip by traveling through the U.S.. U.S. domestic traffic traveling through Canada made up 1 in 168 of all truck trips in Canada.

Trucks of the tractor and 1 trailer configuration made up the bulk of the trips on the SHS, 72.6% (see Table 4). This

Table 3 - Trip Sector Estimates		
Trip Sector	Trip Estimate	Distribution (% of Total)
IntraProvincial		
- Via Canada	410.9 ± 3.8	58.1
- Via U.S.	3.5 ± 50.4 %	0.5
InterProvincial		
- Via Canada	88.1 ± 7.1 %	12.4
- Via U.S.	2.6 ± 22.1 %	0.4
Transborder		
- To U.S.	77.9 ± 22.1 %	11.0
- From U.S.	61.3 ± 9.7 %	8.7
- To Mexico	0.0 ± 107.8 %	0.0
- From Mexico	0.1 ± 117.1 %	0.0
U.S. Domestic via Canada	4.2 ± 15.2 %	0.6
Unknown (but crossed border)	4.0 ± 80.0 %	0.6
Unknown	54.9 ± 10.9 %	7.8
Total	707.6 ± 2.6 %	100.0

configuration was the largest component of trips regardless of the trip length involved.

Table 4 -- Truck Traffic By Truck Configuration & Trip Length[12]							
Truck Configuration	Trip Length Range (kilometres)						Trip Length Distribution (%)
	0001-0250	0251-0750	0751-1500	1501+	Unknown	Total	
Tractor 1 Trailer	211.4	154.9	60.0	32.9	54.7	513.9	72.63
Tractor: 2 Trailer	16.2	17.2	5.5	1.3	5.4	45.6	6.45
Tractor; 3 Trailer	0.0	0.0	0.0	0.0	0.2	0.2	0.02
Sta. Truck	96.2	21.8	2.3	0.5	9.3	130.2	18.40
Str. Truck; 1 Trailer	7.0	2.1	0.3	0.1	0.8	10.3	1.46
Tractor Only	5.4	1.0	0.2	0.0	0.5	7.1	1.00
Other	0.1	0.0	0.0	0.0	0.1	0.3	0.04
Total	336.4	197.1	68.3	34.8	71.0	707.6	100.00
Trip Length Distribution (%)	47.54	27.85	9.65	4.92	10.04	100.00	

Tractor and 1 trailer configuration trucks were mainly van bodies, as were the straight trucks (see Table 5). Tractor and 2 trailer configuration trucks were spread among van, flatbed and tanker body styles.

Table 5 -- Body Style			
Body Style	Truck Configuration		
	Tractor & 1 Trailer	Tractor & 2 Trailers	Straight Truck
Van	62.0%	20.2%	65.5%
Flatbed	11.2%	25.9%	9.3%
Tanker	7.6%	21.7%	5.1%
Dump	4.2%	4.5%	10.4%
Other	14.9	27.6	9.2
Unknown	0.0%	0.0%	0.5%

One quarter of the trucks were traveling empty (25.3%) while one third were fully loaded with cargo (37.0%). When a truck was full, space available limited the load 46.6% of the time, weight allowable was the limit 22.7% of the time, and the load was both space and weight limited 30.7% of the time.

For-hire carriers' trucks generated 63.9% of the trips and private carriers' trucks 27.8%, a 2 to 1 split. The driver was more likely to own the truck if the carrier was

Table 6 -- Driver		
Driver Owns Truck	Driver's Carrier Is	
	For-Hire	Private
Yes	26.3%	12.8%
No	72.8%	86.1%
Unknown	1.0%	1.1%

[12] *The 95% confidence interval for the estimates has not been included due to the limited space in the table.*

for-hire rather than private (see Table 6). If the truck was operating for a private carrier and was carrying cargo, there was a 1 in 7 chance (13.6%) that it was carrying the cargo on a for-hire basis.

Two thirds (66.2%) of the drivers had training during the previous 3 years in one or more of 5 subject areas[13]. Dangerous goods handling was the most common subject (57.1%), followed by National safety Code requirements (37.3%), driving skills (33.8%), electronic equipment usage (12.7%) and business skills (9.1%). A significant portion of the drivers, 43.0%, had training in 2 or more areas, while 27.8% had training in 3 or more, 9.8% in 4 or more and 3.3% in all 5. Three quarters (78.8%) of drivers operating for for-hire carriers had training, a significantly higher portion than drivers operating for private carriers (58.9%).

8. Experience of the Survey

Driver Participation

Participation in the survey was voluntary. It took 8 minutes on average for the driver to answer the questions,[14] though the time depended on the specifics of the truck. Compared to an empty truck, a truck carrying cargo took 2 minutes longer to interview due to the additional questions related to the cargo. When the times taken to enter and exit the interview area and to weigh and measure the vehicle were added, a truck could spend a total of 10 to 15 minutes at the survey site

The cooperation received from the trucking industry was excellent. Overall 95% of the drivers agreed to the interview, and 6 of the 12 jurisdictions had a participation rate greater than 90%. The lowest rate was 47%.

---

[13] *These statistics exclude the interviews collected by MTO in its CVS. Driver training questions were not asked in the CVS.*

[14] *These times are derived from interviews conducted in jurisdictions other than Ontario.*

Hand Held Computers

Direct entry of information into hand held computers provided better quality data than filling out paper questionnaires. This was demonstrable because paper questionnaires were used as a back-up to the computers and 674 paper questionnaires were completed. There appeared to be two principle advantages with the computer.

- The logic built into the software ensured that all questions were answered, and prevented entry of inconsistent answers (in selected situations).
- Three data bases built into the software allowed the surveyor to select location, border crossing and company name answers from a list rather than write the answer (and potentially misspell or use local terminology).[15]

Additionally, because computer stored information avoided the computer coding of answers step that is required with paper questionnaires:

- a potential source of data errors was eliminated; and
- data processing could begin sooner.

The computers did not eliminate all data problems, and surveyors did not always take maximum advantage of the software features. Answers could still be incorrectly entered to questions that were not checked for consistency. Surprisingly frequently the surveyors chose to write an answer rather than select from the list.[16]

The computers were more difficult to use than expected. It took a surveyor several days to become proficient with the computer. The computer could not be used outside during rain, and operated erratically in extremely hot conditions or in cold conditions. Stability of the software also appears to have been a problem. Judging from the

---

[15] *The location data base contained 214,439 entries. The border crossings data base contained 138 entries. The company names data base contained 143,717 entries.*

[16] *During data editing 9,703 handwritten locations were electronically or manually matched to entries in the location database.*

computer files received from the surveyors, 1 out of 14 attempts[17] to enter an interview were interrupted by a computer failure for one reason or another, though some "failures" may have been intentional.[18]

9. Concluding Comments

The data analysis package produced from the 1995 CCMTA National Roadside Survey will provide transportation officials with a view of truck traffic that was previously unavailable. It offers a wide range of characteristics which can be interrelated, several measures of activity, and estimates of statistical confidence.

Future work may examine linking the survey data with other transportation and economic data bases, thereby extending the value of the survey product.

---

[17] *This rate is derived from interviews conducted in jurisdictions other than Ontario. MTO had fewer problems with the computers than experienced during the CCMTA exercise. Apparent contributing factors to MTO's success were: better weather (summer for MTO versus fall for CCMTA); more experienced surveyors; and a longer software testing period.*

[18] *Turning off the computer was the quickest way to break out of an interview since the software prevented the surveyor from terminating an interview until the questions were answered.*

AN EXPERIENCE IN MODELLING COMMERCIAL VEHICLE SURVEYS

Robert J. Tardif
Ministry of Transportation, Ontario *

Don L. McLeish
Department of Statistics & Actuarial Science
University of Waterloo

ABSTRACT

The Ministry of Transportation of Ontario conducted an inter-city roadside survey of commercial vehicles over a period of 167 days in 1995 at a cost of approximately $600,000. In total, 31,860 surveys were conducted at 68 sites throughout Ontario. Through a combination of; weighting each axle, observations and interviews, information was collected within five major areas; vehicle, trip, commodity, carrier and driver.

On completion of previous commercial vehicle surveys, a series of expansion factors were created using the counts collected at the time of survey. These factors merely expanded the survey sample to represent the total truck population observed at each survey station for the same period. Additional factors were then generated using broader time periods, daily, seasonal and annual. These were generally created in isolation of a host of potential problems, primarily related to bias in sampling, over counting in route assignment and without regard to trip travel time on the highway network. The data was analyzed using crude expansion factors. As a result, many of the estimates, even percentages, were inaccurate.

Realizing that flaws existed in the previous survey expansion methodology, the ministry embarked on a review of the survey design and analysis methodologies. The primary purpose of this paper is to explain how the survey expansion process was improved by adding additional trip routing questions and identifying the stations passed for each trip.

* The views expressed are those of the author and not necessarily those of the ministry.

Once additional routing information was collected, four Data Enhancement factors were created, they are described in greater detail within. The first two factors, accounting for station specific hour and daily truck population, are consistent with previous expansion methods. The remaining factors, remove sampling bias and over counting which hold the key to matching the count data for all survey locations involved, especially when all surveys are used.

The factors were evaluated within our modelling framework, where the results proved that no degree of error exists in the Data Enhancement process. However, deficiencies with the model were highlighted. The model is now the weak link in the process with an overall accuracy rate of 92% for representation of truck volumes at each station. The focus now turns from calibrating the Data Enhancement process to calibration of the model for future applications.

INTRODUCTION

In an effort to update and increase our knowledge of the trucking industry, the Ministry of Transportation of Ontario completed a roadside survey of commercial vehicles. The information collected from roadside surveys is viewed as being the most comprehensive and reliable method to obtain provincial wide information on trucking activity trip behaviour. These data are required for a variety of purposes, primarily for policy and planning initiatives. In total, 31,860 surveys were completed over 167 days in 1995. Prior to '95, as part of this survey program, a similar survey was conducted in 1993 and another of similar scale in 1988. In '95 surveys were primarily completed at permanent Truck Inspection Stations (97%) with the remaining 3% completed at temporary sites at strategic highway locations.

To initiate the survey, trucks were alerted by flashing lights and requested at random to enter the inspection station. Each of the axles were weighed and visual observations of the vehicle characteristics recorded. The driver was then signalled to bring the truck into the compound where the verbal portion of the survey took place; this was completed in approximately six minutes. Equipped with a PenDOS computer, the interviewer asked the driver a series of questions related to characteristics of driver, employment, trip, commodity and vehicle.

Of particular interest for this paper are the details surrounding the trip characteristics. The drivers were asked a series of detailed trip routing information, including: vehicle base, commodity origin, trip origin, border crossing on entry to Ontario, last stop prior to taking part in the survey, survey location, next stop immediately after the survey, exit border crossing, trip destination and commodity destination. The computer system used for the

survey contained an extensive database of placenames within North America. Along with the placename the corresponding longitude and latitude coordinates and a predefined zone system were simultaneously added to the survey record. If the trip was not a line haul (direct trip without pick-ups or drop-offs on route), and was of the peddle-run variety, the first seven stops were added.

This wealth of geographical data provides both challenges and opportunities. This origin-destination (OD) data is essential to accurately represent the survey sample trip routing within the model and most importantly to facilitate a revised Data Enhancement process. The term "Data Enhancement" refers to a series of measures which minimize survey bias, increase statistical reliability and expand the sample.

Given the cost of this massive effort to conduct the surveys and counts, it seemed appropriate to maximize our investment and increase the reliability of the data while extending the application of its use. The inspiration to develop a trucking model was spurred from the requirement to evaluate this relatively untested "Data Enhancement" process.

This Data Enhancement process was activated in the initial planning stages and carries over to the data expansion and model calibration stages. Thus, it was tied to the survey program from start to end. Initially, an assessment of the potential survey sites was undertaken. The selection criteria for optimal survey locations was based on the strategic level of truck information required, historical data characterising truck movements and daily truck count data. On average, seven 24-hour periods of detailed classification counts were collected for each survey site. Similar to the survey collection, detailed classification counts were collected throughout the summer survey period. As much as possible, we attempted to balance sample size for surveys and hours of counts for weekdays, weekends, nights and days.

On completion of field data collection of surveys and counts, a thorough review of the data was initiated. This review was not limited to merely reviewing and correcting obvious errors. In terms of the trip and commodity OD data, PC*MILER® routing software was used to verify the survey data and generate additional details to be added to the survey record. This detail includes: highways and local roads used (identifying the jurisdiction) and the sections used (from intersection entered to exited). Trip distances were then calculated for distance travelled in Ontario and outside of Ontario. The next and perhaps the most useful piece of data constructed was an inventory of the survey stations passed along the entire route. This inventory consists of a string of station

numbers in which this truck was not necessarily surveyed at, but rather the potential existed for the truck to be surveyed. This inference enables development of a trip relationship with all possible stations passed along the trip route. The code in the inventory is referred to as "SSS", the directional station number.

This paper deals with the advantages afforded by the Data Enhancement factors. The base model developed is more a mechanism to graphically present, test and validate results of the Data Enhancement process. Often in our applications, a model is developed only to emulate a provincial perspective from limited amount of OD data.

Counts were applied to the sample, first on an hourly basis then beyond, to day of survey, average day of week over the survey period, weekly, seasonal and annual. For analysis purposes, if only one survey site was involved this approach would be sufficient. However, this is not the case and an alternative process is required.

The goal in this survey was to provide an accurate measurements of truck movements through each of the 68 directional survey sites. In previous truck OD surveys, data were collected from a number of locations and a conventional model used to simulate the movements. Since the basic expansion factors were developed for hour, day and year for each station to represent the truck population at each site, it was considered logical that this would represent province wide commercial vehicle flows. This assumption was found to be false, as illustrated in Figures 1 and 2. Further, the duration of a truck trip on the network was never accounted for. In the past, the only time value used to expand the survey sample was the hourly count at the specific station in which the survey was collected and counted. In this exercise, the duration of the trip on the network and the relationship to the truck passing each of the other survey stations along the entire route is factored in. This is achieved by constructing a travel time and distance matrix for each time iteration. This hourly offset identifies the impact of the trip for each hour previous to and after the survey was conducted for the trip duration.

The two figures below best illustrate the scope of the benefit awarded by the Data Enhancement process. Figure 1, contains the desired flows at each survey station resulting from the Data Enhancement process and Figure 2 dramatically demonstrates the effects of double counting within the survey sample without the Data Enhancement process applied. It is obvious in Figure 2 that the actual truck volumes are misrepresented at all survey sites and that the order of

magnitude varies throughout the province. The surveys may represent trucking activity at each station but when all OD surveys are combined a province wide perspective becomes misrepresented. Figure 2 includes survey bias and over-counting. Previously, this was considered to be an accurate 24 hour snap-shot of provincial trucking activity, obviously it was not.
Figure 1

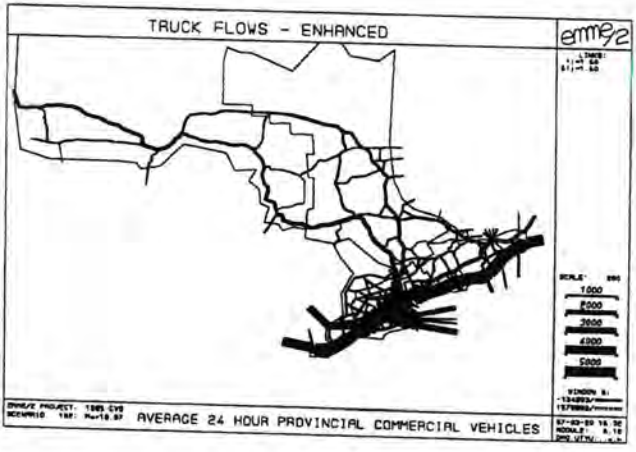
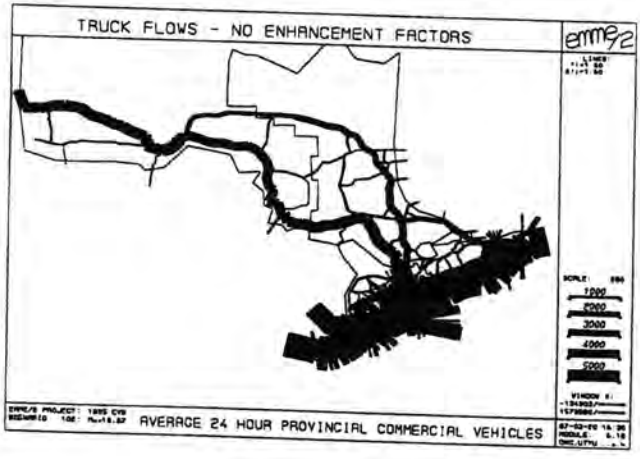


Figure 2



DATA ENHANCEMENT FACTORS

The Data Enhancement process involves the creation of four factors to eliminate bias in sampling and deal with over counting. The desired result is achieved by applying combinations of these factors. Applying sums of the factors to any of the data collected for analysis purposes adjusts the sample proportionally to the truck population. For all applications, including analysis of truck flows in modelling, the factors achieve an assignment value equal to the traffic count at each survey station.

Definition of DE, the Data Expansion Factor

We begin by determining the total truck volume at each station in a week. Let $V(ss,h)$ be the (estimated) truck traffic volume at station/direction combination ss and hour of week (HOW) h where $h = 0, 1, \ldots 167$ and $h = 0$ denotes Sunday, 12:01 a.m. - 1:00 a.m., $h = 167$ denotes Saturday 11:01 p.m. - 12:00. In an ideal world the predefined count schedule would have been carried out flawlessly. Unfortunately this was not the case. On review of the classification counts collected, we refer to these as ground counts, it became obvious that for relatively short time periods, automatic count systems failed or provided unrealistic data at some stations for various hours of the week. In order to overcome these deficiencies, a count model was used to identify errors through statistical analysis and to estimate traffic counts at station-hour combinations where no counts were taken.

In addition to loss of count data during the actual survey period unforseen and uncontrollable circumstances forced cancellations in the preplanned schedule. The causes were varied, ranging from sudden changes in staff schedules, adverse weather, prolonged power outages from lightening strikes at remote sites, and as the summer wore on the exhausted survey crews required additional days off.

The count model approximated the expected counts at all station-hour combinations. Counts at a given station were approximated by regressing on all counts at all stations in the same region, and then smoothing the residuals. Summing over all hours of the week, we obtain total weekly truck traffic volume

$$V(ss)=V(ss,0) + V(ss,1) + \ldots + V(ss,167).$$

Data Expansion Factor: DEF

The purpose of the hourly station data expansion factor DE is to expand all surveys at a given station and given hour of the week to all expected truck traffic at that station/direction/HOW combination. The data expansion factor

for a survey taken at station *ss* and HOW *h* is the ratio of the expected truck volume *V(ss,h)* at that station-HOW to the number of surveys collected at that station-HOW. Thus, if *NS (ss,h)* = total number of surveys collected at station *ss*, in
HOW = h, we define the data expansion factor for a record at *(ss,h)* to be

$$DE(ss,h) = V(ss,h) / N(ss,h).$$

In cases where the traffic counts were unavailable or unreliable, there was an additional possible source of information about counts available from the surveys themselves. In particular, if there are 10 surveys in a given HOW at a given station-direction, then it is reasonable to assume that the expected truck traffic volume at that station is greater than or equal to 10 (even though actual count data on some other week or estimated counts may have been less than 10). Thus, we ensure that all factors are greater than or equal to 1 (since the average volume of traffic is assumed at least equal to the number of surveys). Therefore, replace DE above by:

$$DE = min(1,DE) \text{ and } V(ss,h), NS(ss,h))$$

Definition of The Station Sampling Fraction

Let *V(ss)* be the total weekly truck volume at station *ss*. We define the weekly sampling fraction as the *proportion of weekly traffic sampled*. We take as an example one of the stations with the least surveying, station *ss* = 134. We recorded the following number of surveys by hour of week:

Hour of Week	Number of Surveys	Expected Truck Count
106	3	13
107	1	13
111	3	13
112	1	12
113	1	11

Adding the expected truck count over all 168 hours of the week we obtain the weekly count 1,582. These nine surveys need to be expanded over the whole weekly traffic. The weekly station sampling fraction is the ratio

$$SF(134) = 9/1582 = .005689$$

In general:

$$SF(ss) = \frac{\sum_{h=0}^{167} N(ss,h)}{V(ss)}$$

When a full week of surveys was not obtained, resulting in only a partial weeks coverage, an adjustment factor was created and applied to expand the surveyed hours to all 168 HOW. For example, at station 134, the factor DE expands to the surveyed hour, but there were only 5 of the 168 possible HOW in which surveys were taken. Thus, to expand to the weekly traffic, we need to multiply by the ratio of the expected traffic counts for the whole week to those in the hours surveyed. To expand to an average day, we must divide this again by 7, the number of days in a week.

$$DEF = 1582/7*(13+13+13+12+11) = 1582/7*62=25.516/7=3.645$$

Definition of OCF, The Overcount Correction Factor

The purpose of this factor is to adjust for the bias caused by the probability that a given trip could have been sampled at two or more locations *when records at all survey stations are combined.* Thus, although a given truck may be surveyed only once, if it passes several survey stations, *trips of this type* have a higher probability of appearing in the sample than do trips that pass a single station. Suppose, for example, a trip passes two stations, number 105 and 167, with sampling fraction at station 105 of 0.066 and at station 167 of 0.064. Then roughly .066/(.066+.064) =50.8% of these surveyed trip types will be surveyed at station 105, and 49.2% at station 167. For a survey conducted at station 105 with these properties, 50.8% is the OCF. Moreover, a trip type through two or more survey stations has a higher probability of being sampled in the survey as a whole than does one that only passes a single survey station. The OCF is defined for a given survey at station *ss*. Assume this trip type passes stations *sspass* on this trip. Then we define OCF by

$$OCF = \frac{SF(ss)}{\sum_{i \text{ in } sspass} SF(i)}$$

Notice that this factor only depends on the relative sampling fractions at the various stations, not on their magnitude. In other words, it is unaffected by computing sampling fractions on daily, weekly, monthly periods, or indeed computing them for the whole survey period.

Example: Record number 26814. Survey conducted at station 105. *sspass* = 105 167. There were 18 surveys conducted at this station in *HOW* = 69. The expected truck count at *HOW* = 69 is 147, so *DE* = 147/18 = 8.17. The sampling fraction at station 105 is 0.066 and at station 167 is 0.064. Therefore

$$OCF = 0.066/(.066 + .064) = 0.5077$$

Definition of OCFR, The Regional Overcount Correction Factor

The purpose of this factor is to adjust for the bias caused by the possibility that a given trip could have been sampled at two or more locations *when records within the corresponding region only are combined.* The justification for the factor is similar to that for OCF, only it assumes that for information at a given station say in region 2, we only wish to combine records surveyed in the same region 2.

The regions were defined by *ss* number as follows:

1. Southwest	102 103 109 111 112 117 142 148 149 150 151 158 165 166
2. Central	101 104 105 118 119 131 132 137 139 140 161 162 167
3. Eastern	100 108 115 116 120 121 122 123 125 126 135 136 138 141 163 164
4. Northern	106 107 110 124 127 128 133 134 143 144 147 152 153 159 160
5. Northwest	113 114 129 130 145 146 154 155 156 167 113 114 129 130 145 146

For example, for a station *ss* in the central region

$$OCFR = \frac{SF(ss)}{\sum_{i \text{ in sspass and } i \text{ in central}} SF(i)}$$

Example: Record number 8, station *ss* = 101, *sspass* = 138 121 101 132 149 151. This station is in the central region. Those stations in *sspass* that are also in the central region are stations 101 and 132. Their sampling fractions are respectively, 0.0406 and 0.0421. Therefore

$$OCFR = .0406/(.0406 + .0421) = .4909.$$

Summary of Factors

DE: used to expand all records at a given location/direction/hour to all traffic passing that station in the hour

DEF: expands all traffic in all hours surveyed to average daily traffic (multiply by 7 to get average weekly traffic)

OCF: overcount correction factor. Use when combining all surveys (in all regions) that have a given property P and have the given *ss* number anywhere in sspass field.

OCFR: overcount correction factor. Use when combining all surveys in a single region that have a given property P.

Truly, this survey was a departure from our common OD survey methodology. This extraordinary detail related to the routing opens many opportunities to accurately eliminate survey bias and network wide trip overcounting. Detail characteristics include, total trips, specific commodity flows, vehicle characteristics, driver characteristics, and just about any of the approximately 200 variables. As a result of the Data Enhancement process all trips represent a 24 hour period and most importantly, are systematically calibrated at each and every survey station.

DATA ENHANCEMENT VERIFICATION\MODEL CALIBRATION

The resulting OD matrix total after applying the Data Enhancement process, inclusive of the OCF stage expands the sample to represent approximately 51,000 truck trips for a 24 hour period. Once this matrix is assigned through the model onto the provincial road network with external centriod connectors, the combined volume at each station amounts to 91,000 trips. This is due to trips that pass more than one survey station and on the other hand, total trips are reduced as a result of the coarse zonal definitions and resulting inter-zonal trips.

As discussed earlier, the factors related to the Data Enhancement process are attached to every record. Using a combination of factors for the desired time reference any combination of flows can be realistically represented. The same process can be tested within the model. To verify this statement, provided in Table 1, are model results using Select Link analysis for each survey station and results generated from the survey data. In both cases the Data Enhancement factors have been applied.

Initially, the trip matrix was prepared through a typical approach, with each OD pair summed and weighted by the data enhancement factors. This matrix was then assigned to the provincial network as free-flows. Assignment results were compared against the 24 hour truck volumes at each station, border crossings and transportation boundary. For ease of repeating calibration procedures, survey stations, border crossings and transportation boundaries were tagged as screen lines. A screen line is defined as a series of points on a road network forming a cordon or net around a particular area of interest. For analysis purposes, this cordon is often consistent with existing political boundaries or other similar set of features.

Select Link analysis was performed to verify the routings at all screen lines identified. In addition, assignment volumes were checked against actual volumes. In general, initial model results were within 18%, but this was not satisfactory. When these system wide results were displayed graphically and compared to the border crossing movements, it became obvious that changes in the matrix preparation were necessary. In order to accurately represent border crossings and station flows, it was decided to prepare the matrix as if in two stages. In the first stage, a matrix was constructed from only intra-provincial flows. This was accomplished by coding Ontario zones between 1 and 60 and using only OD surveys with both origin and destination zones between 1 and 60. This leaves out any trips which, originate outside Ontario, and\or have destinations outside Ontario. For these trips, all border crossings on entry and exit were given unique zone numbers as gateways. Numbers ranged from 69 to 97. Regardless of differing origin zone numbers, as would be the case with a trip from a Western Province or California, but they share the same border crossing, the trip origin was coded as the crossing. This is an excellent method since a number of potential border crossing may exist for like OD pairs and now the trip matrix exactly replicates the each surveys routing, regardless of the assignment technique. Storing this revised OD zoning in a new field permits us to build matrices specifically for the model.

This matrix was then assigned in the same manner as in the first attempt. Verification of results was completed from screen line reports for the three types of screen lines. Overall, model results were improved to within 8% of ground counts or within 92% of the actual daily truck volumes for 68 sites throughout Ontario. Meanwhile, testing the Data Enhancement process based on the new model zone structure, but not running through the model, continues to reach the level of accuracy in matching ground counts at each survey station with 100%.

Unique to this methodology, from the survey database possesses the ability to construct a trip matrix for any station or combination of stations. The result of the Data Enhancement process ("Enhanced Survey"), will produce total flows identical to the 24 hour truck traffic counts ("Ground Counts") for each station. To verify these results I present the following table:

Table 1 - Comparison Between Counts\Data Enhancement\Model

STATION NAME	DIRECTION	GROUND COUNTS	ENHANCED SURVEY	ENHANCED MODEL	% DIF D/E	REAL DIFFERENCE
FORT ERIE	West	1,418	1,418	1,310	1.08	108
WINONA	East	4,100	4,100	4,495	0.91	(395)
FRUITLAND	West	3,934	3,934	4,628	0.85	(694)
OAKVILLE	East & West	11,254	11,254	10,969	1.03	285
THAMESFORD	East	432	432	52	8.31	380
CHATHAM	West	498	498	253	1.97	245
TILSONBURG	West	328	328	169	1.94	159
MORRISTON	North	934	934	285	3.28	649
FERGUS	South	249	249	238	1.05	11
GLEN TAY	East & West	734	734	938	0.78	(204)
NEW HAMBURG	East	370	370	437	0.85	(67)
VICTORIA	South	719	719	474	1.52	245
GRAVENHURST	South	829	829	649	1.28	180
WASI NORTH	North & South	1,056	1,056	847	1.25	209
NEW LISKEARD	North & South	830	830	573	1.45	257
COCHRANE	West	243	243	217	1.12	26
HEARST	East & West	494	494	368	1.34	126
RED ROCK	East & West	889	889	871	1.02	18
MORTON	South	201	201	160	1.26	41
KEMPTVILLE	North & South	690	690	263	2.62	427
HALEY STATION	East & West	870	870	439	1.98	431
NORTH BAY	West	422	422	103	4.10	319
NORTH SHORE	East & West	1,042	1,042	1,028	1.01	14
HEYDEN/SSM	North & South	507	507	612	0.83	(105)
THUNDER BAY	East & West	520	520	0	0	520
DRYDEN	East & West	957	957	793	1.21	164
WINCHESTER	North & South	278	278	12	23.1	266
DOVER CENTRE	North	279	279	113	2.47	166
ELDORADO	North & South	165	165	23	7.17	142
KEARNS	East & West	423	423	29	14.5	394
TIMMINS	East & West	366	366	255	1.44	111
THUNDER BAY	East & West	1,169	1,169	862	1.36	307
WOODBRIDGE	South	2,754	2,754	2,664	1.03	90
KING	North	2,571	2,571	2,728	0.94	(157)
CURRY HILL	East	2,181	2,181	2,111	1.03	70
LANCASTER	West	2,459	2,459	2,386	1.03	73
GANANOQUE	East & West	5,682	5,682	4,820	1.17	842
WHITBY	East	4,413	4,413	4,258	1.04	155
BOWMANVILLE	West	3,880	3,880	3,571	1.09	309
TRAFALGAR	East &West	12,361	12,361	12,855	0.96	(494)
PUTNAM	East & West	10,409	10,409	9,550	1.09	859
WINDSOR	East & West	5,779	5,779	5,875	0.98	(96)
SARNIA	East & West	3,538	3,538	3,342	1.06	196
BRANTFORD	East & West	4,156	4,156	3,497	1.19	659
CASSELMAN	West	743	743	685	1.08	58
					1.08	

The table above contains stations for which no volume in the Enhanced Model column while the Enhanced Surveys column provides the actual number of trucks that were recorded at that station. This is due to a combination of reasons, ie: competing parallel links where all the volume is assigned to the optimal route or the zone system in the matrix and network are too course thus incapable of replicating the real simulation. The fact remains, the Data Enhancement process provides better results than the model can produce.

However, it should be noted that the Data Enhancement process may provide excellent representation of the base situation, but its ability to simulate future year scenarios does not exist. Nor can changes in the actual road network the day after the survey be accurately reflected. In addition, the network in the model contains a vast number of highways not included in the survey. For these highways and highway sections between survey points the model provides us with a relative understanding of travel. Only with a model can we track the value of goods transported or set priorities for maintenance and capital planning or graphically identify strategic corridors.

Commodity Flow Validation

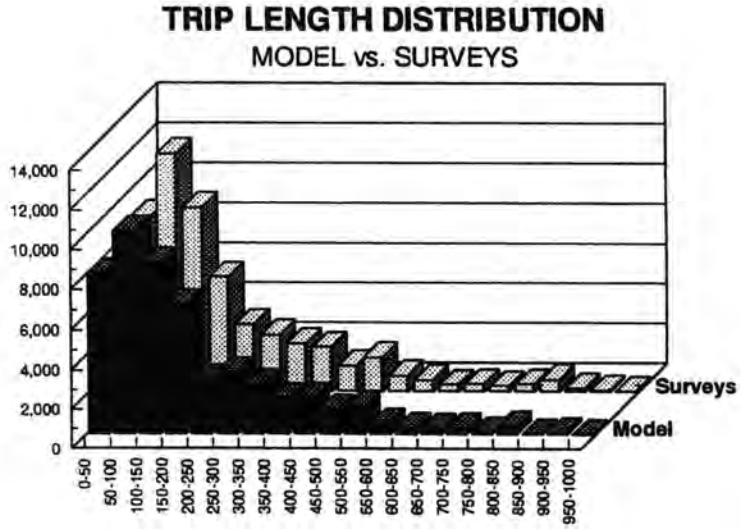
Table 2 - Comparison -Data Enhancement & Model Results - Food Commodity

STATION NAME	DIRECTION	ENHANCED SURVEY	ENHANCED MODEL	% DIF ES\EM	ABSOLUTE DIFFERENCE
FORT ERIE	West	109	92	1.18	17
WINONA	East	511	656	.77	(145)
FRUITLAND	West	635	745	.85	(110)
OAKVILLE	East	937	834	1.12	103
OAKVILLE	West	845	744	1.14	101
DRYDEN	West	103	93	1.11	10
THUNDER BAY N.	West	126	96	1.31	30
WOODBRIDGE	South	279	265	1.05	14
KING	North	372	388	.96	(16)
CURRY HILL	East	393	387	1.02	6
LANCASTER	West	279	274	1.02	5
GANANOQUE	West	487	429	1.14	58
GANANOQUE	East	364	313	1.16	51
WHITBY	East	501	473	1.06	28
BOWMANVILLE	West	884	839	1.05	45
TRAFALGAR	East	1047	1135	.92	(88)
TRAFALGAR	West	903	961	.94	(58)
PUTNAM	West	817	802	1.02	15
PUTNAM	East	729	710	1.03	19
WINDSOR	East	366	421	.87	(55)
WINDSOR	West	314	359	.87	(45)
SARNIA	West	139	165	.84	(26)
SARNIA	East	148	163	.91	(15)
BRANTFORD	East	354	250	1.42	104
BRANTFORD	West	229	213	1.08	16
CASSELMAN	West	112	101	1.11	11
				1.006	

We have established that the Data Enhancement process accurately matches ground counts. The next step, given that we require the model to estimate commodity related activity, is to perform a test with commodity data. The following table provides these results, a comparison of Model results and Survey data for food related products. In preparation of this table, trips carrying food related commodities were filtered into a separate database. Using the data subset, two parallel streams of comparison were undertaken.

The number of trips passing each station was generated from the Survey data with enhancement factors applied are provided in the column "Enhanced Survey". For comparison, a matrix was constructed from the same subset of data, also with the enhancement factors applied and assigned in the model in an "all-or-nothing" assignment, results are provided in column "Enhanced Model".

Another test was conducted to verify travel characteristics related to trip length distribution (TLD). The chart below illustrates and confirms the validity of the Data Enhancement process by checking against the TLD from the model. Given that in general the model results contain a degree of error estimated at 8%, these results are well within that range. This confirms that the model is producing logical results, especially when we consider that peddle trips are included, peddle trips account for 5% of all trips and have numerous stops and illogical routing in relation to the common line-haul OD pairs. In addition, the "all-or-nothing" assignment technique will not perfectly replicate routing diversion of congested urban areas. Even without these two explanations for possible error, the difference generated is less than 1%.



The accuracy of the results provided from the Data Enhancement process are not surprising, nor is the fact that model results are not as reliable. There are a number of explanations for this. The actual survey data contains much more information than can be easily replicated in the model. The model uses only a fraction of the routing information collected to construct the trip matrices.

The nature of the model contains a number of potential errors in terms of: no trip generation model, zonal aggregation problems causing excess intra-zonal trips, precision or simulation of volume delay functions, no distribution model or cell and trip end recalibration, not mode sensitive, auto and truck OD can be combined in one database, not sensitive to trip assignment techniques, DE process precalculates travel time and distance between all OD pairs, adjusting for not only at time of survey flow on network but potential impact on network over time. For all of these reasons how can a model better represent the base situation when the volumes at each survey station are matched.

Benefits of the Data Enhancement Process Are Passed to the Model

- Select link analysis was not conducted on a zonal basis but rather for each survey site and for a series of select links.
- In calibration the matrix was known prior to the assignment stage, double counting was eliminated in the Data Enhancement process.
- Building of trip matrices was not a simple process of summing OD pairs and applying weights
- Border crossings become the entry point regardless of origin zone, this forces all known trips through their respective border crossing rather than allowing an equilibrium assignment or gravity model distort the actual flows
- Data Enhancement factors remove double counting basis for all records at each station.

In conventional models with OD data, the model itself is calibrated as an end product producing an OD matrix resembling flows. Without internal select link analysis only trip ends are calibrated. The information collected in the '95 CVS is not limited to OD and routing information. In fact OD questions relate to only 10% of the survey. In a conventional model the calibration achieved remains within the model and is not automatically related back into the flat data file. Since the Data Enhancement process remains tied to the flat data file, any data element enjoys the same benefits. In addition, each data element becomes tied to the trip. In the conventional model this relationship is very difficult and cumbersome to obtain. In addition, most OD surveys contain only origin and destination.

The additional trip information collected allows creation of multiple trip segment matrices, ie: trip origin to border-crossing, border crossing to survey station. A combination of 12 OD matrices can be constructed while maintaining the flexibility to apply the Data Enhancement factors. The resulting combined matrix more closely replicates each segment of the trip. Assigning this matrix produces a improved association with "U" Trips, those trips which are of shorter duration and not captured in this inter-city survey.

FUTURE SURVEY PROGRAM AND MODEL DEVELOPMENT

The ability exists to identify each stations relative strategic importance while using existing base year information to pre-test the sensitivity of policies and capital expenditures. Thus, the survey program can now be streamlined.

In terms of model development, given that the truck volumes are accurately represented at many screen-lines. It is now desirable to examine economic relationships. These involves developing a model to represent the value-of-goods transported while matching with Import\Export values, GDP and GNP trucking activity indices.

ACKNOWLEDGEMENTS

The authors wish to thank Michele Babin of the Freight and Economic Research Office, MTO for her contributions.

BIBLIOGRAPHY

1. Don L. McLeish. 1994. *Report on The 1993 Pilot Commercial Vehicle Study*

2. Don L. McLeish. 1997. *Data Enhancement Process Documentation For The 1995 Ontario Commercial Vehicle Survey.*

OPERATING COSTS OF TRUCKS IN CANADA - 1996: A SYNOPSIS

by Lloyd L. Ash, P.Eng., P.Log, Trimac Consulting Services Inc.
and Barry Gardner, Transport Canada Motor Carrier Policy

The 1996 edition of Operating Costs of Trucks in Canada is the fifteenth of a series previously published since 1972. The project is sponsored by Transport Canada and is, starting with the 1995 edition, available over the Internet. Historically, this publication has been a widely cited reference on trucking industry productivity for academics, highway planners, and others within the consulting / shipper / carrier community.

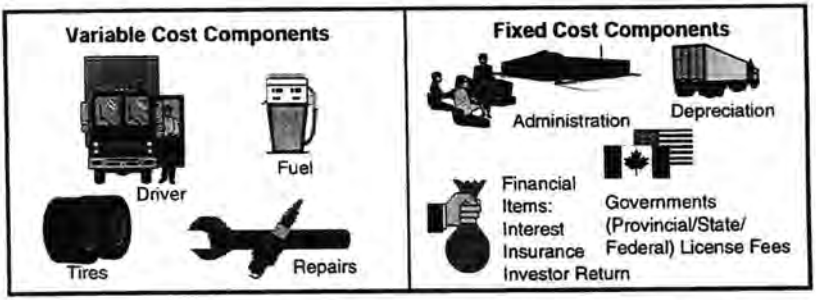


Our paper is a synopsis of highlights from the 1996 study. For common 5, 6, and 8 axle truck combinations, the study provides activity based operating cost estimates in various jurisdictions. These include each Canadian province and territory, five U.S. regions, and several interregional corridors. Cost levels reflect mid year (July 1996) prices for inputs to the trucking industry, within the major population centers in each region, for a medium sized commercial fleet operator. As well as highlighting the over all cost findings from the study, this synopsis reviews major underlying cost trends for wages, fuel, and equipment purchase cost. We also compare cost for a Canadian versus a U.S. based trucker operating on transborder corridors. As this paper is intended only as a synopsis, readers requiring further detailed background information are referred to the Transport Canada website at http://www.tc.gc.ca for the full study report. In the full report, additional intermodal cost structure

comparisons are made and there is more detailed discussion of costing assumptions and basis for analysis.

METHODOLOGY

In this project, a trucking company "rate making model" is used that computes cost as a sum of fixed and variable components in relation to standardized activity based case studies for annual kilometers driven, hours driving, and hours loading / unloading freight. Input cost components reflect a regional database of factor inputs derived from actual trucking company records and consultations, published and unpublished statistics, and consultation with suppliers to the trucking industry for fuel, commercial shop services, equipment, insurance, and financing.



In each detailed case study, the cost for operating one vehicle, as part of a medium sized fleet, is presented in terms of the foregoing components (and others) as total annual costs in Canadian dollars, costs per operating hour, and costs per kilometer. A significant variable, especially for allocating fixed costs such as vehicle depreciation costs, is the annual utilization of the vehicle. Base case studies are provided for:

Low Utilization	80,000 km/year	50,000 mi/year
Medium Utilization	160,000 km/year	100,000 mi/year
High Utilization	240,000 km/year	150,000 mi/year

OVER ALL COST RESULTS

Tables on the following pages present the derived costs for each of the intra regional case studies as cents per kilometer. Note that costs of trucking (including provision for "return on investment" for the business owner) are portrayed for operating ratios of 90, 95, and 97.5 (margins of

10%, 5%, and 2.5% of revenue, respectively). Principal factors that lead to differences in derived operating costs for each provincial and territorial region are:

- Differences in the local wage markets. Wage costs shown are equivalent to union rate levels in each region, with wage burden applied to cover costs for paid holidays, contributions to pension and group insurance, medical coverages, workers compensation, and other "payroll taxes", as applicable. Note that highest Canadian wage levels are in British Columbia (the lower mainland), Ontario, and Quebec. Mid range wage levels occur in the prairie provinces (Alberta, Saskatchewan, Manitoba). Wage levels are lower in the eastern Canadian provinces (Nova Scotia, New Brunswick, Prince Edward Island, and Newfoundland). Direct wage levels are reflected primarily in driver wage costs and the mechanic wage component of repair costs, but are also reflected in regional administrative cost levels.

All costs in cents per km

	British Columbia			Alberta		
Configuration	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)						
80,000 km	182.1	172.5	166.1	155.9	147.7	143.9
160,000 km	166.2	151.8	147.9	135.8	128.6	125.3
240,000 km	152.9	144.9	141.1	129.1	122.3	119.1
5 Axle Semi Unit (Flat Deck)						
80,000 km	185.8	176.1	171.6	158.8	151.4	147.5
160,000 km	165.7	156.9	152.9	141.2	133.8	130.4
240,000 km	158.9	150.6	146.7	135.0	127.9	124.7
5 Axle Bulk Liquid Tanker						
80,000 km	175.4	166.2	161.9	147.3	139.6	136.0
160,000 km	149.4	141.6	138.0	123.4	116.9	113.9
240,000 km	140.8	133.4	130.0	115.4	109.3	106.5
5 Axle Bulk Dry Tanker						
80,000 km	180.5	171.0	166.6	152.0	144.0	140.4
160,000 km	152.1	144.1	140.4	125.8	119.2	116.1
240,000 km	142.6	135.1	131.6	117.0	110.9	108.0
6 Axle (Triaxle) Semi Unit (Van)						
80,000 km	209.4	198.4	193.3	179.3	169.8	165.5
160,000 km	185.1	175.3	170.8	156.8	148.7	144.9
240,000 km	177.0	167.7	163.4	149.5	141.6	138.0
6 Axle (Triaxle) Semi Unit (Flat Deck)						
80,000 km	205.4	194.6	189.6	176.5	167.2	162.9
160,000 km	182.9	173.3	168.8	155.8	147.6	143.8
240,000 km	175.4	166.2	161.9	148.9	141.1	137.5
8 Axle Super B Train Unit (Van)						
80,000 km	232.1	219.9	214.2	195.2	184.9	180.2
160,000 km	202.6	192.0	187.0	168.2	159.3	155.2
240,000 km	192.8	182.7	178.0	158.1	150.8	146.9
8 Axle Super B Train Unit (Flat Deck)						
80,000 km	240.0	227.3	221.5	200.5	190.0	185.1
160,000 km	213.7	202.4	197.2	178.4	167.2	162.9
240,000 km	204.9	194.1	189.1	168.4	159.5	155.5
8 Axle Super B Bulk Liquid Tanker						
80,000 km	227.6	215.8	210.1	188.9	178.9	174.3
160,000 km	187.4	177.8	173.0	151.8	143.8	140.1
240,000 km	174.0	164.9	160.6	139.4	132.1	128.7
8 Axle Super B Bulk Dry Tanker						
80,000 km	226.3	214.4	208.9	187.6	177.7	173.2
160,000 km	186.9	177.0	172.5	151.2	143.3	139.6
240,000 km	173.7	164.6	160.4	139.1	131.8	128.4

All costs in cents per km

Configuration	Saskatchewan 90 Operating Ratio Total Costs (c/km)	Saskatchewan 95 Operating Ratio Total Costs (c/km)	Saskatchewan 97.5 Operating Ratio Total Costs (c/km)	Manitoba 90 Operating Ratio Total Costs (c/km)	Manitoba 95 Operating Ratio Total Costs (c/km)	Manitoba 97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)						
80,000 km	155.6	147.6	143.8	149.6	141.7	138.1
160,000 km	133.3	126.3	123.1	127.6	120.9	117.8
240,000 km	125.9	119.2	116.2	120.3	114.0	111.1
5 Axle Semi Unit (Flat Deck)						
80,000 km	156.1	147.9	144.1	151.8	143.8	140.1
160,000 km	135.4	128.2	124.9	131.6	124.6	121.4
240,000 km	128.4	121.7	118.6	124.8	118.2	115.2
5 Axle Bulk Liquid Tanker						
80,000 km	146.2	138.5	135.0	143.6	136.1	132.6
160,000 km	119.6	113.3	110.4	117.6	111.4	108.5
240,000 km	110.7	104.9	102.2	108.9	103.2	100.5
5 Axle Bulk Dry Tanker						
80,000 km	151.3	143.3	139.6	148.6	140.8	137.2
160,000 km	122.2	115.8	112.8	120.1	113.8	110.9
240,000 km	112.5	106.6	103.8	110.7	104.8	102.1
6 Axle (Triaxle) Semi Unit (Van)						
80,000 km	176.6	167.3	163.0	172.4	163.3	159.1
160,000 km	152.0	144.0	140.3	148.0	140.3	136.7
240,000 km	143.9	136.3	132.8	139.9	132.6	129.2
6 Axle (Triaxle) Semi Unit (Flat Deck)						
80,000 km	173.4	164.3	160.1	169.3	160.4	156.3
160,000 km	150.8	142.9	139.2	146.8	139.1	135.5
240,000 km	143.2	135.7	132.2	139.3	132.0	128.6
8 Axle Super B Train Unit (Van)						
80,000 km	196.7	186.3	181.5	191.6	181.6	176.9
160,000 km	166.5	157.8	153.7	162.0	153.5	149.6
240,000 km	156.5	148.3	144.5	152.2	144.2	140.5
8 Axle Super B Train Unit (Flat Deck)						
80,000 km	201.1	190.6	185.7	195.9	185.6	180.8
160,000 km	174.2	165.1	160.8	169.4	160.5	156.4
240,000 km	165.3	156.6	152.5	160.6	152.2	148.3
8 Axle Super B Bulk Liquid Tanker						
80,000 km	192.3	182.2	177.5	188.3	178.4	173.8
160,000 km	151.3	143.3	139.7	148.0	140.2	136.6
240,000 km	137.6	130.4	127.0	134.5	127.4	124.2
8 Axle Super B Bulk Dry Tanker						
80,000 km	190.9	180.9	176.2	186.9	177.0	172.5
160,000 km	150.7	142.7	139.1	147.3	139.6	136.0
240,000 km	137.2	130.0	126.7	134.1	127.1	123.8

Configuration	Ontario 90 Operating Ratio Total Costs (c/km)	Ontario 95 Operating Ratio Total Costs (c/km)	Ontario 97.5 Operating Ratio Total Costs (c/km)	Quebec 90 Operating Ratio Total Costs (c/km)	Quebec 95 Operating Ratio Total Costs (c/km)	Quebec 97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)						
80,000 km	161.9	153.3	149.4	151.3	143.4	139.7
160,000 km	140.1	132.7	129.3	129.5	122.7	119.6
240,000 km	132.8	125.8	122.6	122.3	115.9	112.9
5 Axle Semi Unit (Flat Deck)						
80,000 km	169.2	160.3	156.2	160.2	151.8	147.9
160,000 km	149.1	141.2	137.6	140.1	132.7	129.3
240,000 km	142.4	134.9	131.4	133.4	126.3	123.1
5 Axle Bulk Liquid Tanker						
80,000 km	165.4	156.7	152.6	158.7	150.4	146.5
160,000 km	139.4	132.1	128.7	132.8	125.8	122.6
240,000 km	130.8	123.9	120.7	124.2	117.7	114.6
5 Axle Bulk Dry Tanker						
80,000 km	170.5	161.5	157.3	163.8	155.1	151.2
160,000 km	142.1	134.6	131.1	135.4	128.3	125.0
240,000 km	132.6	125.6	122.4	126.0	119.4	116.3
6 Axle (Triaxle) Semi Unit (Van)						
80,000 km	192.2	182.0	177.4	183.2	173.6	169.1
160,000 km	168.0	159.2	155.1	158.8	150.5	146.6
240,000 km	160.0	151.6	147.7	150.7	142.8	139.1
6 Axle (Triaxle) Semi Unit (Flat Deck)						
80,000 km	189.1	179.2	174.6	180.1	170.6	166.2
160,000 km	166.9	158.1	154.0	157.5	149.2	145.4
240,000 km	159.5	151.1	147.2	150.0	142.1	138.5
8 Axle Super B Train Unit (Van)						
80,000 km	199.4	188.9	184.0	189.0	179.0	174.4
160,000 km	170.6	161.7	157.5	160.3	151.9	148.0
240,000 km	161.1	152.6	148.7	150.8	142.9	139.2
8 Axle Super B Train Unit (Flat Deck)						
80,000 km	207.5	196.5	191.5	191.4	181.4	176.7
160,000 km	181.9	172.3	167.9	166.0	157.2	153.2
240,000 km	173.4	164.3	160.1	157.5	149.2	145.4
8 Axle Super B Bulk Liquid Tanker						
80,000 km	209.1	198.1	193.0	200.0	189.4	184.6
160,000 km	169.6	160.6	156.5	160.6	152.2	148.3
240,000 km	156.4	148.1	144.3	147.5	139.8	136.2
8 Axle Super B Bulk Dry Tanker						
80,000 km	207.8	196.8	191.8	198.6	188.2	183.3
160,000 km	169.0	160.1	156.0	160.1	151.6	147.7
240,000 km	156.1	147.8	144.1	147.2	139.4	135.9

All costs in cents per km

Configuration	New Brunswick 90 Operating Ratio Total Costs (c/km)	New Brunswick 95 Operating Ratio Total Costs (c/km)	New Brunswick 97.5 Operating Ratio Total Costs (c/km)	Nova Scotia 90 Operating Ratio Total Costs (c/km)	Nova Scotia 95 Operating Ratio Total Costs (c/km)	Nova Scotia 97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)						
80,000 km	143.9	136.4	132.9	141.0	133.6	130.2
160,000 km	121.3	114.9	112.0	118.9	112.6	109.7
240,000 km	113.7	107.8	105.0	111.5	105.6	102.9
5 Axle Semi Unit (Flat Deck)						
80,000 km	145.5	137.9	134.3	142.6	135.1	131.6
160,000 km	124.6	118.1	115.0	122.2	115.7	112.8
240,000 km	117.6	111.4	108.6	115.4	109.3	106.5
5 Axle Bulk Liquid Tanker						
80,000 km	147.9	140.1	136.5	145.4	137.7	134.2
160,000 km	120.9	114.5	111.6	118.9	112.6	109.7
240,000 km	111.9	106.0	103.3	110.1	104.3	101.6
5 Axle Bulk Dry Tanker						
80,000 km	153.0	145.0	141.3	150.5	142.6	138.9
160,000 km	123.6	117.1	114.1	121.6	115.2	112.2
240,000 km	113.7	107.8	105.0	111.9	106.0	103.3
6 Axle (Triaxle) Semi Unit (Van)						
80,000 km	165.8	157.1	153.1	165.8	157.1	153.0
160,000 km	140.8	133.4	130.0	141.2	133.8	130.4
240,000 km	132.5	125.5	122.3	133.0	126.0	122.8
6 Axle (Triaxle) Semi Unit (Flat Deck)						
80,000 km	162.3	153.8	149.8	162.2	153.7	149.8
160,000 km	139.2	131.9	128.5	139.6	132.3	128.9
240,000 km	131.5	124.6	121.4	132.1	125.1	121.9
8 Axle Super B Train Unit (Van)						
80,000 km	179.5	170.1	165.7	176.5	167.2	162.9
160,000 km	149.7	141.8	138.1	147.3	139.6	136.0
240,000 km	139.7	132.3	128.9	137.6	130.4	127.0
8 Axle Super B Train Unit (Flat Deck)						
80,000 km	180.9	171.3	167.0	175.1	165.9	161.6
160,000 km	154.3	146.2	142.4	149.2	141.4	137.8
240,000 km	145.4	137.8	134.2	140.6	133.2	129.8
8 Axle Super B Bulk Liquid Tanker						
80,000 km	191.9	181.8	177.1	188.0	178.1	173.6
160,000 km	150.9	142.9	139.3	147.7	139.9	136.3
240,000 km	137.2	130.0	126.6	134.3	127.2	123.9
8 Axle Super B Bulk Dry Tanker						
80,000 km	190.5	180.4	175.8	186.6	176.8	172.2
160,000 km	150.2	142.3	138.6	147.0	139.3	135.7
240,000 km	136.8	129.6	126.3	133.9	126.8	123.6

Configuration	Prince Edward Island 90 Operating Ratio Total Costs (c/km)	Prince Edward Island 95 Operating Ratio Total Costs (c/km)	Prince Edward Island 97.5 Operating Ratio Total Costs (c/km)	Newfoundland 90 Operating Ratio Total Costs (c/km)	Newfoundland 95 Operating Ratio Total Costs (c/km)	Newfoundland 97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)						
80,000 km	133.9	126.9	123.6	144.3	136.7	133.2
160,000 km	112.1	106.2	103.5	121.9	115.5	112.5
240,000 km	104.9	99.3	96.8	114.4	108.4	105.6
5 Axle Semi Unit (Flat Deck)						
80,000 km	134.8	127.7	124.4	143.6	136.1	132.6
160,000 km	114.7	108.6	105.9	123.0	116.5	113.5
240,000 km	108.0	102.3	99.7	116.1	110.0	107.2
5 Axle Bulk Liquid Tanker						
80,000 km	139.3	131.9	128.5	152.3	144.3	140.6
160,000 km	113.2	107.3	104.5	125.6	119.0	116.0
240,000 km	104.5	99.0	96.5	116.7	110.6	107.8
5 Axle Bulk Dry Tanker						
80,000 km	144.3	136.7	133.2	157.5	149.2	145.4
160,000 km	115.8	109.7	106.9	128.3	121.6	118.4
240,000 km	106.3	100.7	98.1	118.6	112.3	109.5
6 Axle (Triaxle) Semi Unit (Van)						
80,000 km	154.4	146.3	142.6	164.6	155.9	151.9
160,000 km	130.4	123.5	120.3	139.8	132.5	129.1
240,000 km	122.3	115.9	112.9	131.6	124.7	121.5
6 Axle (Triaxle) Semi Unit (Flat Deck)						
80,000 km	150.8	142.9	139.2	160.9	152.4	148.5
160,000 km	128.7	121.9	118.8	138.1	130.8	127.5
240,000 km	121.3	114.9	112.0	130.5	123.6	120.4
8 Axle Super B Train Unit (Van)						
80,000 km	170.3	161.4	157.2	177.7	168.3	164.0
160,000 km	141.3	133.9	130.5	148.1	140.3	136.7
240,000 km	131.7	124.8	121.6	138.2	130.9	127.6
8 Axle Super B Train Unit (Flat Deck)						
80,000 km	168.2	159.3	155.3	178.8	169.4	165.1
160,000 km	142.5	135.0	131.5	152.5	144.5	140.8
240,000 km	133.9	126.9	123.6	143.8	136.2	132.7
8 Axle Super B Bulk Liquid Tanker						
80,000 km	182.0	172.5	168.0	196.7	186.4	181.6
160,000 km	142.1	134.6	131.1	156.0	147.7	144.0
240,000 km	128.8	122.0	118.8	142.4	134.9	131.4
8 Axle Super B Bulk Dry Tanker						
80,000 km	180.6	171.1	166.7	195.3	185.0	180.3
160,000 km	141.4	134.0	130.5	155.3	147.1	143.4
240,000 km	128.3	121.6	118.5	142.0	134.5	131.0

All costs in cents per km

Configuration	Yukon 90 Operating Ratio Total Costs (c/km)	Yukon 95 Operating Ratio Total Costs (c/km)	Yukon 97.5 Operating Ratio Total Costs (c/km)	Northwest Territories 90 Operating Ratio Total Costs (c/km)	Northwest Territories 95 Operating Ratio Total Costs (c/km)	Northwest Territories 97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)						
80,000 km	185.8	176.0	171.5	168.4	160.5	156.4
160,000 km	166.1	157.4	153.3	149.7	141.8	138.2
240,000 km	159.5	151.1	147.2	143.2	135.6	132.1
5 Axle Semi Unit (Flat Deck)						
80,000 km	189.7	179.7	175.1	170.9	161.8	157.7
160,000 km	171.6	162.5	158.4	152.7	144.7	141.0
240,000 km	165.5	156.6	152.8	146.8	138.9	135.4
5 Axle Bulk Liquid Tanker						
80,000 km	172.4	163.3	159.1	151.7	143.7	140.1
160,000 km	148.9	141.0	137.4	128.2	121.4	118.3
240,000 km	141.0	133.6	130.2	120.3	114.0	111.1
5 Axle Bulk Dry Tanker						
80,000 km	177.1	167.8	163.5	156.4	148.2	144.4
160,000 km	151.3	143.3	139.7	130.8	123.7	120.6
240,000 km	142.7	135.2	131.7	122.0	115.6	112.6
6 Axle (Tridem) Semi Unit (Van)						
80,000 km	216.0	204.6	199.4	192.2	182.1	177.4
160,000 km	194.2	184.0	179.2	170.5	161.5	157.4
240,000 km	186.9	177.1	172.5	163.3	154.7	150.7
6 Axle (Triaxle) Semi Unit (Flat Deck)						
80,000 km	213.2	202.0	196.8	189.4	179.4	174.8
160,000 km	193.2	183.0	178.3	169.4	160.5	156.4
240,000 km	186.5	176.7	172.1	162.7	154.2	150.2
8 Axle Super B Train Unit (Van)						
80,000 km	235.4	223.0	217.3	221.3	209.7	204.3
160,000 km	209.4	198.4	193.3	195.3	185.0	180.2
240,000 km	200.7	190.2	185.3	186.6	176.8	172.2
8 Axle Super B Train Unit (Flat Deck)						
80,000 km	238.0	225.5	219.7	218.2	206.7	201.4
160,000 km	215.0	203.7	198.5	195.1	184.9	180.1
240,000 km	207.3	196.4	191.4	187.4	177.5	173.0
8 Axle Super B Bulk Liquid Tanker						
80,000 km	219.3	207.7	202.4	196.4	186.1	181.3
160,000 km	183.3	173.7	169.2	160.4	151.9	148.0
240,000 km	171.3	162.3	158.1	148.3	140.5	136.9
8 Axle Super B Bulk Dry Tanker						
80,000 km	218.0	206.6	201.3	195.1	184.9	180.1
160,000 km	182.7	173.1	168.7	159.9	151.4	147.5
240,000 km	171.0	162.0	157.8	148.0	140.2	136.6

Principal factors leading to regional differences, contd.:

- Differences in fuel prices in each region. Note that significant differences in fuel prices occur within each region. The cost assessed for our case study was the local fuel cost, including all applicable provincial taxes and bulk purchase discounts, for the major population center of each region.
- Equipment purchase costs in each region vary according to applicable provincial sales taxes.
- Licensing costs vary by jurisdiction.
- Repair costs reflect (as noted before) local wage markets as well as costs for shop facilities as reflected in commercial shop rates that are generally applicable in the major urban centers of each region.

The following page tabulates local intra regional cost results for the five United States regions. As before, all costs are in cents per kilometer (converted to Canadian funds at the applicable mid 1996 currency exchange rate of 1$ Cdn = 0.722 cents per U.S. $).

All costs in cents per km

	U.S. Northeast			U.S. Great Lakes		
Configuration	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)						
80,000 km	186.9	179.0	174.4	186.2	176.4	171.8
160,000 km	162.1	153.8	149.7	160.1	151.7	147.8
240,000 km	153.2	145.1	141.4	151.4	143.4	139.8
5 Axle Semi Unit (Flat Deck)						
80,000 km	182.7	182.6	177.9	190.1	180.1	175.5
160,000 km	167.4	158.6	154.5	165.8	156.8	152.8
240,000 km	158.9	150.6	146.7	157.4	149.1	145.3
5 Axle Bulk Liquid Tanker						
80,000 km	175.7	166.5	162.2	172.4	163.3	159.1
160,000 km	144.2	136.6	133.1	141.4	133.9	130.5
240,000 km	133.7	126.6	123.4	131.1	124.2	121.0
5 Axle Bulk Dry Tanker						
80,000 km	181.3	171.7	167.3	178.0	168.7	164.3
160,000 km	147.0	139.3	135.7	144.3	136.7	133.2
240,000 km	135.6	128.5	125.2	133.1	126.1	122.9

	U.S. Midwest			U.S. South		
Configuration	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)						
80,000 km	160.4	152.0	148.1	160.6	152.1	148.2
160,000 km	138.0	127.9	124.6	133.9	126.8	123.6
240,000 km	126.5	119.9	116.8	125.0	118.4	115.4
5 Axle Semi Unit (Flat Deck)						
80,000 km	161.9	153.4	149.5	162.1	153.6	149.6
160,000 km	138.0	130.7	127.4	137.0	129.7	126.4
240,000 km	130.0	123.2	120.0	128.6	121.8	118.7
5 Axle Bulk Liquid Tanker						
80,000 km	156.2	148.0	144.2	160.0	151.6	147.7
160,000 km	126.0	119.4	116.3	128.3	121.5	118.4
240,000 km	116.0	109.9	107.0	117.7	111.5	108.7
5 Axle Bulk Dry Tanker						
80,000 km	161.7	153.2	149.3	165.7	157.0	153.0
160,000 km	126.9	122.1	118.9	131.2	124.3	121.2
240,000 km	117.9	111.7	108.8	119.8	113.5	110.5

	U.S. West		
Configuration	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
5 Axle Semi Unit (Van)			
80,000 km	183.2	173.8	169.1
160,000 km	156.8	148.5	144.7
240,000 km	147.9	140.2	136.6
5 Axle Semi Unit (Flat Deck)			
80,000 km	186.2	176.4	171.9
160,000 km	161.2	152.7	148.8
240,000 km	152.9	144.9	141.1
5 Axle Bulk Liquid Tanker			
80,000 km	172.5	163.4	159.2
160,000 km	141.1	133.7	130.3
240,000 km	130.7	123.8	120.6
5 Axle Bulk Dry Tanker			
80,000 km	178.1	168.7	164.4
160,000 km	144.0	136.5	133.0
240,000 km	132.7	125.7	122.5

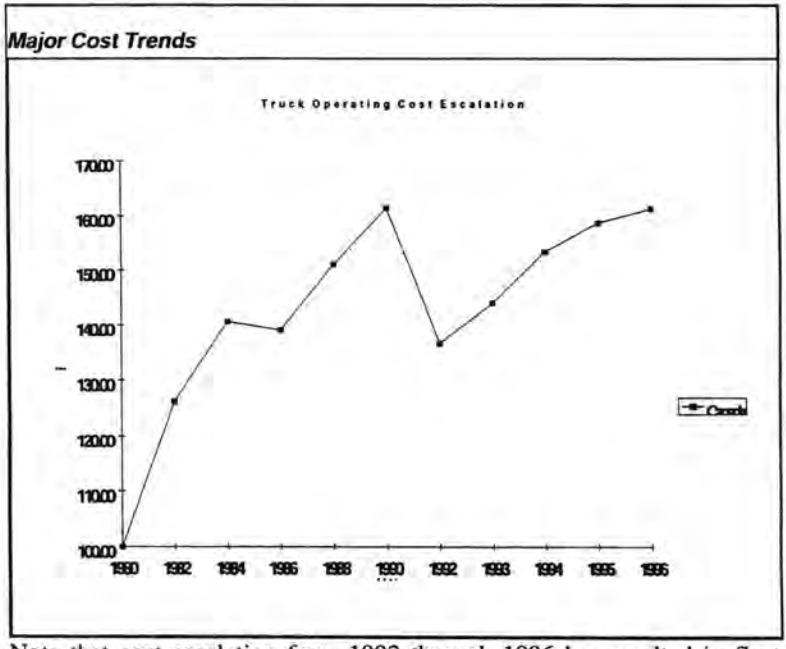
Note that the case studies for U.S. regions do not include the 8 Axle Super B-Train configuration that is in common use throughout Canada since the late 1980's RTAC studies on sizes and weights.

OPERATING COST RESULTS OVER TIME

As noted previously, the report goes back to 1972 for the Canadian regions, a period of 25 years. Some changes to configurations have occurred over this period. We have primarily seen a shift to include larger vehicle configurations, but also have witnessed an increase in the allowable dimensions and GVW (Gross Vehicle Weight) for a standard 5 axle semi-trailer configuration. Furthermore, some of the prior report

editions included various 2 axle straight truck configurations operating in a suburban setting. Evaluating these straight truck configurations has since been discontinued.

The following graph depicts how intra-regional (Canadian) truck operating costs have escalated since 1980 (the base year indexed as 100).

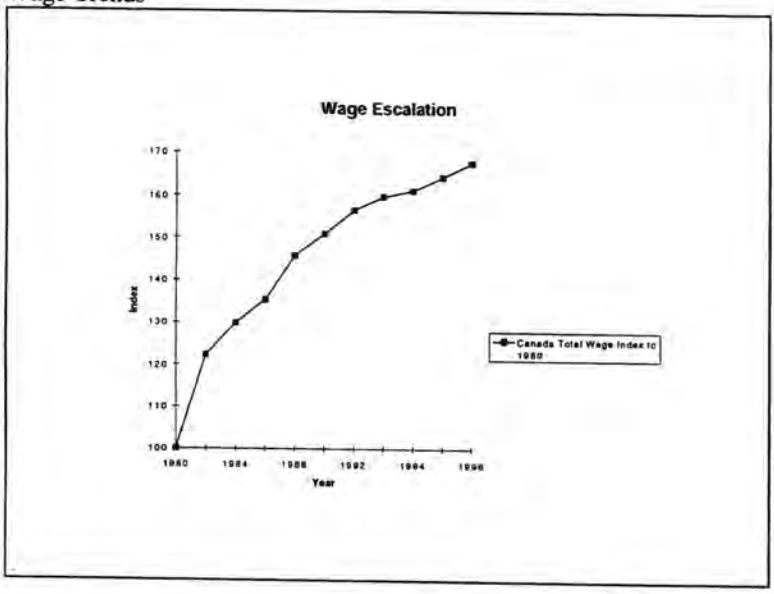


Note that cost escalation from 1992 through 1996 has resulted in fleet cost levels that are almost back to the 1990 (pre recession) cost levels. In addition to the 1990-92 recessionary correction in cost levels, enhanced fuel consumption and reduced repair costs have accompanied various technological enhancements such as the widespread use of electronic fuel injection diesel technology by the industry.

SAMPLE INPUT COST TRENDS

The following graphs illustrate trends in some of the important input cost components that underlie the over-all operating costs results presented.

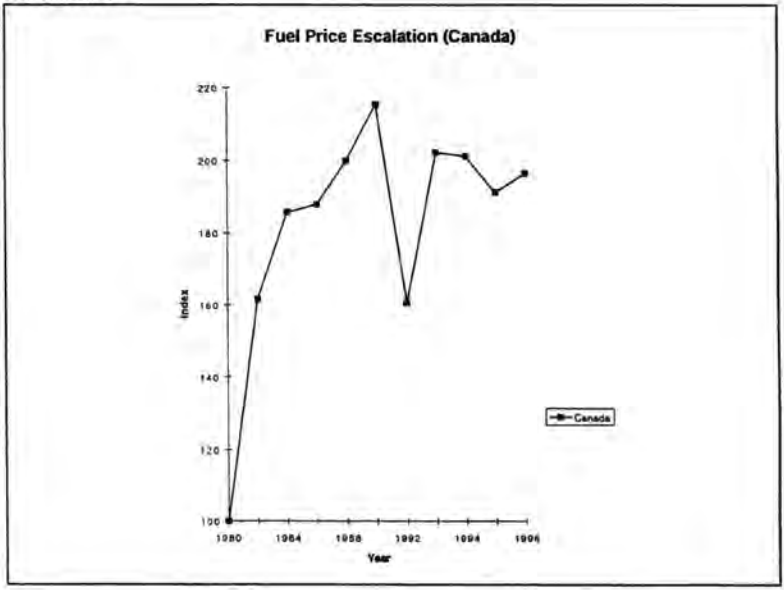
Wage Trends



Wage escalation in the industry tends to remain moderate, reflecting an average increase of 2% in Canada and approximately 2.5% in the U.S. between 1995 and 1996.

Looking at the longer trends, one notes that the graph was steeper in the early 1980's, reflecting high rates of wage inflation, than is the case in the later 1980's and into the 1990's.

Fuel Prices

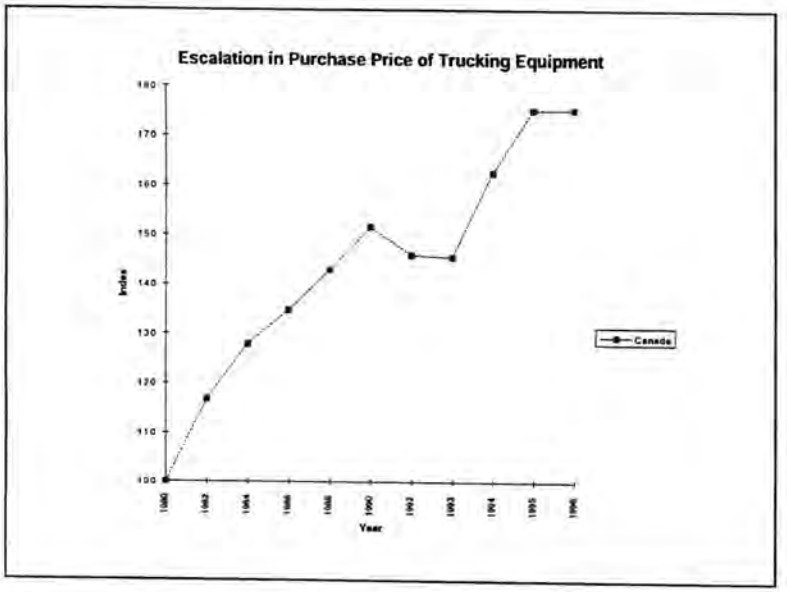


In the above graph of mid year fuel prices, costs are seen to be somewhat volatile. Although costs appear to have increased only slightly since 1995, users of the costing results in the 1996 study are cautioned that the (mid year) cost levels shown do not reflect escalation in fuel price (approximately a further 20%) that have occurred since July 1996.

In addition to the price factor (charted), over-all truck fuel costs have moderated since 1980 with the introduction of various fuel saving technologies, including:

- Aerodynamic Shapes of Vehicles and Add on Devices
- High Torque Diesel Engines
- Electronic Fuel Injection Technologies
- Increased Use of Radial Ply Tires

Equipment Purchase Cost Trends



Comparing 1996 cost levels to 1995, the principal finding noted was the "leveling off" of escalation in equipment prices. This factor reflects a slight easing in "orders" for new equipment as the trucking industry has steadily climbed out of recession, and replaced equipment that was being held on to during prior lean years. In late 1994 and 1995, average order lead times had been estimated by some suppliers as 12 to 18 months -- these have now eased to in the order of 3 to 6 months.

CANADA - US INTERNTIONAL TRUCKING COST COMPARISONS

In the 1996 study, we have separately considered all combinations of the following scenarios:

- Hauling distances of 160, 400, 800, 1600, and 3200 one way kilometers
- Three Corridors designated International West, International Central, and International East
- Operation on each of the corridors by a Canadian or a U.S. based trucker.

The Canada to U.S. dollar exchange rate resulted in substantial cost advantages (lower cost) for Canadian based operators on the Central and Eastern corridors.

On the Western Corridor, there was a slight favoring of U.S. based operators reflecting generally higher cost levels in B.C. for driver and mechanic wages.

The foregoing relative cost profile depends on exchange rates. Earlier Trimac studies (when the Canadian dollar value approached 90 cents, U.S.) have demonstrated the lack of competitiveness and the resultant business distress for Canadian trucking companies and owner-operators when there is a strong Canadian $. Specific case study summaries are tabulated on pages following.

CONCLUSION

The results from the study, "Operating Costs of Trucks in Canada – 1996", presented herein, are an overview. Interested readers are referred to the principal study document, distributed by Internet, for more detailed presentation of all cost components, by scenario. Also included are more detailed discussion of all costing assumptions and methods used, together with a comparison of linehaul trucking costs to intermodal COFC (Container on Flat Car) and TOFC (Trailer on Flat Car) truck/rail technologies.

Comments from users of the data are welcomed by both authors, who can be reached by e-mail at the following addresses:

- Lloyd L. Ash <lash@cadvision.com>
- Barry Gardner <gardneb@tc.gc.ca>

As previously stated, the full study is available from the Transport Canada website at http://www.tc.gc.ca.

Province : International West: Canada Based

Long Distance Hauling Case

One Way Distance (km)	Configuration and Annual Distance in (km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
	5 Axle Semi Unit (Van)			
160	240,000	150.7	142.7	139.1
400	278,581	141.3	133.8	130.4
800	294,677	136.9	129.7	126.3
1600	303,958	134.6	127.6	124.3
3200	308,290	133.5	126.5	123.3
	5 Axle Semi Unit (Flat Deck)			
160	240,000	156.4	148.2	144.4
400	292,491	142.0	134.6	131.1
800	315,829	136.0	128.8	125.5
1600	329,506	139.5	132.2	128.8
3200	336,221	137.9	130.6	127.3
	5 Axle Bulk Liquid Tanker			
160	240,000	139.8	132.5	129.1
400	256,991	138.1	130.8	127.5
800	263,501	136.3	129.1	125.8
1600	267,343	135.3	128.2	124.9
3200	268,838	134.9	127.8	124.5
	5 Axle Bulk Dry Tanker			
160	240,000	141.7	134.2	130.8
400	257,537	139.7	132.3	128.9
800	264,266	137.7	130.5	127.1
1600	268,227	136.7	129.5	126.2
3200	269,782	136.2	129.1	125.8

Province : International West: USA Based

Long Distance Hauling Case

One Way Distance (km)	Configuration and Annual Distance in (km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
	5 Axle Semi Unit (Van)			
160	240,000	148.4	140.6	137.0
400	281,202	139.6	132.3	128.9
800	298,597	134.7	127.6	124.3
1600	308,650	132.2	125.2	122.0
3200	313,394	131.0	124.1	120.9
	5 Axle Semi Unit (Flat Deck)			
160	240,000	153.1	145.0	141.3
400	292,491	140.1	132.7	129.3
800	315,829	133.7	126.7	123.5
1600	329,506	136.6	129.4	126.1
3200	336,221	134.9	127.8	124.6
	5 Axle Bulk Liquid Tanker			
160	240,000	131.0	124.1	121.0
400	256,991	131.9	124.9	121.7
800	263,501	130.1	123.3	120.1
1600	267,343	129.2	122.4	119.3
3200	268,838	128.8	122.0	118.9
	5 Axle Bulk Dry Tanker			
160	240,000	133.0	126.0	122.8
400	257,537	133.6	126.6	123.3
800	264,266	131.8	124.8	121.6
1600	268,227	130.8	123.9	120.7
3200	269,782	130.3	123.5	120.3

Province : International Central: Canada Based Long Distance Hauling Case				
One Way Distance (km)	Configuration and Annual Distance in (km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
	5 Axle Semi Unit (Van)			
160	240,000	122.7	116.2	113.3
400	268,857	117.6	111.4	108.6
800	280,395	114.6	108.6	105.8
1600	287,034	113.1	107.2	104.4
3200	289,971	112.4	106.5	103.7
	5 Axle Semi Unit (Flat Deck)			
160	240,000	132.0	125.1	121.9
400	292,491	119.3	113.0	110.1
800	315,829	113.8	107.8	105.0
1600	329,506	116.7	110.5	107.7
3200	336,221	115.2	109.1	106.3
	5 Axle Bulk Liquid Tanker			
160	240,000	125.5	118.9	115.9
400	256,991	123.8	117.3	114.3
800	263,501	121.9	115.5	112.6
1600	267,343	121.0	114.6	111.7
3200	268,838	120.5	114.2	111.3
	5 Axle Bulk Dry Tanker			
160	240,000	127.4	120.7	117.6
400	257,537	125.4	118.8	115.7
800	264,266	123.4	116.9	113.9
1600	268,227	122.4	116.0	113.0
3200	269,782	121.9	115.5	112.6

Province : International Central: USA Based Long Distance Hauling Case				
One Way Distance (km)	Configuration and Annual Distance in (km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
	5 Axle Semi Unit (Van)			
160	240,000	152.8	144.8	141.1
400	281,202	142.3	134.8	131.3
800	298,597	136.7	129.5	126.2
1600	308,650	133.9	126.9	123.6
3200	313,394	132.5	125.6	122.3
	5 Axle Semi Unit (Flat Deck)			
160	240,000	158.5	150.2	146.3
400	292,491	143.1	135.6	132.1
800	315,829	136.0	128.8	125.5
1600	329,506	138.6	131.3	127.9
3200	336,221	136.7	129.5	126.2
	5 Axle Bulk Liquid Tanker			
160	240,000	132.4	125.5	122.2
400	256,991	132.7	125.8	122.5
800	263,501	130.8	123.9	120.7
1600	267,343	129.8	123.0	119.8
3200	268,838	129.3	122.5	119.4
	5 Axle Bulk Dry Tanker			
160	240,000	134.5	127.4	124.1
400	257,537	134.5	127.4	124.1
800	264,266	132.4	125.5	122.2
1600	268,227	131.4	124.5	121.3
3200	269,782	130.9	124.0	120.8

Province : International East: Canada Based

Long Distance Hauling Case

One Way Distance (km)	Configuration and Annual Distance in (km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
	5 Axle Semi Unit (Van)			
160	240,000	111.3	105.4	102.7
400	275,703	106.7	101.1	98.5
800	290,408	103.9	98.5	96.0
1600	298,872	102.5	97.1	94.6
3200	302,769	101.9	96.5	94.0
	5 Axle Semi Unit (Flat Deck)			
160	240,000	115.0	109.0	106.2
400	292,491	106.5	100.9	98.4
800	315,829	102.5	97.1	94.6
1600	329,506	105.9	100.3	97.7
3200	336,221	104.8	99.2	96.7
	5 Axle Bulk Liquid Tanker			
160	240,000	110.7	104.9	102.2
400	256,991	110.3	104.5	101.8
800	263,501	108.8	103.1	100.5
1600	267,343	108.1	102.4	99.8
3200	268,838	107.8	102.1	99.5
	5 Axle Bulk Dry Tanker			
160	240,000	112.6	106.6	103.9
400	257,537	111.8	105.9	103.2
800	264,266	110.3	104.5	101.9
1600	268,227	109.6	103.8	101.1
3200	269,782	109.2	103.5	100.8

Province : International East: USA Based

Long Distance Hauling Case

One Way Distance (km)	Configuration and Annual Distance in (km)	90 Operating Ratio Total Costs (c/km)	95 Operating Ratio Total Costs (c/km)	97.5 Operating Ratio Total Costs (c/km)
	5 Axle Semi Unit (Van)			
160	240,000	153.6	145.6	141.8
400	281,202	143.3	135.7	132.2
800	298,597	137.8	130.6	127.2
1600	308,650	135.1	128.0	124.7
3200	313,394	133.8	126.7	123.5
	5 Axle Semi Unit (Flat Deck)			
160	240,000	159.1	150.7	146.8
400	292,491	144.0	136.4	132.9
800	315,829	137.0	129.8	126.5
1600	329,506	139.8	132.4	129.0
3200	336,221	137.9	130.6	127.3
	5 Axle Bulk Liquid Tanker			
160	240,000	134.0	127.0	123.7
400	256,991	134.2	127.2	123.9
800	263,501	132.3	125.4	122.2
1600	267,343	131.4	124.4	121.2
3200	268,838	130.9	124.0	120.8
	5 Axle Bulk Dry Tanker			
160	240,000	136.0	128.8	125.5
400	257,537	135.9	128.8	125.5
800	264,266	133.9	126.9	123.6
1600	268,227	132.9	125.9	122.7
3200	269,782	132.4	125.4	122.2

Focusing on Growth through High Quality Service and Innovative Products at Low Cost

Gerry Kolaitis

Department Director
Corporate Planning and Financial Support
VIA Rail Canada Inc.

May 27, 1997

Facing the funding challenge while keeping the network intact

In 1993, in view of over $100 million annual reductions in government funding for passenger rail contained in the Federal budget, VIA faced the difficult choice of either eliminating a quarter of its train services network or taking steps to lower costs dramatically while maintaining and improving service to customers and meeting high quality standards. VIA opted for the latter and this strategy has proven to be the right one. Ridership and revenues are up while operating costs and government funding are down dramatically.

- For most of VIA's history, the company earned only 30 cents of revenue for every dollar of expenses, while requiring very large amounts of funding from the Government of Canada. For example, in 1988 VIA required $637 million of subsidy (over $800 million in today's dollars), which in 1990, prompted a 50% cut to its train services network and resulted in reducing funding to $389 million by 1992.

- By 1996 things had changed. VIA earned 47 cents of revenue for every dollar of expenses, and that will increase to over 60 cents in the next three years. This is reflected in the reduction in government

funding to $245 million in 1996, and that will be down to $170 million by 1999. This is in large part due to a cultural change within the corporation concerning finances. In effect, VIA now acts like any other company trying to improve its bottom line and has adopted an aggressive, private enterprise approach towards customer service, revenue generation and cost reduction.

- Over the next few years VIA will focus on the achievement of common goals and objectives as outlined in its 1996-99 Strategic Plan entitled *Service for Success*. This plan confirms VIA's vision "to be the best passenger transportation company in Canada" and its mission "to deliver high quality, low cost passenger service". The strategic plan recognizes the importance of each employee's contribution to the success of the corporation, and makes special efforts to this effect through improved information technology, effective communications and an emphasis on training and celebration of team successes.

Emphasis on the bottom line

Between 1992 and 1996, VIA's bottom line improved by $126 million annually --- operating costs were reduced by $97 million while revenue increased by $29 million. This represents a major turn-around by any standard.

- Further improvements will see operating costs down by a remarkable 24% between 1992 and 1999 (after absorbing another 12% of inflationary pressures on costs) while revenues are projected to increase by 47%. These goals will be achieved through the continued efforts of the corporation to build on the solid financial framework that has been developed.

- In 1996 alone, VIA's funding requirements from the federal government dropped by $50 million; the largest year over year

reduction since 1990. Total government funding in 1996 was $245

million, representing a decrease of almost 37% from the 1992 funding.

- By 1999, revenues are projected to be up by another $45 million (an increase of 24%) while operating costs will be down by another $19 million (a decrease of 5%) versus 1996 levels (despite the large increase in revenue and ridership).

- VIA's government funding will decrease to $170 million by 1999, down 60% versus 1992 and 80% versus 1988, in real terms.

The increase in revenue is the result of superior customer services and on-time performance, innovative marketing and promotion initiatives as well as increased train capacity. Costs have been reduced as a result of organizational streamlining, modernization of collective agreements, lower maintenance costs and numerous other productivity measures.

- The streamlining of the organization, which began in 1993, has resulted in the reduction of administrative and management staff by almost 800 or over 60%, saving $46 million annually.

- New collective agreements have allowed VIA to modernize its labour practices. Shopcraft unions have been consolidated, reducing the number of job classifications and increasing flexibility. The method of payment for running trade employees has changed from a mileage to an hourly basis and assistant conductor positions were eliminated on most trains. Customer service staff in the Corridor has been rationalized. The number of employees eligible to receive employment security has been reduced.

- In 1997, a further step in modernizing VIA's crewing operations will take place with the merger of the conductor and engineer positions thus eliminating 227 train conductor/assistant conductor positions, while improving safety and saving $15 million annually. The merger, a first in North America, means that the two engineers in the

cab will be responsible for the safe operation of the train and for coordination of emergency response in case of accidents. Passenger-related duties currently performed by the conductor will be handled by customer service employees. Service managers on the Eastern and Western transcontinental trains and a designated in-charge employee on Quebec-Windsor Corridor trains will be responsible for all customer service matters. Furthermore, customer service personnel will be given enhanced safety training to ensure the security of passengers.

- The number of employees stood around 4,500 in 1992. This declined to 3,000 by the end of 1996 and will decline to 2,700 by the end of 1997, all with the same train services network.

- In 1996, VIA retired the last of its steam-heated equipment with the introduction of refurbished and modernized stainless steel cars on its services in northern Manitoba, northern Quebec and southwestern Ontario. (The refurbishment of 33 stainless steel cars was internally-financed through productivity savings.) The refurbished equipment ensures greater reliability and improved comfort for passengers and represents significant savings for VIA. In southwestern Ontario the introduction of the refurbished cars increased VIA's first class capacity by over 40%. The cost of refurbishing the cars was about two thirds of the cost of buying new equipment, resulting in savings of $30 million. Operating savings are expected to be $2 million annually.

- The elimination of steam-heated equipment along with better equipment cycling and improved availability and reliability of the fleet have allowed VIA to reduce the number of active cars and locomotives in its fleet to 400 --- 323 cars and 77 locomotives. This compared to 577 pieces of equipment in 1991, a reduction of 31% with the same train services network.

- In June 1996, VIA announced the closing of the Halifax and Toronto maintenance centres which will be completed this spring. The consolidation of maintenance activities in Montreal, Winnipeg and Vancouver will result in annual savings of up to $7 million. VIA has also introduced various cost reduction initiatives in equipment maintenance including the reduction of unscheduled repairs, improving the quality of work and streamlining maintenance and servicing practices.

- In September 1996, VIA introduced a new economy class food service on all Corridor trains. This service eliminated complimentary snacks while offering a better selection of sandwiches and other items for sale. This change represents an overall annual benefit of $3.5 million --- $1.7 million in savings and a $1.8 million increase in revenue.

The customer service results speak for themselves

Today, with less than 3,000 employees, VIA operates 429 weekly trains serving more than 400 communities across the country. In 1996, VIA carried 3.7 million passengers and earned $185 million in revenues. While this represents only a 2% increase in the number of riders since 1992, passenger miles are up 9% and total revenue is up by 18% despite extremely competitive markets and difficult economic conditions.

- Passenger satisfaction ratings in 1995 and 1996 averaged 95% (customers who said VIA service met or exceeded their expectations). Of note are the over 30% of passengers on the western and eastern transcontinental trains who reported that the overall train trip exceeded their expectations. And some 35 to 40% of these same passengers reported that VIA exceeded their expectations in terms of value for money.

- Operating performance and service quality have continued to be high. VIA's on-time train performance has averaged 84 to 90% since 1991. Equipment performance has been excellent with 88% equipment availability and equipment reliability averaging 229,000 miles per vehicle, or almost two years, between major defects.

- At the same time safety improved in 1996 versus the previous year. The number of lost-time accidents was reduced by 31% and the number of lost days due to those accidents by 32%. By this spring, all on-train employees will have received the Emergency Response Training Program.

Setting --- and achieving --- goals for success through people

VIA's success has been made possible by the commitment and contributions of all VIA employees, reinforced through effective two-way communications, improved information technology and enhanced training, and celebration of team and individual successes and achievements. It is also a result of total commitment to quality and an emphasis on customer services and cost control. For example, front line staff are given a "License to Serve" which gives precedence to superior customer service at the expense, in certain circumstances, of policies and procedures.

VIA's efforts on these fronts have not gone unnoticed. In 1996, VIA was selected by the Canada Quality Council to chair its public awareness campaign and Canada Quality Month. VIA's campaign theme was "Quality comes from within". VIA's new uniforms and the refurbished stainless steel equipment won international acclaim for outstanding visual design in railway transportation from the Watford Conference of European Railway Architects and Designers.

New strategies to increase ridership and revenues include innovative marketing and promotional programs, excellent on-time performance and equipment reliability, better utilization of equipment to take advantage of peak demand and superior customer service at all levels from economy class to "Silver and Blue" sleeping car service on the western transcontinental train *The Canadian*.

Launching new products and services

- In June 1996, VIA introduced a frequent-user program called VIA Preference. By year-end, the program had attracted some 50,000 members. In addition to increasing ridership and revenue, the program will build a customer profile base that will allow VIA to develop more cost-effective and focused promotional activities and market research efforts.

- VIA is the first railway in the world to offer an Internet booking and payment system called VIA Resernet. During a two-month test period, over 30,000 users visited the site which resulted in some 1,000 confirmed reservations.

- In 1996, VIA began accepting reservations for coach travel, formerly limited to VIA 1 and sleeping car passengers. This year, a new reservations system will be introduced which will reinforce the linkages already created with the Sabre, Apollo and Worldspan airline reservation systems (representing 35,000 U.S. and 5,000 Canadian travel agents and tens of thousands worldwide) and allow expansion of the international sales network through travel agents. The new system will provide features such as inventory and yield management, with full reservation and seat selection for all classes of service.

- Last year, VIA introduced a two-day, daylight, touring-oriented service on the *Skeena* (Jasper, Alberta-Prince Rupert B.C.) and ridership rose by close to 40% during the key summer season. New and more attractive daylight schedules were introduced for the

northern Quebec services. Along with the Winnipeg-Churchill trains, these services are being promoted to the growing "adventure" and "ecotourism" markets.

- Capacity was increased on both the Eastern and Western transcontinental trains to take advantage of high summer demand. In 1996, bookings by western wholesalers increased peak passenger mile volumes by 18% on *The Canadian* while price promotions in the Maritimes increased year-round passenger volumes by 14%. This summer, the consist of the western transcontinental train between Jasper and Vancouver will be up to 25 cars (almost 1/2 mile long, carrying 500 passengers) as additional sleepers are being assigned to this popular segment.

- Frequencies were added in the Corridor to test the long-distance commuter market in key southern Ontario markets. At the same time, VIA tested the self-propelled European IC3 Flexliner trains which are designed to operate at lower cost. It was found that although the equipment is less expensive to operate, it is not yet economic enough to generate a profit given the low average fares that these commuter markets could sustain, but we may continue the test on the Kingston-Toronto service because of its potential. Nevertheless, VIA will continue to explore options to serve this market segment.

Developing partnerships for growth

- Pricing promotions included a promotion with the International Students' Association which enabled students to more easily qualify for off-peak fares, family fares offering free travel for children to stimulate summer traffic in the Corridor and, as a joint effort with the Canadian Tourism Commission, "Rediscover Canada" fares to encourage round-trip, inter-provincial travel.

- Other promotional efforts included joint sponsorship of ABC TV's Good Morning America's tour of Canada along with the Canadian Tourism Commission, Canadian Airlines International and CP Hotels and Resorts. The show, with an audience of some 30 million, featured segments on rail travel over a five-day period. Seasonal campaigns were launched with tourism authorities across Canada as well as numerous joint travel promotions for specific events and attractions. These kinds of successful promotions will continue this year. VIA is also aiming at the business community in the Corridor this year with a program that will target Chambers of Commerce with executive presentations to members, and sponsorship of events.

- VIA also works closely with industry partners, government and police agencies to improve safety. This includes the Operation Lifesaver campaign to educate the public about the danger of ignoring crossing signals and trespassing. Along with CN, local police forces and representatives from Operation Lifesaver, VIA has participated in an Officer on the Train program, an approach to reducing level crossing incidents. VIA was also a partner in the Canalert research project, a comprehensive study of the factors affecting locomotive engineers' performance. The results will help to improve the health and welfare of employees, enhance railway safety and the security of passengers.

Investing for a better future

- Currently, VIA carries 3 million passengers in the Quebec-Windsor Corridor and this is expected to increase to 3.5 million by 2001. VIA will be introducing changes to both its economy class service throughout the system and VIA 1 (first class) services in the Corridor in response to the need to significantly increase revenues during the coming years. All aspects of passengers' experience from the first phone call to the final farewell are under review. Among the improvements planned for VIA 1 service are visually distinctive interiors providing more seating capacity as well as electrical outlets at

seats, specialized meeting/conference areas and upgraded meal service. Furthermore a state of the art system to track customer complaints, commendations and feedback is under development.

- The public address system on the LRC trains was upgraded last year. As well, VIA introduced a series of improvements to emergency-handling procedures and equipment including trauma kits, new safety and emergency signs and pamphlets, and redesigned emergency lighting systems. This year, LRC cars will have a better margin of safety with a new much more economical, air-only braking system.

- Investment in training continues to be a top priority at VIA. Training programs are focused on safety, customer service and supervision skills, including announcement skills and selling skills. Over 600 participants have taken part in the "Achieving Excellence in Customer Service", a two-day workshop focused on improving direct customer contact and providing people with the tools to create positive and memorable customer experiences. Some 75 employees have completed "Supervising for Service Success", an 11-day program designed to create "missionaries for service". In 1996, individuals and work teams developed and implemented more than 400 Performance Enhancement Projects, designed to make the best use of employees in an environment of innovation and participation as well as recognition of employee efforts.

- Besides the new state-of-the-art reservation system, VIA has invested in other new technologies including an Executive Information System that gives management daily or monthly access to detailed financial, revenue, ridership and operating performance data through their desktop or portable computers. A self-monitoring diagnostics system on LRC cars and F40 locomotives helps the shop floor in the trouble-shooting of failures and assessment of performance.

- A comprehensive review of VIA's network of stations is underway to streamline operating costs and improve station facilities and their related services including their esthetics. The goal is to ensure that VIA operates, upgrades and effectively manages quality passenger service

facilities, satisfying customer and operational requirements at the lowest cost.

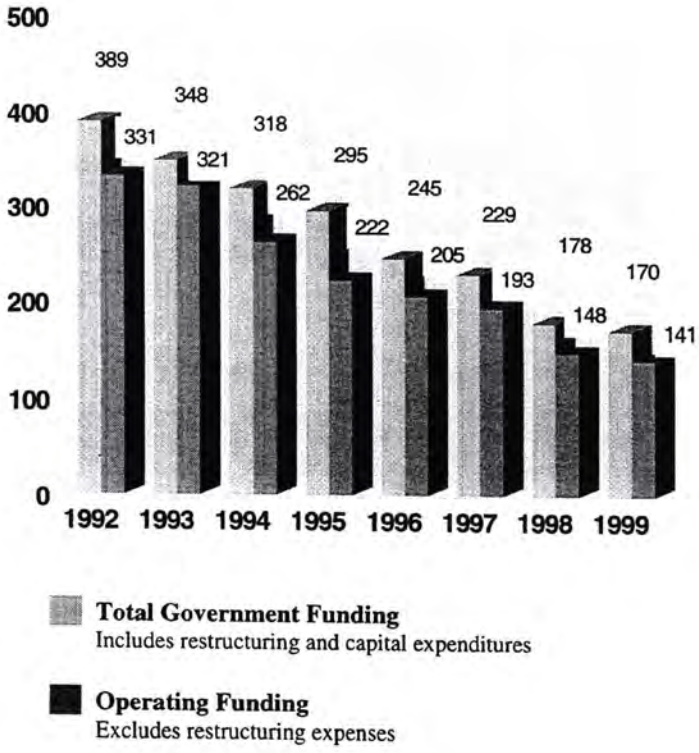
Confidence in the future

A major priority that VIA faces is developing and funding capital investments which will allow the corporation to meet the more demanding expectations of the travelling public, in a highly competitive market. To meet customer needs for faster trip times, more convenient departures and enhanced on-board amenities - planning for infrastructure improvements and replacement of the fleet must now begin. Our ability to satisfy the demand for more capacity, while continuously lowering costs and generating a positive contribution to the bottom line, depends on targeted capital investments.

VIA still faces challenges in the years ahead if it is to meet its funding targets by the end of this century, but it has turned a very important corner in the past few years and there is no going back. Its revenue and ridership targets are ambitious --- as are its projections for further operating cost savings. This will require a commitment by all employees to continuously improve products and services and find new innovative ways to eliminate waste in the drive towards self-sufficiency. However, there is no doubt that a very important change to the culture of the organization has taken place that will allow these things to happen. And if VIA's strong performance of the past few years is a sign of things to come, there is little doubt that we will succeed.

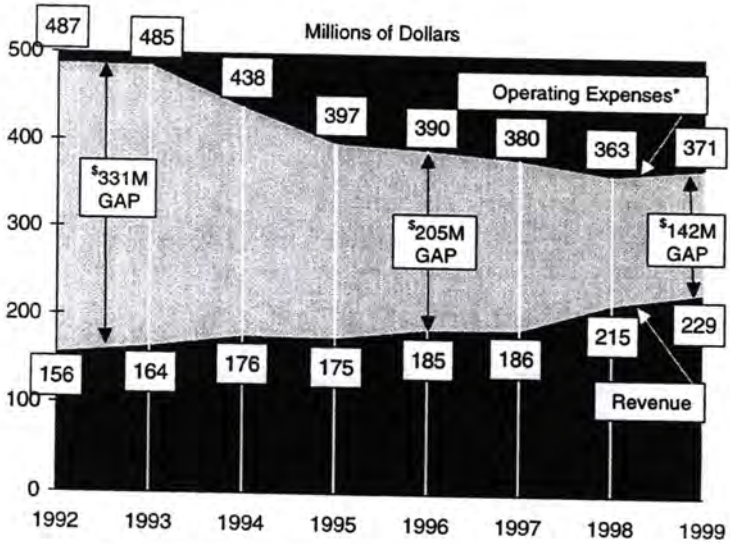
VIA RAIL CANADA INC.

TOTAL GOVERNMENT AND OPERATING FUNDING



VIA RAIL CANADA INC.

REVENUE & OPERATING EXPENSES

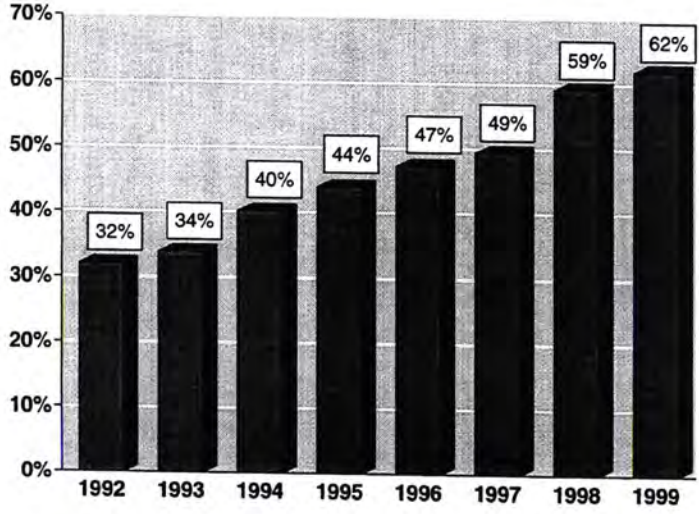


* Excludes restructuring & non-cash expenses

Operating Deficit (Gap)

VIA RAIL CANADA INC.

REVENUE / EXPENSE RATIO



Excludes restructuring and non-cash expenses

Logistics Defined

Logistics has been defined as procuring, maintaining, and transporting materiel, personnel and facilities. Its roots are in military science - an army marches on its stomach. Buxton traces the evolutions of management focus since the beginning of the industrial revolution from production to finance to marketing. Only since the 1960's has logistics begun to emerge as the fourth evolution in managerial practice[1]. The Mission of the Logistician is: "to get the right goods or services to the right place, at the right time, and in the desired condition, while making the greatest contribution to the firm"[2]. Few would dispute that the success of any firm in a competitive environment of free and global trade depends on customer satisfaction, the customer's view of the price, quality and service offerings of the firm. All three elements may only be achieved through logistics activities.

In a world where change is the only constant, all product offerings are subject to the product life cycle. Much of the post WW II focus of management has been on the challenges of the introduction and growth phases. The 1990's have reminded us of the limits to growth as

---

[1]Buxton, Graham, *Effective Marketing Logistics, the MacMillan Press, London, 1975*

[2]Ballou, Ronald J., *Business Logistics Management*, Third Edition, Prentice-Hall Inc., 1992

more and more products pass through the maturity stage and enter the decline phase. Down sizing has been the response of many companies as they struggle to stay alive. But there is another way. "Re-engineering" is the innovation of finding new ways to do business, of re-inventing the product, giving it new forms (packaging), new uses (needs), of adding value for new levels of customer satisfaction, giving it a boost on the product life cycle curve, a new growth phase. Many industries have done this through the better understanding and wider application of logistics in the management of the delivery of the firm's raison d'etre. In the 1980's, innovation and innovative response was evident in an increasingly dynamic and unforgiving environment, what came to be known as "new business practices". The RAND Institute termed this "lean logistics", a network that links suppliers, producers and carriers in closely integrated transportation and distribution systems for enhanced customer service while at the same time achieving reduced costs[3]. Do logistics principles also apply to the movement of people? to the provision of travel services?

The Principles of Logistics Management

Logistics is a broad subject. Ballou identifies ten activities involved in the logistics process. Customer

---

[3]Levine, Arnold and Luck, Jeff, *The New Management Paradigm, A review of Principles and Practices,* The RAND Institute, 1994

service, transportation, inventory management and order processing are the major activities, supported by: warehousing, materials handling, purchasing, packaging, coordination with production and information management. However, a simple comparison of how these activities are carried out for the movement of people would be a dry paper. Instead, this paper takes the perspective of management, of the processes and tools offered in this fourth evolution of management focus. Every logistics decision is a *trade-off* between competing requirements. It is a "satisficing" of a mix of needs as opposed to maximizing a single goal. Indeed, this is the essence of management. It is simplistic to say the purpose of management is to maximize profits. As soon as the qualifier "in the long term" is added, then management also needs to consider investment in R&D, development of human resources and being a good corporate citizen, all considerations that may cost money and detract from profit in the short term. The most fundamental trade-off for the logistics manager is *cost versus service* (the same trade-off from the customer's view is the price versus value optimum). The Total Distribution Cost (TDC) concept is the trade off between direct transport rates and the indirect costs associated with service, usually a function of time, which include inventory, insurance, packaging, etc.

In optimizing trade-off choices, logistics applies *engineering* approaches to quantify and analyse. For example, the 80:20 rule suggests 20% of your customers produce 80% of your profits. Certainly, this justifies greater efforts for those most important customers, but not at the neglect of the other 80%, who likely include

tomorrow's premium customers. Equally important are *marketing approaches*, which may complement the quantified approach or require its compromise. Logistics adds to the marketer's opportunities for further product differentiation to satisfy finer nuances of market segmentation. The same product or service becomes three different things when there is the choice of receiving it today, tomorrow or next week. Logistics also relies on the application of *information management* systems for the audit and control function of management. Wal-Mart revolutionized retail through advanced logistics based on information technology. Engineering, marketing and information management are the key principles of modern logistics practice.

Today's Travel Industry

Travel and tourism is a big business. Foreign visitors and Canadian travellers spend $21 billion; direct employment is 600,000; tourism is Canada's third largest "export" industry in terms of foreign exchange[4]. People move for a variety of reasons. There is business travel and personal travel; personal travel may be for vacation or visiting friends and relatives. But all travel is undertaken to achieve some experience, be it closing a business deal or Caribbean romance. Private road vehicles still provide almost 90% of inter city passenger kilometres, but the air mode has steadily been increasing its share (while bus and

[4]*Starting a Travel Agency,* Entrepreneur's Handbook, Manitoba Industry, Trade and Tourism, 1991, p29.

rail have stayed fairly constant)[5]. In fact, travel by bus is now being supplanted by discount air travel that is price competitive with bus for the long haul. That is why Greyhound has entered the air business to carry passengers long haul and then distribute them over their local bus systems; it is the first inter-modal passenger carrier. Rail and sea, the mainstays up to forty years ago, are now used more as attractions; scenic trains through the Rockies or cruise ships as floating resorts.

The travel industry is first made up of service providers. This sector can also be considered in the transport carriers such as the airlines, and in the destination or en route/stopover services such as hotels, restaurants and attractions. The distribution of travel services includes wholesalers and retailers. The wholesalers are a recent phenomenon of the past thirty years. They act much like a freight forwarder by buying travel services in bulk (blocks of air seats, hotel rooms, etc.) and repackaging these as individual vacations. Retailers are the travel agents found in most malls and office complexes.

While airlines, hotels, and car rental companies all market directly through 800 numbers and their own premises, about 80% of travel arrangements are processed by travel agents on a commission basis, paid by the service provider. Prior to the development of the

---

[5]Hyndman, Louis D., Chair, *Directions, the Final Report of the Royal Commission on National Passenger Transportation*, Supply and Services Canada, Ottawa, 1992, p.3

wholesaler, travel agents provided wider services in making arrangements with a variety of suppliers to put together an individual's trip. The advent of the tour company wholesalers greatly reduced the need for travel agents to provide such detailed services to the client as there is likely to be a wholesaler somewhere with a prepackaged tour to satisfy the customer's needs. The retail travel agent, whether in the mall or office complex, has become less a problem solver and more an order taker. The counselling function may be limited to information about what is being ordered or advising on the merits of alternative packages. The retailer's cost pressures force them to concentrate on volume. In the long term, their role is jeopardized because they add less value themselves in the distribution chain. Retail travel is already a tough business. There are some 140 travel agencies in Manitoba competing for about $60 million in business; return on sales is typically only 0.75%[6].

It has long been held that many consumers do not use a travel agent because they believe it will cost more; a recent U. S. Travel data Centre survey found only 5% believe this. At the same time as technology threatens to replace them, Judson Green, President, Walt Disney Attractions, has advised travel agents that consumers "don't trust you" (neither do they trust lawyers, bankers, brokers or other agents). He advises agents to: "Entice,

---

[6]*Starting a Travel Agency,* Entrepreneur's Handbook, Manitoba Industry, Trade and Tourism, 1991,p 1.

Engage, Educate, Entertain."[7] Another trend is industry consolidation where the industry is dominated by a few large chains or consortia. Related to this is the trend to vertical integration where a major wholesaler also operates retail outlets that appear as independents while obviously favouring the products of their parents[8]. The Transat AT Group has grown in a few short years to become a fully integrated retailer, wholesaler, and airline covering the Canadian market and is now focussing on foreign acquisitions and hotel management[9].

Today's travel agency industry is largely a product of deregulation: "The ever-changing array of fares, schedules, service terminations, route entries, and entirely new airlines that came in the wake of deregulation threw increasing numbers of confused consumers into the arms of travel agents, the official interpreters of the mysteries of air travel."[10] From 1977 to 1985, the number of

---

[7]Green, Judson, *The Travel Agency Business - The Family Market,* Vital Speeches of the Day, Volume 62, Issue 8, pp247-250, February 1, 1996

[8]*OFT Eyes Travel Agent Mergers,* Marketing, p7, February 8, 1996

[9]Jean Marc Eustache, President, interviewed on *Venture*, CBC, March 2, 1997

[10]Meyer, John R. and Oster, Clinton V. Jr., *Deregulation and the Future of Intercity Passenger Travel,* Massachusetts Institute of Technology, 1987 (Chapter 8. *Changes in Distribution Channels and*

accredited agencies in the U.S. more than doubled from 13,000 to 27,000 (airlines were deregulated in 1978). There are 36,000 IATA accredited agencies serving the rest of the world. Technology has been the travel agent's tool for growth with productivity gains over 40%. In fact, it was technology that permitted decentralization of airline ticket distribution to travel agents. The changes in the 1970's and 1980's enhanced customer service while lowering distribution costs. But it could be an extension of these technological changes that permit a further decentralization from agents directly to the travelling public[11].

Is the retail travel agent to become an anachronism? Or are there opportunities, in applying the principles of logistics management, to add new value in the overall travel service product?

Logistics In Passenger Transportation

Engineering Approaches in Travel Distribution

Most examples of cost optimizing can be observed at the level of the service provider, particularly airlines. Aircraft size and frequency selection on a particular city pair is the primary tool in matching fluctuating demands to productive capacity. The use of hub and spoke networks would be another example of cost-service trade-offs. In most cases, a trip between a city pair will be less

---

the Travel Agency Business)

[11]Ibid.

direct and will involve the extra time of a connection. However, the trip can be taken more often (all the traffic from a point is routed over the hub permits greater frequency) and at lower cost (consolidation of demand over the hub permits a larger aircraft with lower unit cost where the saving can be passed along to the customer in lower fares).

Airlines are applying engineering approaches in their focus on distribution costs, which now exceed fuel costs, and can run 20% of revenue[12]. Greyhound has recently introduced an airline travel service by-passes the traditional retail travel agent. Their call centre distribution channel is backed up by their own network of 600 bus stops. However, Robert Christie, President of ACTA Manitoba, reported that Greyhound is negotiating

---

[12]Airlines claims of high distribution costs can be inflated. Where pricing remains regulated, such as on the Pacific, airlines dump blocks of seats for marketing by consolidators (also known as "bucket shops"). Agents will then buy a ticket for a customer at a net cost and add their own mark-up. For example, a ticket with a fare of $2000 may be available from a consolidator at $1000. An agent adds $100 "commission" and sells it for $1100. The actual sale may be reported through BSP as a $2000 sale with a $1000 commission paid while the customer's invoice shows $2000 as the price with a $900 "discount" applied. The airline reports a distribution "cost" of $1000 when in fact the agent only made $100.

to access the travel agent distribution channel; it may be that their attempt to totally bypass the industry's retail level has been premature. The increasing commission expense of U. S. carriers since 1978 reversed course in 1994 and further decreased in 1995, from 11% in 1993 to 10.4% in 1994, it is now down to 9.3% of passenger revenues[13]. Travel agent commissions are not the only element of airline distribution costs which now total 18.6% of operating expenses. Morris Air is pioneering home based reservations agents as more adaptable and lower cost than a call centre[14]. Bar coding is another technological innovation that is evidence of engineering approaches in airline distribution. Electronic ticketing, real time wireless baggage scanners and a variety of self service devices permit a traveller to go from home computer to an airline seat without any intervening stops in the terminal[15]. Yield management is an airline black science in the effort to match demand to capacity with

---

[13]Henderson, Danna K., *Commissions in decline,* Air Transport World, pp 108-112, October 1995

[14]Flint, Perry, *Homework - Is the traditional airline res center on the endangered species list?,* Air Transport World, pp138-140, June 1996

[15]Henderson, Danna K., *On the bar code road,* Air Transport World, pp 88-90, September 1996

maximum revenue.

A wholesaler might be applying an engineering choice in a strategic decision to concentrate on fewer destinations where consolidation allows it to purchase more product (e.g. hotel rooms) to achieve lower costs which it can pass on as lower prices to capture greater market share for that destination. However, this can be a trade off if fewer destinations reduce its overall product line and positioning in consumer awareness. Recent consolidations have seen the larger wholesalers acquiring the smaller ones which could not compete in these economies of scope (e.g. Regent by Transat, Alba by Sunquest). At the retail level, the best example of engineering is the specialisation in either the business or personal travel markets. The first retailer will locate in an office complex while the latter will look for a shopping mall location. Engineering approaches at the retail level are most common in the corporate travel segment where management reporting and analysis of expenses is as important to corporate management as the travel services themselves. The Ryder Group, the current holder of the federal government travel service contract recently took over DND's travel management on a fee for service basis instead of commission.

The CBC has recently signed onto Air Canada's Corporate Travel System. In conjunction with the Apollo Computer Reservations System (CRS), and the CBC's travel agent cum travel manager, American Express, CBC frequent travellers will be booking trips themselves through their office or portable computers. The system is designed to let the typical booking be done in under two minutes. From the first group of thirty to be trained to use

the system, half of the Corporation's 2000 frequent travellers will be using the system within a year. Air Canada expects such systems will reduce its distribution costs. CBC expects to save money because the system "forces" use of the preferred suppliers offering discounted rates while automating travel expense accounting.[16]

This brief survey would conclude that the type of management tools typical of engineering approaches in logistics are common and well established in the people distribution business.

Marketing Approaches in Passenger Logistics

Airlines use pricing to differentiate the identical aircraft seat for different market segments. Passengers requiring flexibility must pay more than a customer, generally a leisure traveller, who can meet the lower fare's restrictions (known as fences). Price can also be used to shift demand. The airline product is very capital intensive and requires high rates of daily utilization. Lower fares might attract a passenger to use a less desirable night flight. Airlines do segment their market (e.g "global traveller, road warrior, price driven occasional") but are less successful at product differentiation: "Flying is almost at a commodity level; there's no differentiation"[17].

---

[16]McArthur, Douglas, *CBC tests self-booking technology*, The Globe and Mail, Page A18, March 5, 1997

[17]Wilkie, Michael and Gleason, Mark, *Air travel study broke Burnett's hold on United*, Advertising World, Volume 67, Issue 44, *October*

Wholesalers, dealing with a more homogeneous market of holiday makers, will also use pricing to smooth demand with lower prices in the shoulder (versus peak) seasons or discounts for booking more in advance. Different wholesalers may focus on different segments to position themselves as low cost (e.g. Regent) or up scale (e.g. Horizon). Beyond location considerations for broad segmentation, retailers make little use of product differentiation for market segments. Some may specialize in certain ethnic markets or in limited product such as cruise agencies but for the most part offer the same products to most market segments.

Current corporate travel trends include higher costs but more sophisticated travellers are demanding increased service (Report on Business Travel, Globe and Mail, February 25, 1997). This trend offers opportunities for increasing product differentiation and for travel agents to provide value added services to corporate travellers too busy to consider doing it for themselves.

It is concluded that marketing tools are less employed in the logistics of passenger transportation. The corollary is to suggest opportunities may be found in marketing tools for greater competitive advantage.

Information Technology in Passenger Logistics

Passenger travel logistics is at the forefront of information technology. SABRE, from a tool of American Airlines, has grown into its own mega business

---

28, 1996.

as the world's largest IT system. It only draws 4.3% of its revenue from travel agents , the main users of the system, but has 59.1% from booking fees as a distribution service for the airlines and other service providers. 34.2% of its revenue now comes from IT services. In part due to new regulations restricting bias, it is expected that American will soon take Sabre public[18]. Technological development has been so rapid that the majority of users are far behind in making regular use of existing functionality. Future development will focus on refinement and user friendliness to extend use of what is already available. For example, the internet is still in less than 10% of Canadian homes. The web on the TV/phone stage has not yet been reached with a high degree of penetration, although such products are already seen at consumer "Home" shows (Winnipeg, March 8, 1997 observed the "Starphone", complete with screen). Time is still money and many consumers, especially affluent frequent travellers, value their time too much to become proficient with the technology. Where they value the relationship with their agent, they will continue to book with an agent. After all, direct booking is not new. Airlines have always had phone lines to their reservations centres, yet travel agents sell 80% of tickets. Now that airlines have progressed to "Call Centres", the level of service can be so poor that more consumers will phone travel agents.

Government regulation is also involved with

---

[18]Flint, Perry, *Sabre Unleashed,* Air Transport World, pp95-96, November 1996

technology. Policy is to rely on competitive forces. This trend will support the independence of agents. The primary recent example is regulations issued for Computer Reservations Systems to eliminate bias[19]. As noted above, the prediction is that the large CRS's will be spun off as separate companies[20].

Last year, Cathay Pacific used its web site to "auction" a block of business class seats. The internet offers the potential for airlines to sell almost 100% of their seats. Salomon Brothers Vice President Julius Maldutis believes the airline industry is on the verge of a "third revolution" of new profitability that will far exceed the gains of yield management. "Cyberauctions" hold the potential of 100% load factors, to solve the airline's traditional challenge of perishability of an empty airline seat[21]. But "it is too early to offer travel agents the last rites". The biggest web site is *Travelocity*, a combination of Sabre and 200,000 pages of destination information

---

[19]Hyndman, Louis D., Chair, *Directions, the Final Report of the Royal Commission on National Passenger Transportation,* Supply and Services Canada, Ottawa, 1992. P.23

[20]Flint, Perry and Henderson, Danna K., *Crandall speaks his mind,* Air Transport World, pp42-43, March 1997

[21]Flint, Perry, *Cyber hope or cyber hype?,* Air Transport World, pp25-27. October 1996

from Worldview Systems. With reservations capability on 370 of the 700 airlines represented, it is still overwhelmingly used for browsing and planning but not for transactions. Successful travel agencies will not fight technology but instead can make it work for them. A retail agency could provide terminals for customers to browse just as they now flip through brochures. Half the agent's job may be done by the time the customer sits down. "Technology offers agents the opportunity to reinvent themselves.... to stop the focus on price and product and to concentrate on value added service"[22], to become information providers and interpreters, at a fee. This view is shared by John Berger, owner of ABC travel in Monroe NY: "What takes consumers a half hour to do on their PCs takes us two minutes"[23]. After all, as Murphy noted, "anyone can make a mistake; it takes a computer to really screw up". Airlines and credit card companies are also discovering the perils of on-line distribution which is "rife with opportunities for electronic mischief making and theft"[24].

It can be seen that Information Technology is

---

[22]*Internet threatens agents' survival,* Asian Business, September 1996

[23]Hoffman, Thomas, *Ticketless Travelers*, ComputerWorld, p.69, July 22, 1996

[24]Flint, Perry, *E-Burglary,* Air Transport World, pp.32-33, November 1996

widely employed in passenger logistics but much opportunity remains to refine its functionality while extending its reach.

Travel Retailing in 2005

The debate over the travel agent's value in the distribution channel has centred on the question: are they only "processors" or are they truly "counsellors"? According to the 1995 National Travel Monitor, 28% of American adults used a travel agent in the previous year. Users tend to be older, more affluent and better educated. "Travel agents will clearly have to reinvent themselves to remain competitive past the millennium. "However, there will be no shortage of consumers who are prepared to pay for their expertise"[25].

Retail travel agents will become more independent of service providers, the suppliers in passenger logistics. The previous partner arrangements (e.g. preferred suppliers) are eroding due to greater competition as providers and wholesalers increasingly market directly to the customer via the internet and call centres. .

Retail travel agents will be less a level in the distribution channel and become more of a facilitator. They will be travel managers for corporations managing their travel costs and consultants for individual leisure

---

[25]Yesawich. Peter C. Ph.D., *Travel Agents: Dinosaurs or Divas*, Lodging Hospitality, p.18. February 1996

travellers sorting through the complexity of the business. Their independence from travel suppliers will also increase as they are increasingly paid by the customer for their professional expertise. The agent marketing to consumers will use their independence from suppliers to broaden their product lines or to specialize with increasing differentiation. Professionalism will increase in importance as the product will shift from being perceived as a trip to more as a trip planning service. Both ACTA and CITC (Canadian Institute of Travel Counsellors) will expand their training and on-going education programs; professional association certification will become essential

An alternative view of retail travel service is the anticipated entry of the "category killer". A precursor may be the store in Texas called *Travel Fest* which sells travel gear, books and software as well as full travel services; it is also open 9 a.m. to 11 p.m., seven days a week[26]. Terry Ohman, Executive Director of ACTA, suggests that retail agencies will supplement their travel services with high volume/low margin products such as event tickets.

Travel agencies experienced growth and profits from the rapid changes following deregulation but the emerging competition related to technology will lead to a

---

[26]Green, Judson, *The Travel Agency Business - The Family Market,* Vital Speeches of the Day, Volume 62, Issue 8, pp247-250, February 1, 1996

major transition. Today's travel agent may become largely extinct as technology permits further decentralization of information dispensing and clerical functions to the household or office. However, "the more personalized aspects of the travel agent's advising function would seem relatively impervious to such developments "[27].. Fundamental attention to service is a core value of logistics and of successful travel agents. Service can not be replaced by technology. Instead, technology needs to be harnessed to enhance service. The video telephone will soon place the customer face to face even when they are far away.

Levine and Luck draw the conclusion that the key to success in business is to "focus on changing what matters to customers". Will the passenger logistics business, and particularly the travel agent, succeed in implementing necessary change? Levine and Luck also note that "firms losing competitiveness are more likely to risk the organizational pain that comes from adopting this paradigm"[28]. We can not say whether the retail travel

---

[27]Meyer, John R. and Oster, Clinton V. Jr., *Deregulation and the Future of Intercity Passenger Travel,* Massachusetts Institute of Technology, 1987 (Chapter 8. *Changes in Distribution Channels and the Travel Agency Business)*

[28]Levine, Arnold and Luck, Jeff, *The New Management Paradigm, A review of Principles and Practices,* The RAND Institute, 1994

industry will succeed in its re-invention of itself, incorporating the lessons of lean logistics. It can be concluded that they should have the motivation.

MANITOBA-RELATED TRUCKING IN THE MID-CONTINENT CORRIDOR

Jeannette Montúfar[1] and Alan Clayton[2]

Introduction

This paper analyzes Manitoba-related trucking activity and freight movements to and from the U.S. through the Mid-continent corridor. Detailed results of two surveys are presented. The first is a survey of 3,922 trucks at the major Manitoba-U.S. border crossing of Emerson-Pembina, and the second is a survey of nine of the major Manitoba-based truck load carriers.

The Mid-continent corridor runs from Manitoba to Mexico through the Mid-continent states of North Dakota, South Dakota, Nebraska, Minnesota, Iowa, Missouri, Kansas, Oklahoma, and Texas. The total distance of the corridor from Winnipeg, Manitoba to Monterrey, Nuevo León is 3,000 kilometers.

The principal road networks serving Manitoba and the corridor area consist of Manitoba Provincial Trunk Highways, and in the U.S., Interstate System highways, National Highway System highways, and National Network highways.

The Quantity of Manitoba-U.S. Crossborder Trucking Activity

Total crossings of the Manitoba-U.S. border averaged 1,294 trucks per day in both directions in 1996 (472,310 trucks for the year). This represents about 20 percent of all truck crossings of the western U.S.-Canada border and about 4 percent of all truck crossings of the Canada-U.S. border.

There are 17 highway crossings on the Manitoba-U.S. border. Twelve of these crossings are on the Manitoba-North Dakota border, and 5 are on the Manitoba-Minnesota border. In 1996, 8 of the 17 crossings accounted for 96 percent of total truck traffic crossing the Manitoba-U.S. border. These

are: (1) West Hope-Coulter--31 trucks per day; (2) Dunseith-Peace Garden--91 trucks per day; (3)Walhalla-Winkler--78 trucks per day; (4) Neche-Gretna--50 trucks per day; (5) Pembina-Emerson--813 trucks per day; (6) Lancaster-Tolstoi--40 trucks per day; (7) Roseau-South Junction--41 trucks per day; and (8) Warroad-Sprague--103 trucks per day (the number of trucks per day corresponding to each of the crossings represents two-way truck volumes for 1996). Truck traffic crossing the Manitoba-U.S. border between 1991 and 1996 has grown at an average rate of 14.3 percent per year.

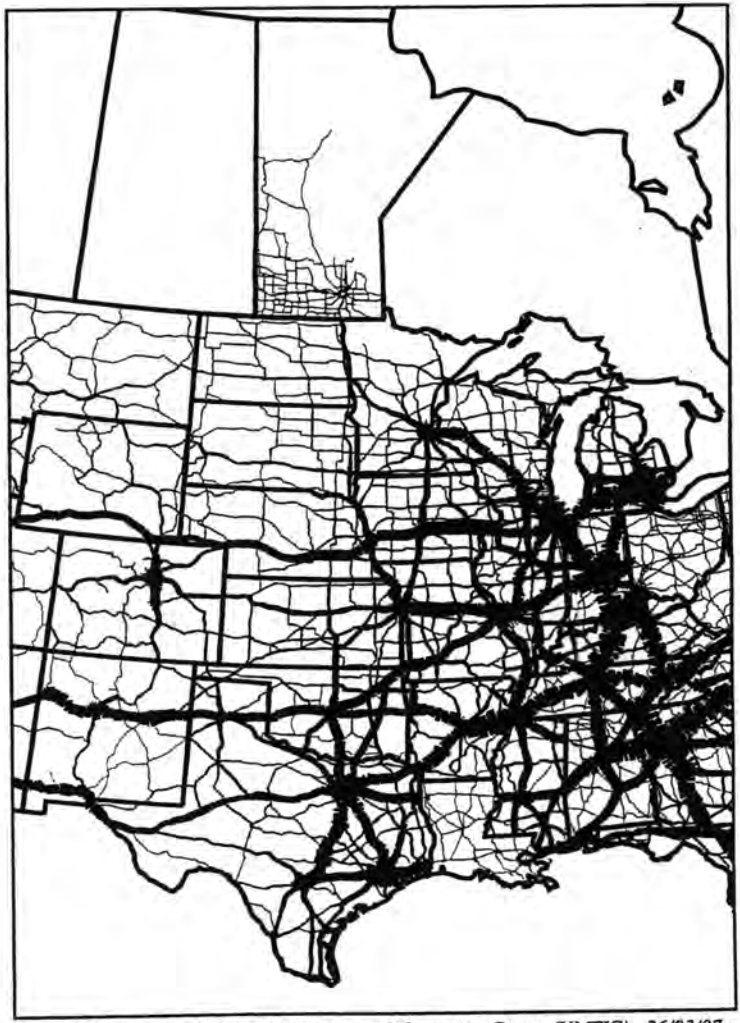
The Pembina-Emerson crossing, located on I-29 (on the U.S. side) and on Provincial Trunk Highway 75 (on the Canada side), is the highest volume crossing on the Manitoba-U.S. border. After Blaine, British Columbia, it is the second busiest crossing on the western border. This crossing averaged a two-way total of 813 trucks per day in 1996. This volume is 10 percent higher than in 1995.

The Warroad-Sprague border crossing is the second highest volume crossing on the Manitoba-U.S. border, with a two-way truck traffic of about 100 trucks per day in 1996. Being a short-cut for Canadian east-west traffic, much of the truck movement (about 80 trucks per day in both directions) is in-transit between eastern and western Canada. These trucks are primarily vans and some specialized bodies. [3]

The remaining 15 crossings account for about one-quarter of truck crossings on the Manitoba-U.S. border. Truck characteristics at most of these crossings are primarily controlled by the U.S. Federal law, U.S. regulations, and in particular those of North Dakota, South Dakota, and Minnesota.

Based on state and provincial data, a truck flow map was developed for all National Highway System highways in the Mid central U.S. and all highways of the Provincial Trunk Highway system in Manitoba.[4] Figure 1 shows the truck flows for the Mid central U.S. and for Manitoba. The highest truck volumes in Manitoba occur on Highway 1 west of Winnipeg. These volumes are equivalent to those observed on I-29 between Grand Forks and Fargo (approximately 1,000 trucks per day). Truck volumes in Manitoba are minor compared to the volumes registered south of Fargo, North Dakota. The lowest truck volume along the I-29 route occurs at the Manitoba-U.S. border. The highest truck volume occurs just north of Omaha, Nebraska--approximately 2,980 Average Annual Daily Truck Traffic (AADTT). On the I-35 segment of the corridor, the lowest truck

volumes occur at its two ends and in its middle (between Minneapolis and Duluth--about 1,450 AADTT; on the east side of Wichita--about 1,750 AADTT; and between Laredo and Cotulla, Texas--about 1,700 AADTT).



Source: University of Manitoba Transport Information Group (UMTIG). 26/03/97

Figure 1 - Commercial Vehicle Flows

Truck traffic between Kansas City, Missouri and St. Louis, Missouri is about 7,500 trucks per day. Between Fargo, North Dakota and Minneapolis, Minnesota, truck traffic is about 2,200 trucks per day.

Types of Manitoba-U.S. Trucking

There are five categories of Manitoba-related trucking across the western U.S.-Canada border. These are: (1) local-regional trade with North Dakota, South Dakota, Minnesota, and Iowa, principally involving farm products; (2) Manitoba western oriented trade, most of which does not take place across the Manitoba-U.S. border but through crossings in Saskatchewan and Alberta. This trade principally involves specialized bulk commodity movements, typically at gross vehicle weights greater than 80,000 pounds; (3) truck load (TL) movements into the U.S. Midwest, the Great Lake region and north east region. These movements are generally restricted to gross vehicle weights (GVWs) of 80,000 pounds; (4) trucking destined for places south of Kansas City and in some cases into Mexico; and (5) trucking to and from eastern Canada, in-transit through the U.S. These movements typically occur at gross vehicle weights of 80,000 pounds.

A TL carrier survey was conducted in Winnipeg, Manitoba in October, 1996.[5] Nine of the major Manitoba-based carriers were interviewed to obtain information regarding origins-destinations, commodities hauled, equipment used, method of operation, intermodal operations, activity levels, and other information regarding trucking between Manitoba and the U.S.

Most of the carriers specialize in truckload movements from Canada to the U.S. and across Canada. The quantity of movements across Canada is not as high as the quantity of movements across the Manitoba-U.S. border.

Origin-Destination and Routing

For eight of the nine carriers interviewed, more than one-half of their operations originate or terminate in the U.S. The major destinations of these carriers are the U.S. Midwest, northeast, and the Great Lake Region (mainly Minnesota, Iowa, Illinois, Wisconsin, Indiana, Michigan, New York, Ohio, Kentucky, Iowa, Georgia, and Tennessee).

Eight of the nine carriers also have operations that are destined for places south of Kansas City. However, these are minor in terms of quantity of

shipments. One of the carriers mentioned that one of the reasons why they have very few movements to places south of Kansas City is that the southern market is completely dominated by U.S. carriers such as Schneider and J.B. Hunt. Of the eight carriers that operate south of Kansas City, four have the U.S.-Mexico border as the major destination/origin. Loads are taken from Manitoba or Toronto to: (1) Texas--to Laredo, El Paso, and McAllen; (2) Arizona--to Nogales (approximately three-quarters of all produce consumed in the U.S. and Canada during the winter is imported through Nogales)][6]; and (3) California. There are very few movements to places directly south of Manitoba (North Dakota and South Dakota), and almost no movements to the west of the U.S.

Only one of the carriers has operations into Mexico. The other carriers either: (1) are not attracted to the idea of starting businesses in Mexico; (2) have not had the contacts to start operations into Mexico; or (3) do not want to go into Mexico because they think that operating costs are higher due to a lot of hidden costs involved with these operations (for example, damage of trailers).

Most of the carriers travel on I-29 from the Manitoba-U.S. border to Fargo, North Dakota and mainly on I-94 from Fargo through Minneapolis to the destinations in the U.S. Midwest or northeast. Traffic running between Winnipeg and eastern Canada is often routed across U.S. Route 2. This is because: (1) it is shorter in distance; (2) the travel time is shorter; (3) the roads are in better condition; (4) the operating cost is lower than running on the TransCanada Highway; (5) fuel in the U.S. is cheaper; and (6) the roads are safer in winter.

Most of these in-transit operations take place across the Pembina-Emerson crossing. However, some trucks move across the Roseau-South Junction crossing and the Warroad-Sprague crossing. The selection of the crossing depends on: (1) the destination of the load; (2) the origin of the load; (3) customer requests; (4) where loading took place; and (5) the weight of the load. For example, the operation of one of the carriers surveyed is as follows: (1) if a load is going from Winnipeg to Chicago, it crosses at the Roseau-South Junction port of entry; (2) if the load is moving in transit from Ontario to Winnipeg, the Warroad-Sprague port of entry is used; and (3) if the load is moving into Manitoba and it has not been pre-cleared, then it goes through the Pembina-Emerson crossing. Two other carriers discussed the crossing selection based on weight. These carriers use mostly the Pembina-

Emerson crossing, except when the vehicles are overweight. If it is an in-transit operation, the Warroad-Sprague crossing is used. If it is a southbound operation, the Gretna-Neche crossing is used.

Triangulation Operations

For all the carriers, most of the operations into the U.S. involve triangulation. A basic triangulation operation from a terminal in Winnipeg consists of the following: (1) Winnipeg with paper to Wisconsin; (2) building materials pick up in Wisconsin for drop off in Toronto; (3) retail merchandise pick up in Toronto for drop off in Winnipeg.

A different type of triangulation operation involves: (1) Manitoba with farm equipment to Kansas; (2) the vehicle dead-heads from Kansas to Nebraska; (3) load pick up in Nebraska for drop off in western Canada; (4) lumber pick up in western Canada for drop off in Manitoba.

Fleet, Equipment and Truck Size and Weight (TS&W) Issues

The nine carriers operate a combined fleet of about 2,160 tractors and 4,425 trailers, for an average trailer-to-tractor ratio of about 2. Approximately 80 percent of the trailers are vans. More than one-quarter of these are equipped with refrigerating units. Flat beds account for about 10 percent of the trailer fleet. More than one-half of the trailers are 53 feet in length, about 45 percent are 48 feet and the remaining are other lengths, less than 48 feet. All new semitrailers in the past year have been 53 footers for most of the companies. The standard trailer width is 102 inches and the standard van height is 13.5 feet. Most of the fleet consists of tandem-axle semitrailers.

During the past few years, all of the surveyed carriers have shifted to lighter equipment. With this lighter equipment, typical payloads range from 43,000 pounds to about 48,000 pounds. At 80,000 pounds GVW, this allows a tare weight of about 32,000 to 37,000 pounds. All carriers surveyed are running average tare weights of approximately 33,500 pounds for the unit.

Three of the surveyed carriers operate tridem-axle semitrailers into the U.S. Typical payloads with this type of trailers range from 44,000 pounds to 57,000 pounds. One of these carriers thinks that with the use of tridems, there is more flexibility on axle requirements in terms of weight. The other six carriers either do not operate tridems at all or only operate them within

Canada because: (1) they are not allowed in certain States; (2) "most of the U.S. is restricted to 80,000 pounds"; (3) "the extra axle decreases the miles per gallon by about 10 to 12 percent"; or (4) "the wear in the middle axle is excessive".

Three of the carriers operate B-trains, mainly across Canada (some B-trains are taken into North Dakota). Another carrier operates rocky mountain doubles from Winnipeg to Alberta. None of the carriers has considered the possibility to start operating triple trailer combinations. The main reason is that in their opinion, triples are good for LTL movements, not for TL. In nearly all the cases, the less-than-truck load (LTL) component of their operations accounts for 10 to 15 percent of the total.

In terms of gross vehicle weight, all companies operate most of the time at 80,000 pounds or less everywhere in the U.S. The 10 percent winter weight allowance in Minnesota is of no consequence to most carriers because most of the operations (1) are not destined for Minnesota; (2) are destined for states that have a maximum allowable GVW of 80,000 pounds; or (3) need to go through States that have a maximum allowable GVW of 80,000 pounds. There is one carrier that makes use of this 10 percent weight allowance but only in its intermodal operation. In this operation, between 10 and 15 tractors with vans are used to haul agricultural products from Minnesota and North Dakota to Winnipeg. The vans are then shipped by rail from Winnipeg to eastern Canada, for final delivery, by truck, in the eastern U.S. In winter, the company operates at GVWs of about 85,000 pounds only in the intermodal operation.

In terms of split tandems (these are axles created by increasing the spacing between the two axles in a tandem group from a typical standard of about 50 inches to 8, 9, or 10 feet), most carriers do not use them, and think that at the moment this type of axle configuration is not critical to their operations. One of the carriers makes use of split tandems in its Ontario operations. Another carrier makes use of split tandems in its operations within the U.S. Five of the nine carriers think that if this type of axle arrangement were allowed in Manitoba, it could increase the company's opportunities in the movement of freight. They also think that split tandems are useful because: (1) there is more flexibility of loading (specially for weight-out carriers); and (2) the payload can increase. The other four carriers would not make use of split tandems in their operations mainly because the fleet would become "equipment specific", therefore restricting its operations and access area.

Commodity Movements Across the Manitoba-U.S. Border

According to the TL carrier survey, the major commodities moved into the U.S. include farm machinery, construction equipment, steel, bulk grain, fertilizer, auto parts, furniture, paper, chemicals, hay, retail merchandise, meat (from Alberta), frozen french fries (from Manitoba), agricultural commodities, dry food products, peat moss, fish, and other general freight.

The major commodities moved into Canada from the U.S. are produce (from Nogales, Arizona), books, electrical appliances, fertilizer, construction equipment, farm products, farm equipment, lumber (from Washington), and parts for manufacturing (from the Ohio area).

An extensive survey (three parts) was conducted at the weigh scale in Emerson, Manitoba. This scale is located approximately one-half kilometer north of the U.S.-Manitoba border. The first part of the survey was conducted from February 12 to February 18, 1996. The second part took place from June 5 to June 15, 1996; and the last part was conducted from August 14 to August 15, 1996. All days of the week were covered by each survey, as well as the times of the day. The purpose of the survey was to obtain information regarding: (1) origin-destination of trucks going through the scale; (2) name of carriers; (3) base of operation of those carriers; (3) vehicle configuration; (5) body types of the vehicles; (6) loading conditions; (7) axle group weights; (8) commodities being hauled; and (9) routing from origin to destination.

A total of 3,922 trucks were classified in this survey. This represents almost one-quarter of the total truck traffic through the Emerson-Pembina crossing during the survey period. Of the 3,922 trucks classified, 966 of the truck drivers were personally interviewed. This represents 6 percent of the movements through the Emerson-Pembina crossing for the survey period.

Of the 3,922 trucks that were classified, 2,687 (68.5 percent) were loaded, 774 (19.7 percent) were empty, and loading conditions were unknown for the remaining 461 (11.8 percent) trucks.

Since this section presents findings pertaining to commodity movements including their origin-destination patterns, only the loaded trucks are considered in the analysis.

Of the 2,687 loaded trucks, 1,167 were traveling northbound and 1,292 were traveling southbound. Direction of travel was unknown for the remaining trucks.

Of the 1,167 loaded trucks traveling northbound, almost 10 percent haul machinery; 7 percent haul farm products; 5 percent haul paper and paper products; 4 percent haul animal feed; another 4 percent haul lumber and fabricated wood; other types of commodities account for about 25 percent of the movements. It was not possible to determine the commodity being hauled for 40 percent of these trucks.

Origin-destination information was obtained for 525 (45 percent) of the 1,167 loaded trucks traveling northbound through the Emerson scale. From the 525 trucks, it was observed that more than one-quarter of the movements originate in Minnesota (mainly Minneapolis); 12 percent originate in North Dakota (mainly Fargo); another 12 percent originate in Illinois (mainly Chicago); 6 percent originate in both Wisconsin and Iowa. The remaining 38 percent of the movements originate in other States or Provinces. Figure 2 shows the origin-destination of movements through the Emerson scale for northbound loaded trucks, for all commodities combined. Also as illustrated in Figure 2, almost 85 percent of the movements are destined for Manitoba and the remainder are destined for the other western provinces.

Commodity-specific origin-destination information was also obtained from the survey. Machinery, farm products, and commodities moved in hopper bottom trailers (grain, fertilizer, and animal feed), account for about one third of all northbound movements through Emerson.

One-third of machinery movements originate in Illinois (mainly in Chicago); 13 percent originate in Minnesota; 11 percent in Iowa; 9 percent in North Dakota and 7 percent in South Dakota. The remaining 40 percent of the movements originate in Texas, Kansas, Oklahoma, Missouri, Wisconsin, Ohio, and other States. Nearly two-thirds of the movements are destined for Manitoba; 18 percent are destined for Saskatchewan; 13 percent for Alberta; and 4 percent are destined for British Columbia.

More than one-quarter of farm product movements originate in Minnesota (mainly in Minneapolis); nearly one-quarter originates in North Dakota (mainly in Fargo), and 11 percent in Texas. The remaining 39 percent of the movements originate in Ontario, California, Iowa, South Dakota, Nebraska,

Florida, and other States. Ninety percent of the movements are destined for Manitoba and the remainder are destined for the other western provinces.

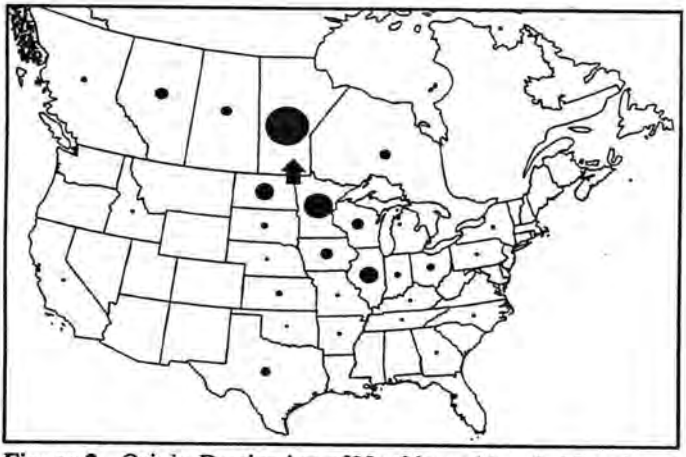


Figure 2 - Origin-Destination of Northbound Loaded Trucks Through Emerson (all commodities)

About 60 percent of the movements of commodities transported in hopper bottom trailers (here referred to as grain bodies) originate in Minnesota (mainly in Minneapolis); 16 percent originate in North Dakota (mainly Fargo and Minot); 12 percent in Iowa; 5 percent in South Dakota and 3 percent in Nebraska. The remaining 4 percent of the movements originate in Arizona and Arkansas. Effectively all of these movements are destined for Manitoba.

Of a total of 1,292 loaded trucks traveling southbound, 10 percent haul live animals; 7 percent haul machinery; another 7 percent haul lumber and fabricated wood; 6 percent haul fresh farm products; 3 percent haul paper and paper products; other types of commodities account for about 20 percent of the movements. It was not possible to determine the commodity being hauled for 48 percent of these trucks.

Origin-destination information was obtained for 385 (30 percent) of the 1,292 loaded trucks traveling southbound through the Emerson scale. From the 385 trucks it was observed that 85 percent of the movements originate in Manitoba; 9 percent originate in Saskatchewan; about 6 percent originate

in Alberta; and less than one percent originate in British Columbia. This is illustrated in Figure 3.

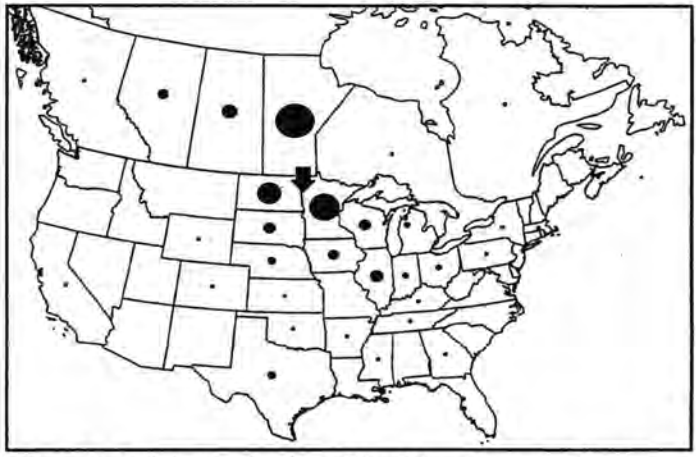


Figure 3 - Origin-Destination of Southbound Loaded Trucks Through Emerson (all commodities)

More than one-quarter of the movements are destined for Minnesota (mainly to Minneapolis); 16 percent are destined for North Dakota (mainly to Fargo); 6 percent are destined for each, Illinois, Wisconsin and South Dakota; 4 percent are destined for Iowa. About 38 percent of the southbound movements through the Emerson scale are destined for other states and provinces. This is also shown in Figure 3.

Live animals, machinery, and commodities moved in hopper bottom trailers (grain, fertilizer, and animal feed), account for about one third of all southbound movements through Emerson.

There are only two origins for livestock movements through Emerson. Nearly all of the live animal movements originate in Manitoba. The major destinations for live animal movements are: Minnesota (mainly Minneapolis)--one-third of the movements; South Dakota (mainly Sioux Falls)--29 percent; Iowa (mainly Sioux City)--11 percent; Nebraska--11 percent and North Dakota--9 percent. The remaining 7 percent of the movements are destined for Wisconsin, Kansas and Kentucky.

Eighty-seven percent of machinery movements originate in Manitoba; 6 percent originate in Alberta; 3 percent in Saskatchewan; and 3 percent in Quebec. Major destinations are North Dakota (mainly Fargo)--19 percent; Minnesota (mainly Minneapolis)--13 percent; Kentucky--10 percent; California--7 percent; and Ontario--6 percent. The remaining 45 percent of the machinery movements are destined for Nebraska, Kansas, Illinois and other States in the Midwest.

Eighty percent of commodities moved in hopper bottom trailers originate in Manitoba; 18 percent originate in Saskatchewan; and 2 percent in Alberta. Most of the hopper bottom movements are destined for Minnesota (principally Minneapolis)--55 percent; almost one-third are destined for North Dakota (mainly Fargo and Grand Forks); 5 percent for Iowa (mainly Sioux City); another 5 percent are destined for Wisconsin; and 3 percent are destined for Nebraska.

From the survey it is evident that most of the movements either originate or end in Manitoba and the U.S. Midwest. Major trading states with Manitoba are North Dakota, South Dakota, Minnesota, Illinois, Wisconsin and Iowa.

Truck Types and Truck Size and Weight Regulations

In addition to commodity movements, origin-destination patterns, and routing, the Emerson Scale survey classified trucks by configuration and body type. Information was also obtained regarding GVW of these trucks as they drove through the scale.

Truck characteristics at the Emerson-Pembina crossing are primarily controlled by the U.S. Federal law and State regulations, in particular, those of North Dakota, Minnesota, and the rest of the corridor states. Being on the I-29, this crossing is subject to the Intermodal SurfaceTransportation Efficiency Act (ISTEA) freeze of 105,500 pounds GVW and 103-foot cargo-carrying unit length for a truck tractor and two trailing units applicable for North Dakota. Truck movements east of I-29 between Manitoba and Minnesota are restricted to 80,000 pounds GVW and 88,000 GVW in the winter period.

About 86 percent of the trucks using this crossing are tractor-semitrailers, primarily with five axles. These vehicles and their loadings are primarily controlled by U.S. Bridge Formula B and the U.S. Federal weight limits of

80,000 pounds GVW, 20,000 pounds on a single axle and 34,000 pounds on a tandem axle.

Of the total 3,922 trucks that were classified, the fleet is as follows: 2-, 3-, and 4-axle straight trucks combined account for 2.7 percent; 5-axle tractor semitrailers (3-S2s) account for 85.7 percent; 6-axle tractor semitrailers (3-S3s) account for 2.1 percent; 7- and 8-axle B-trains combined account for 1.8 percent; A-trains are almost as common as B-trains. They account for about one percent of the total fleet. other configurations account for 6.7 percent.

In terms of body type, vans account for 42.7 percent of the fleet (one-quarter of those vans are equipped with refrigerating units); flat beds account for 18.9 percent of the trucks surveyed; grain bodies account for 16.3 percent; livestock trucks account for 9.5 percent; other body types account for 12.6 percent of the trucks surveyed.

Of the 319 hopper bottom trucks operating southbound, 80 percent were loaded, 5 percent were empty, and loading conditions were unknown for the remaining 15 percent. Of the 310 hopper bottom truck operating northbound, more than one-quarter were operating empty, about three-quarters were carrying a load, and loading conditions were unknown for less than one percent of those trucks. The direction of operation was unknown for the remaining 8 percent of hopper bottom trucks operating through the Emerson scale.

In addition, there are some longer combination vehicle (LCV) operations which were not captured in this survey. LCVs operate to and from Manitoba, North Dakota, South Dakota, into Sioux City, Iowa.

In terms of weight distribution, it was found that for loaded 3-S2s operating northbound, the GVW ranged from a low of 15,500 kg (34,100 pounds) to a high of 47,100 kg (103, 837 pounds). The average GVW for this units was 30,650 kg (67,570 pounds). Thirteen percent of these vehicles operated at a GVW greater than 80,000 pounds. For 3-S3s operating northbound, the GVW distribution ranged from a low of 28,800 kg (63,490 pounds) to a high of 45,500 kg (100,310 pounds). The average GVW was 37,435 kg (82,530 pounds). Ten percent operated at a GVW greater than 90,000 pounds.

For eight-axle B-trains operating northbound, the GVW ranged from a low

of 28,900 kg (63,712 pounds) to a high of 48,300 kg (106,482 pounds). The average GVW was 41,830 kg (92,225 pounds). One of nine trucks going through the scale operated at a GVW greater than 105,000 pounds.

For southbound moving vehicles, it was found that the GVW of 3-S2s ranged from a low of 15,400 kg (23,880 pounds) to a high of 43,300 kg (95,460 pounds). The average GVW was 32,600 kg (71,870 pounds). Nearly 20 percent operated at more than 80,000 pounds GVW. For 3-S3s, the GVW ranged from a low of 28,700 kg (63,272 pounds) to a high of 41,600 kg (91,712 pounds). The average GVW was found to be 37,500 kg (82,706 pounds). Eight percent operated at more than 90,000 pounds GVW.

Finally, in the case of eight-axle B-trains, GVW ranged from a low of 31,000 kg (68,343 pounds) to a high of 48,700 kg (107,365 pounds). The average GVW was 45,475 kg (100,251 pounds). Four of nineteen trucks going through the scale operated at more than 105,000 pounds GVW.

Concluding Remarks

On an average day, truck movements crossing the Manitoba-United States border are more than twice as frequent as truck movements crossing the Manitoba-Ontario border and somewhat greater than the number of truck movements crossing the Manitoba-Saskatchewan border. There are as many trucks crossing the Manitoba-U.S. border as cross the combination of the Saskatchewan-U.S. and Alberta-U.S. borders. Manitoba-U.S. truck crossings have grown at an average rate of more than 14 percent per year for the past five year period. On the other hand, east-west truck traffic to and from Manitoba is growing only marginally.

The size and growth of Manitoba-U.S. trucking suggests that the province's highway transport policy interests and highway development efforts should and probably will increasingly focus on U.S. highway and transportation policy considerations.

Examples of developments in Manitoba's transport policies which could benefit Manitoba's trade and transport into and through the Mid-continent corridor are: (1) greater harmonization of Manitoba's TS&W laws with U.S.

TS&W laws; (2) increased cooperation with North Dakota and Minnesota on truck enforcement programs; (3) enhancement of the Emerson border crossing scale site using advanced intelligent transport system technologies to minimize the need for trucks to stop; (4) harmonizing City of Winnipeg TS&W regulations with provincial regulations, and making "transparent" special permitting requirements and services for trucks operating between the U.S., the province, and Winnipeg.

In the area of TS&W regulations, Manitoba should consider allowing for the use of split tandem axles; 14-foot vehicle heights; air-suspension lift axles; short wheelbase tractors which could be combined with 53-foot semitrailers within an overall length restriction of 65 feet; 57-foot semitrailers on major routes; and any vehicle which complies with the requirements of U.S. Bridge Formula B (given acceptable safety performance and bridge loading levels). These developments would enhance the compatibility of Manitoba-U.S. TS&W regulations, and facilitate more efficient cross border trucking. As a final observation, promoting an expanded highway network and normalization of regulations relating to the crossborder movement of long combination vehicles (rocky mountain double, turnpike double and triple trailer combinations) could also be of significant benefit to Manitoba.

Endnotes

1. Jeannette Montufar, Ph.D. student. Faculty of Civil Engineering, University of Manitoba.
2. Alan Clayton, professor. Faculty of Civil Engineering, University of Manitoba.
3. Intermodal Surface Transportation Efficiency Act: Section 6015 "Profiles of Western U.S.-Canada Border Crossings", FHWA-PL-94-009-017. September, 1993.
4. University of Manitoba Transport Information Group
5. Montufar, Jeannette. "Trucking and Size and Weight Regulations in the Mid-Continent Corridor". Master of Science Thesis, University of Manitoba. November, 1996.
6. Intermodal Surface Transportation Efficiency Act: Section 6015 "Assessment of Border Crossings and Transportation Corridors for North American Trade (West)", FHWA-PL-94-009-018.

THE "GRAIN TRAIN" IN WASHINGTON STATE: AN EVALUATION OF INTERNAL AND EXTERNAL BENEFITS

Kenneth L. Casavant, Washington State University
Richard Mack, Central Washington State University
Gary Richardson, Central Washington State University

A current issue affecting all railroads, but of special interest to shortlines, is the availability of rail cars to operate on those lines. It does little good to save trackage if the rolling stock (cars and locomotives) are not available to serve shippers. The rail car supply problem limits a region's ability to consistently supply grain and other products to world buyers and to pass price increases back to our farmers. This loss of capacity in the rail sector of the multimodal system can cause increased shipper costs, road damage, excessive energy utilization, safety concerns, etc.

The Grain Train project, discussed in greater detail later in this paper, involved the purchase of 29 state-owned rail grain cars. Those cars are utilized by the members of a voluntary shippers association.

OBJECTIVES OF RESEARCH

The overall goals of the grain train project are to preserve rail service in a multimodal system, provide additional rail cars to alleviate existing shortages, and to redistribute energy rebate funds in accordance with acknowledged constraints. The general purpose of this research is to analyze and evaluate the operation of the Washington State Department of Transportation (WSDOT) purchased covered hopper rail cars for service to shippers in eastern Washington during calendar year 1995.

EFFICIENCIES, UTILIZATION, AND PERFORMANCE

Physical Utilization

Grain train carloads by the month for 1995, as reported by WSDOT, are shown in Table 1. Rates on the Union Pacific (UP) system

are based on 25-car blocks called unit trains. All shipments were comprised of this size unit except for several larger trains in February, August, and October, as well as a four-car unit that was tagged to non-WSDOT grain train cars in June.

Table 1. WSDOT Grain Train Carloads, by Month, 1995.

Month	Grain Train Carloads	Month	Grain Train Carloads
January	49	July	0
February	50	August	48
March	25	September	26
April	48	October	29
May	50	November	53
June	17	December	28
Six Month Subtotals	**239**		**184**
1995 TOTAL			**423**

There were 326 car-months of service available in 1995, rather than the 348 car-months if all cars had been available all year. The preliminary target was 1.5 loadings per car per month, or 489 carloads. The 423 carloads that were achieved represent 86.5% of the initial target.

Financial Performance

The purpose of this analysis was to examine the current financial performance of the grain car project and to make general forecasts of anticipated future profitability. The fleet description is shown in Table 2.

Revenues and Expenses

Over the first year of this project, the total operating revenues were $173,805.09. Operating expenses totaled $20,112.83. Revenues minus expenses amounted to $153,692.26. We assume that revenues per car will decline by 1% per year of age. Over the past 70 years, annual inflation, based on the CPI, has averaged 3.1%. We used this rate to increase both operating expenses and purchase prices for grain cars when the existing cars are assumed to be scrapped.

Table 2. Fleet Description.

Total Cars*	29
Average Age	17 Years
Remaining Life Expectancy	23 Years
Purchase Price	$634,074.02
Capital Improvements	$119,957.02
Total Capital Expenditure	$754,031.04

Investment theory suggests that today's 25-year bond yields are the market's consensus opinion for the yield on one year investments, 24 years from now. This would suggest that the market consensus for one year investments at the end of the life of the existing cars should be slightly over 6%. Given the room for error that is inherent in any long-run projection of interest rates, we have assumed that the excess funds will be able to be invested at an average rate of 4.75% over the next 23 years.

Findings

Operating revenues are found to be more than adequate to cover operating expenses while also increasing the investment account to an amount adequate to make the project self-sustaining (i.e., to replace existing cars at the end of their useful lives). The investment account reaches its threshold amount in the ninth year and replacement/additional cars begin to be purchased out of project cash flows beginning in the tenth year. Tables 3 and 4 detail revenues, expenses, profits, and car purchases over the life of the project.

Sensitivity Analysis

We evaluated the sensitivity of our results to changes in revenues, growth rate in expenses, reinvestment rate, and rate of decline in revenues. In each scenario, we manipulated only a single variable while holding all else constant. Although this project was not financed with debt, we also projected the results under the assumption that the state had issued municipal bonds to acquire the necessary funding. Table 5 clearly demonstrates the viability of the grain car project.

Table 3. Revenue/Cost Streams and Replacement Potential, No Interest, WSDOT Grain Cars.

	Original Values	Year 1	Year 8	Year 16	Year 23
Revenues		$173,805.09	$161,997.70	$371,129.01	$672,615.41
Operating Expenses		$20,112.83	$24,904.99	$78,938.75	$190,062.50
Interest Expense		$0.00	$0.00	$0.00	$0.00
Profit		$153,692.26	$137,092.71	$292,190.26	$482,552.90
Number of Cars	29	29	29	72	140
Revenue (per car)	$5,993.28	$5,993.28	$5,586.13	$5,154.57	$4,804.40
Operating Expense (per car)	$693.55	$693.55	$858.79	$1,096.37	$1,357.59
Interest Expense (per car)	$0.00	$0.00	$0.00	$0.00	$0.00
Profit (per car)	$5,299.73	$5,299.73	$4,727.33	$4,058.20	$3,446.81
Running Total Profit	$143,395.05	$143,395.05	$1,367,349.56	$1,542,166.34	$1,558,993.50
Total Replacement Cost	$763,032.04	$777,406.00	$974,119.19	$1,243,602.02	$1,539,898.43
Per Car Replacement Cost	$26,001.07	$26,807.10	$33,194.07	$42,376.97	$52,473.56
Excess Over Total Replacement Requirement	$0.00	($1,378,338.18)	($154,383.67)	$20,433.11	$37,260.27
Total Add'l. Cars Purchased		0	0	52	121

Table 4. Revenue/Cost Streams and Replacement Potential, With Interest, WSDOT Grain Cars.

	Original Values	Year 1	Year 8	Year 16	Year 23
Revenues		$103,695.00	$96,650.52	$89,183.75	$83,125.08
Operating Expenses		$20,112.83	$62,903.99	$69,793.77	$77,369.09
Interest Expense		$37,999.00	$37,999.00	$37,999.00	$37,999.00
Profit		$45,583.17	$33,746.53	$19,389.98	$5,756.00
Number of Cars	29	29	29	29	29
Revenue (per car)	$5,993.28	$3,575.69	$3,332.78	$3,075.30	$2,866.38
Operating Expense (per car)	$693.55	$693.55	$2,169.10	$2,406.68	$2,667.90
Interest Expense (per car)	$0.00	$1,310.31	$1,310.31	$1,310.31	$1,310.31
Profit (per car)	$5,299.73	$1,571.83	$1,163.67	$668.62	$198.48
Running Total Profit	$143,395.05	$143,395.05	$565,190.81	$1,067,207.12	$1,521,786.84
Total Replacement Cost	$763,032.04	$777,406.00	$974,119.19	$1,243,602.02	$1,539,898.43
Per Car Replacement Cost	$26,001.07	$26,807.10	$33,194.07	$42,376.97	$52,473.56
Excess Over Total Replacement Requirement	$0.00	($1,378,338.18)	($956,542.42)	($454,526.11)	$53.61
Total Add'l. Cars Purchased		0	0	0	1

Table 5. Sensitivity Analysis.

Variable	Boundary	Year Cars Begin To Be Added From Cash Flows
No Interest Charges		
Beginning Revenues	$64,865	23
Rate of Increase in Expenses	18.87%	13
Reinvestment Rate	0%	12
Annual Rate of Decline in Revenues	13.84%	23
With Interest Charges		
Beginning Revenues	$103,695	23
Rate of Increase in Expenses	13.87%	17
Reinvestment Rate	0%	16
Annual Rate of Decline in Revenues	7.05%	23

REGIONAL AND STATE IMPACTS

Rail Service Preservation

WSDOT has a mission, defined in RCW 47.76, to encourage preservation of light density rail lines and thereby provide shippers with competitive transportation alternatives. The impact of losing rail service is far broader than just the grain industry. Consider the following sequential relationship.

The overall goal is rail service preservation as part of a multimodal transportation system. Given the regulatory freedom existing for Class I railroads, abandonment of light density lines can be expected to continue. An attractive and successful alternative to abandonment is "selling off" of the lines or the creation of shortline railroads such as the Blue Mountain Railroad, the Washington Central Railroad, or the Eastern Idaho Railroad. These shortline railroads, if rail cars are available, provide grain service to shippers. The grain service produces revenue for the shortline railroad and allows continuing service on that line to not just grain shippers, but to other shippers, the general agricultural and rural community, and even to those entities working on

rural roads and economic development. This interconnected relationship is complex, but definite.

Highway Damage Savings

Highway use by heavily loaded trucks has long been associated with damages to road surfaces and understructures. When rail substitutes for trucking, damages to county roads and state highways are avoided. Damages caused by grain trucks with GVWs that range in gross weight from 80,000 to 105,000 pounds can be readily quantified for both the truck ton miles actually displaced by the grain train in 1995 and for the potential displacement if loadings occur at optimal 1.5 carloads per month discussed earlier.

The methodology used in this section of the analysis is straightforward and traditional; estimates of grain tonnage that would have been transported by truck from elevators to river ports are converted into ton miles travelled on state and county roads, and multiplied by coefficients that express road damages per ton mile travelled. Road damage coefficients of $0.071 per ton mile for state highways and $0.1065 per ton mile for county roads were obtained from a previous research study by Ken Casavant and Jerry Lenzi titled, "Procedure for Predicting and Estimating the Impact of Rail Line Abandonments on Washington Roads, Final Report--1989."

The first two columns of Table 6 show distances and ton miles shipped by point of origin. These ton miles are based upon an assumed gross vehicle weight of 40 tons (80,000 lbs.) and a net weight of 33 tons. The percent of the trip accounted for on state highways versus county roads is shown in column 3. The fourth and fifth columns show the ton miles travelled loaded on state roads from each origin and the resultant damages at $0.071 per ton mile. Similarly, the next two columns show total ton miles travelled loaded on county roads and the resultant damages calculated at $0.1065 per ton mile. Total damages to state and county highways from loaded shipments were $144,039 and $17,943, respectively, or $161,982 in total. Since we assumed an average gross vehicle weight of 80,000 pounds, our damage estimate can be considered a fairly conservative lower bound, because trucks with legal gross vehicle weights of up to 105,000 pounds are often used in such shipments.

Table 6. Avoided Road Damages to State Highways and County Roads by Loaded Trucks.

	One Way Miles	Loaded Ton Miles	Percent State Road	Loaded Ton Miles, State Roads	Damage to State Roads	Loaded Ton Miles, County Roads	Damage to County Roads
Prescott	46	$561,502	100	$561,502	$39,867	$0	$0
Thornton	54	$659,155	100	$659,155	$46,800	$0	$0
Endicott	29	$460,188	69	$317,530	$22,545	$142,658	$15,193
Willada	47	$516,338	95	$490,521	$34,827	$25,817	$2,750
Total					**$144,039**		**$17,943**

Note: Total damage to state and county roads from loaded trucks: $144,039 (state) + $17,943 (County) = **$161,982**.

Return trips (backhauls), although usually empty, also cause damage to roads. Accordingly, backhaul damages were calculated based upon a 6.67 ton truck weight. Table 7 shows elevator, backhaul miles, and number of truckloads. That, when converted into ton miles empty on both state and county roads, results in damages to state highways and county roads. As noted in the table, damages to state highways from the return hauls were $23,782 and damages to county roads from backhauls were $2,963 annually.

In summary, total avoided damages to state highways by the 1995 grain train activity were $167,821; similarly, avoided damages to county roads was $20,906, resulting in total savings in 1995 of $188,727. For the optimal 1.5 carloadings per month total damages avoided would have been $232,888.

Valuation of Safety Benefits

Methodology

Two sets of safety valuations again are developed in this analysis. The first is the expected safety benefits attributable to the 1995 grain train program. The second is the annual safety benefits associated with the potential target operating year.

Data

State accident data reflect the configuration of truck as tractor semi-trailer or tractor double trailer. Separate data by vehicle weight (among semi-trailer combinations) or by type of commodity were not available from state sources, but federal data were available and were used to examine whether our estimates of trucking accidents were conservative or aggressive estimates.

Barge and rail safety data are far less sensitive over time and far less region-specific than are trucking data. Furthermore, property damaging accidents and fatalities per ton mile are so low in comparison to trucking accidents and fatalities as to be almost incalculable with statistical significance. Accordingly, we used federal data for both rail and barge; the source of these data was: U.S. Department of Transportation, *Transportation Safety Report.*

Table 7. Avoided Road Damages to State Highways and County Roads by Returning Trucks.

	One Way Miles	Empty Ton Miles	Percent State Road	Empty Ton Miles, State Roads	Damage to State Roads	Empty Ton Miles, County Roads	Damage to County Roads
Prescott	46	$92,712	100	$92,712	$6,583	$0	$0
Thornton	54	$108,836	100	$108,836	$7,727	$0	$0
Endicott	29	$75,983	69	$52,429	$3,722	$23,555	$2,509
Willada	47	$85,255	95	$80,992	$5,750	$4,263	$454
Total					**$23,782**		**$2,963**

Note: Total damage to state and county roads from returning trucks: $23,782 (state) + $2,963 (county) = **$26,745**.

Analysis

Table 8 shows mileage by truck from origin to barge facility on the way to Portland. Modal total mileages were in turn multiplied by accident rates to obtain total expected accidents by mode. Expected value of property, injury, and fatality accidents were estimated. Finally, expectations of alternative safety costs by truck-barge versus rail were calculated for (1) actual 1995 tonnage and (2) potential tonnage.

Table 8. Mileage and Tonnage Coefficients for Calculations of Safety Impacts, 1995.

Origin	Round Trip Miles	Truckloads	Truck Miles	Million Truck Miles
Prescott	92	302.17	27,800	0.0278
Thornton	108	302.17	32,634	0.0326
Endicott	58	392.82	22,784	0.0228
Willada	94	271.95	25,564	0.0256
Total			**108,782**	**0.1088**

Trucking Calculations

For the 0.1087812 million trucking miles displaced, there would be an annual accident expectation of 0.20668 accident per year at the Washington State non-interstate accident rate of 1.9 accidents per million miles. Thus, over a five year period, about one trucking accident can be expected from these displaced activities. Accordingly:

0.20668 accidents x 0.50 (percent of property damage accidents) x $5,300 = $548.

0.20668 accidents x 0.01368 (percent of accidents with fatalities) x $700,000 per fatality = $1,979.

0.20668 accidents x 0.4939 (percent of accidents with injury) x $57,000 per injury = $5,819.

Expected annual value of trucking accidents: $8,346.

Railroad Calculations

For the 0.00145 million miles of grain train travel (displacing the trucked segment) one can expect:

0.00145 x 0.0140 fatalities per million miles x $700,000 per fatality = $14.
0.00145 x 0.8336 injuries per million miles x $57,000 per injury = $69.
Property damage was so small it was not estimated.
Expected annual value of rail accidents: $83.

Therefore, the net expectations of accidents for the segment where trucking is displaced by rail is $8,346 - $83 = $8,263. As rail and barge safety performance is approximately equal, there is no need to calculate the displacement of the barged segment by rail.

Evaluation of Transportation Cost Savings to Growers

Transportation cost savings to growers is shown in Table 9. Each of the grower organizations was surveyed to determine the cost per bushel of shipping by rail (column 3) and by truck-barge (column 5). Column 7 shows savings to each of the shippers, which when summed, equals total savings of $92,320 for the 1995 year. This averages $0.066/bu. If the target 1.5 carloadings per month were achieved, the savings to shippers would total $115,158.

Energy Savings

The range of energy savings can be directly estimated from ton miles transported via each mode and the comparative modal energy efficiencies. Data used were a mixture from local and national sources. The major source of fuel saving stems from the replacement of trucking by rail, due to the magnitude of comparative energy consumption. Energy savings stemming from replacement of barge transport by rail, although positive, are considerably less than savings over trucking.

Fuel displaced by the grain trains for the actual 1995 operations was the sum of the truck-rail and barge-rail displacements, or 10,953 gallons. The value of these savings at $1.29/gallon is $14,129. Projecting these societal savings to the 1.5 potential target, the value would be 13,516 gallons and $17,436.

Table 9. Mileage, Tonnage, and Rate Coefficients for the Calculation of Shipper Savings, 1995.

Origin	Bushels	Rail Rate/Bushel	Total Cost by Rail	Truck-Barge Rate/Bushel	Total Cost by Truck-Barge	Savings
		(¢)	($)	(¢)	($)	($)
Prescott	335710.9	0.285	95,678	0.310	104,070	8,392
Thornton	335710.9	0.298	100,042	0.375	125,892	25,850
Endicott	436424.2	0.270	117,835	0.320	139,656	21,821
Willada	302139.8	0.270	81,578	0.390	117,835	36,257
Total						**92,320**

Summary of Quantifiable Impacts and Potential Impacts

The findings of this section are summarized in Table 10. The first column shows a total of $303,439 as valuation of positive impacts of the actual grain train use in 1995. Column 2 shows potential positive impacts of $375,678 if the cars were to be loaded at the potential 1.5 times per month. Savings per grain car for 1995 of $720 per carload is calculated in column 3.

Were the fleet to grow to 140 cars, as indicated in the finance section of this report, and were the cars utilized at 1.5 carloadings per month in approximately the same nature of usage, the net positive impact in the twenty-third year would be: $720 x 1.5 x 12 x 140 = $1,814,400 per year.

Table 10. Loadings and Savings, for Actual 1995 Activities, for Ideal Activities, and by the Grain Carload Shipped.

	1995	Ideal	1995 Savings by the Grain Carload
Loadings Per Car Month	1.3	1.5	
Loadings Per Year	423	522	
Safety Savings	$8,263	$10,196	$20
Fuel Savings	$14,129	$17,436	$33
Transportation Cost Savings to Growers	$92,320	$115,158	$221
County Road Damage Avoidance*	$20,906	$25,797	$49
State Highway Damage Avoidance*	$167,821	$207,091	$397
Total	**$303,439**	**$375,678**	**$720**

*Sensitivity analysis of a 50% reduction in road damage per ton mile indicates total road damage for 1995 and the ideal would be $94,363 and $116,444, respectively.

IMPLICATIONS

This analysis of efficiency, management and regional/state impacts of the grain train project reveals a project that has succeeded in

meeting most of its original program goals: rail car capacity has been increased in a period of continuing car shortage; rail service on light density rail lines has been sustained, maybe even saved, thereby generating benefits that reach far beyond the grain industry; and, the project has saved energy, increased safety, and returned benefits to a population segment overcharged for petroleum products.

The State of the Canadian Intercity Highway System, 1986-1993

Bob Leore[1]

Introduction

This paper reports on an attempt to improve the state of knowledge about intercity road transport in Canada. The motivation has grown out of dissatisfaction with conventional statistical sources which are inadequate for making in-depth analyses of the structure and trends in intercity road transport. This unpleasant reality has forced the adoption of a new data source, specifically, the planning information collected by provincial highways departments. This data source, hitherto unexploited, is remarkably rich in detail. Unfortunately, no one has undertaken the task of combining these separate sources into a regular, comparable set of *national* statistics. This was partially rectified with the recent creation of a Special Infrastructure Project by Transport Canada.[2]

This paper, an extension of the previous work, will examine the state of our intercity highway network by presenting facts relating to the level and composition of vehicle traffic as well as to the current condition of the highway infrastructure. In particular, two important qualities of the highway facilities will be evaluated: 1) the level of capacity utilisation of the system and 2) the ride quality of network. It will also be shown that the way in which the data are represented spatially provides not merely statistics but powerful tools in the development of future policy on highways.

---

[1] Transport Canada. The views expressed in this paper are those of the author and do not necessarily reflect the views of Transport Canada. I should like to thank the provincial and territorial highways departments whose cooperation made this project possible.

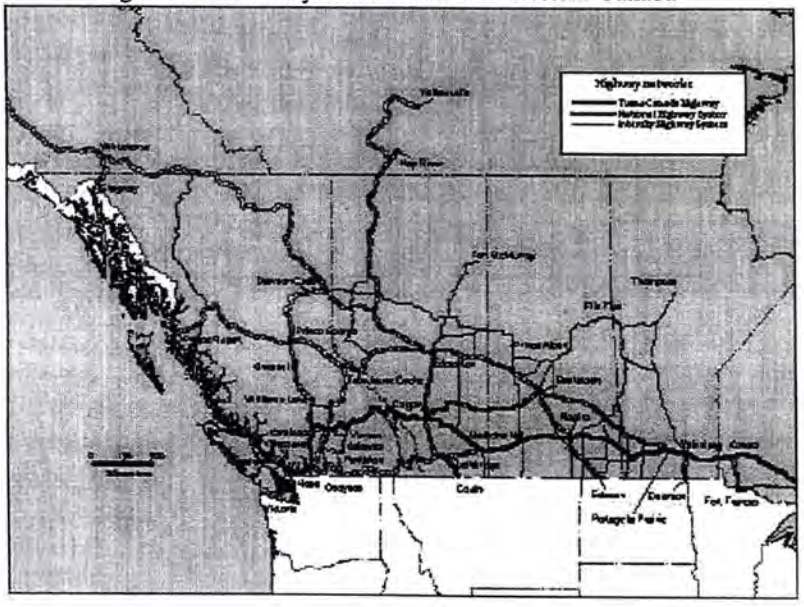
[2] The Special Infrastructure Project (SIP) was created in 1995 to study various highway-related issues to inform future federal policy in this area. A series of analytical reports, summarising this research, was published in the summer of 1996 under the direction of David Stambrook. Some road data discussed in this paper have also appeared in the author's SIP report, "An Economic Model of Inter-Urban Travel on the Canadian Highway Network," TP-12800E, June 1996.

Scope of the project

Owing to the active cooperation of the provincial and territorial highway authorities, the following road data were supplied to Transport Canada: 1) vehicle traffic counts represented by the average annual daily traffic (AADT) passing over the road network; 2) the percentage of the traffic stream represented by commercial vehicles (effectively heavy trucks); 3) the capacity of the network represented by the number of lanes; 4) the quality of the pavement for each road segment in the provincial network as represented by the ride roughness.

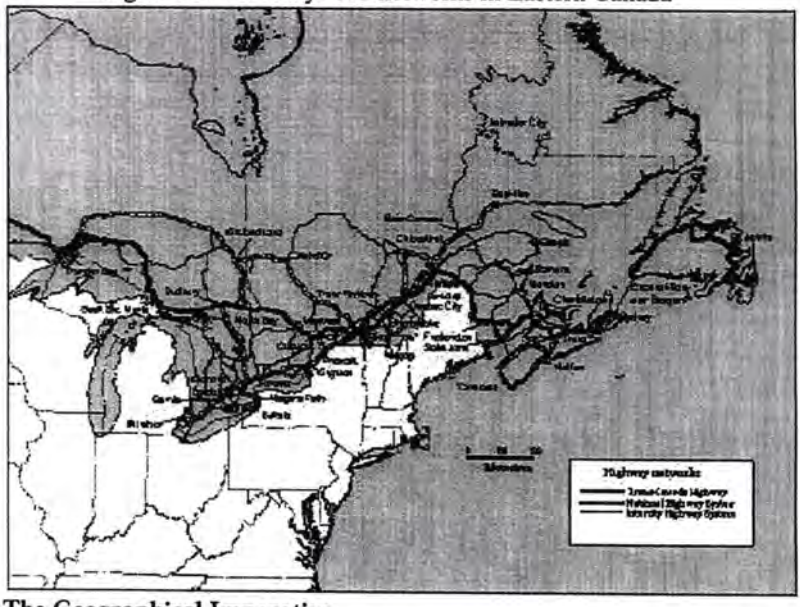
These data were very extensive in scope covering literally all provincial numbered highways. Although statistics could have been assembled for the entire system, time constraints forced the selection of a smaller, more manageable, yet no less significant, sub-network. After very careful consideration, a base network was chosen for analysis joining all major urban areas in Canada. This network, which includes the busiest highways and those most important to Canada's trade and tourism, is termed the Intercity Highway System (IHS).

Figure 1A: Intercity road networks in Western Canada



As can be seen in Figures 1A and 1B, the IHS is a very dense network comprising approximately 45,000 km of numbered highways. Contained within the IHS are two other significant networks which will also be examined here: 1) the National Highway System (NHS) of approximately 24,000 km joining provincial capitals, and 2) the 7,300 km Trans-Canada Highway (TCH), the nation's main street, linking Vancouver Island with Newfoundland.

Figure 1B: Intercity road networks in Eastern Canada



The Geographical Imperative

A notable innovation introduced for this project was the use of a Geographic Information System (GIS) to organise and analyse the data. The compilation of road transport statistics at the level of detail contemplated here would not have been practical without a GIS. The power of the GIS approach is its unique blending of traditional database tools with the visual representation of spatial objects. For any transport application this is an absolute must since knowing *where* the activity is located is just as important as knowing the magnitude of such activity. A GIS is ideal for answering both types of questions; sadly, most road authorities have yet to introduce such a system. In fact, only Ontario was

able to send geo-referenced road data. This meant that the rest of the provinces had to have their statistics geo-referenced manually.

The solution to the lack of spatially referenced data was to obtain a digital map of the North American highway network and match the provincial data to the objects on this map. The map used in this project was obtained from the company American Digital Cartography (ADC). Its Worldmap product is a variant of the well-known Digital Chart of the World (DCW), formatted specially for use in the popular cartographic program MapInfo. Although it has its flaws, Worldmap s finely segmented depiction of the highway network is a suitable platform for capturing the rich provincial databases.[3] In almost all cases, road segment distances on the digital map matched closely with corresponding distances on official provincial highway maps, thus allowing a realistic depiction of the structure and trends in road travel.

Traffic levels on the Intercity Highway System (IHS)

The most important variable supplied by the provincial authorities was the vehicle traffic counts by road segment. Traffic counting has advanced enormously in recent years with the introduction of sophisticated machines capable of counting and classifying vehicles continuously. Although this new technology has done much to improve the accuracy and reliability of counting, continuous counting has yet to become a standard practice throughout the network. Most of the counting is still short-term in nature and vehicle classification is largely a manual process conducted in daylight hours. In general, however, there is sufficient confidence that counts are accurate and classifications reliable enough that at least basic inferences about the characteristics of the traffic stream can be made.

For the IHS, count data were attached to over 4,700 separate line segments for the years 1986 and 1993. Once assigned, it was possible to tabulate aggregate traffic levels nationally and for each jurisdiction by computing an estimate of vehicle-kilometres. Vehicle-km, which are the product of a traffic count and its associated segment length, are the basic summary statistic for road traffic in our system. Table 1 summarises these levels by province and territory.

---

[3] For example, the multi-lane Coquihalla highway in B.C. did not appear on the map and had to be drawn in manually from a paper map.

Table 1: Traffic levels on the Intercity Highway System, 1986 and 1993

Province/ territory	1993 route-kilometres ('000s)	Vehicle-kilometres (billions)		Average annual growth in veh-km	Average annual daily traffic (AADT)		Percentage distribution		
							Route-km	Vehicle-km	
		1993	1986		1996	1986		1993	1986
Nfld.	1.1	0.8	0.7	1.5%	2,000	1,800	2.4	0.8	0.9
P.E.I.	0.2	0.3	0.3	3.2%	3,800	3,100	0.5	0.3	0.3
N.S.	1.4	2.9	2.4	2.9%	5,700	4,700	3.1	2.8	2.9
N.B.	1.8	3.1	2.7	1.9%	4,800	4,200	3.9	3.0	3.3
Que.	7.6	25.5	19.4	3.9%	9,200	7,000	16.9	24.9	23.9
Ont.	9.5	38.4	30.1	3.5%	11,100	8,700	21.0	37.5	37.1
Man.	3.0	2.3	2.0	1.6%	2,100	1,900	6.6	2.2	2.5
Sask.	5.1	4.0	3.4	2.2%	2,200	1,900	11.2	3.9	4.2
Alta.	6.2	9.7	8.0	2.7%	4,300	3,600	13.7	9.5	9.9
B.C.	7.8	15.2	11.8	3.6%	5,300	4,100	17.4	14.8	14.6
Yuk.	0.9	0.2	0.2	1.4%	600	500	2.0	0.2	0.2
N.W.T.	0.5	0.04	0.04	0.1%	200	200	1.2	0.03	0.04
Total	45.0	102.3	81.1	3.4%	6,200	4,900	100.0	100.0	100.0

In 1993, a little over 100 billion vehicle-km were generated on the IHS, an average of over 6,000 vehicles per day. Vehicle traffic was especially concentrated in Ontario and Quebec which together accounted for over 60% of the total even though their road networks accounted for less than 40% of the IHS's route-km. Their daily averages were roughly twice as high as any other province. All the other regions of Canada had traffic shares less than their shares of route-km, a reflection of the dispersed settlement patterns throughout the West and Atlantic Canada.

The growth in vehicle traffic was also noteworthy. Between 1986 and 1993, traffic grew at a brisk average annual rate of 3.4%. The pattern of growth was especially intriguing. The busiest provinces were also the fastest growing with Quebec and Ontario traffic both growing faster than the national average, Quebec at nearly 4% per annum. British Columbia was also relatively rapid at 3.6% per annum. In contrast, the slowest growing provinces were Newfoundland, New Brunswick, and Manitoba all growing by less than 2% per annum.

As all traffic data are geo-referenced, we can also examine the spatial pattern of traffic volumes on the broadly defined IHS. Figures 2A and 2B show the network broken down by selected traffic ranges. The maps reveal immediately that private vehicle travel is a local phenomenon clustering around the largest urban areas in Canada. The traffic in and around the three largest cities, Toronto, Montreal, and Vancouver was especially heavy with 1993 average annual daily traffic exceeding 100,000 vehicles per day over many highway sections. Other than the concentrations in the Big-3 and in other large cities such as Quebec City, Ottawa, or Winnipeg, traffic over the bulk of the network was extremely sparse, usually averaging less than 5,000 AADT.

Figure 2A: Traffic volumes on the IHS, Western Canada, 1993

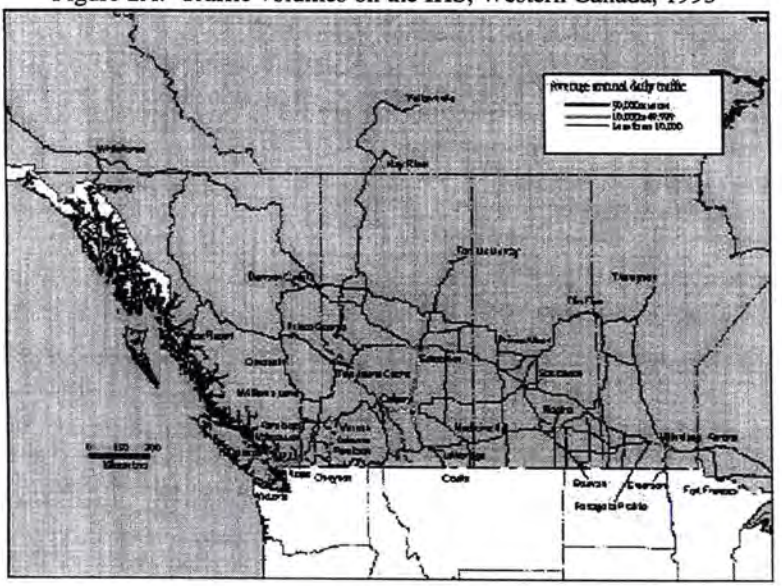
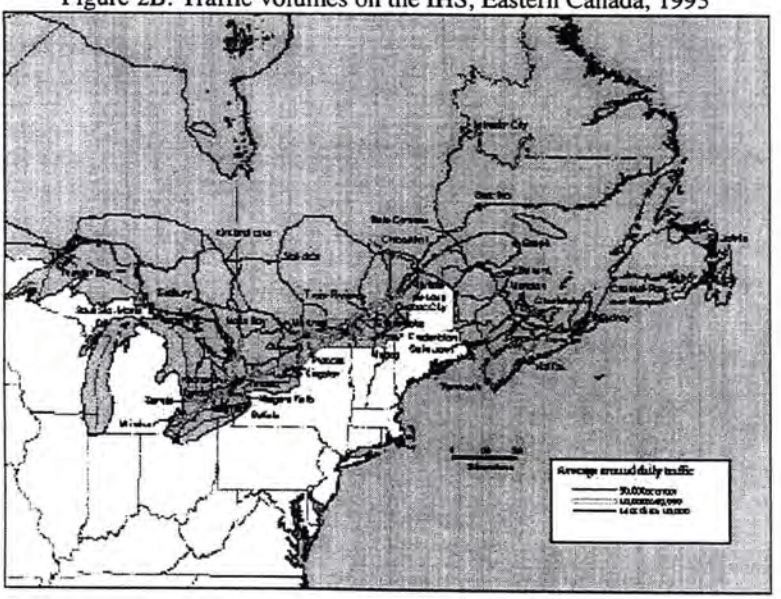


Figure 2B: Traffic volumes on the IHS, Eastern Canada, 1993



This wide disparity in traffic levels is seen more clearly in Table 2 where the data are grouped into AADT bands. Roads with less than 2,500 AADT accounted for 50% of the network in 1993 but only 9% of the vehicle-km. This is in stark contrast to the busiest sections above 50,000 AADT, located largely inside the urban boundaries of the Big-3. These roads made up less than 2% of the route-km but accounted for almost 30% of the traffic. One can also see in Table 2 that the busiest segments of the distribution have begun to hit some capacity constraints. Road sections with AADT over 100,000 actually had below-average growth, no doubt a reflection of growing congestion.

Table 2: Distribution of IHS traffic by traffic range, 1986 and 1993

AADT range	1993 route-kilometres ('000s)	Vehicle-kilometres (billions)		Average annual growth in veh-km	Average annual daily traffic (AADT)		Percentage distribution		
		1993	1986		1996	1986	Route-km	Vehicle-km 1993	Vehicle-km 1986
0-2,500	22.2	9.2	8.2	1.7%	1,100	1,000	49.2	9.0	10.1
2,501-5,000	11.0	14.4	12.3	2.4%	3,600	3,100	24.4	14.1	15.1
5,001-10,000	6.3	15.8	12.6	3.3%	6,900	5,500	13.9	15.5	15.5
10,001-50,000	4.9	34.9	25.9	4.3%	19,600	14,600	10.8	34.1	32.0
50,001-100,000	0.4	9.7	7.2	4.3%	68,600	51,100	0.9	9.5	8.9
Over 100,000	0.3	18.2	15.0	2.9%	146,400	120,300	0.8	17.8	18.4
Total	45.0	102.3	81.1	3.4%	6,200	4,900	100.0	100.0	100.0

Traffic levels on the other sub-networks

By comparing Figures 1 and 2, we can see that all the important concentrations of traffic, with the exception of the Okanagan region in B.C., lie along the National Highway System. As Table 3 shows, over 70% of the IHS traffic in 1993 was carried by NHS highways; moreover, this sub-network was much busier with average volumes 35% higher than the IHS. Traffic was even more concentrated in Ontario and Quebec where the two provinces possessed 32% of the NHS network but 62% of the vehicle traffic. Their overall AADTs where again two to three times larger than the other provinces.

The Trans-Canada Highway, on the other hand, is a curious mix of extremes. It represents a shortest continuous path from coast to coast and thus bypasses most of the busiest parts of Canada, including Southern Ontario, the busiest of all. As a result, Ontario accounts for only 18% of the system s traffic while it possesses 30% of the kilometrage. The TCH does manage, though, to pass through the heart of Quebec following Autoroute 20, one of the busiest highways in the country. Not surprisingly, Quebec dominates TCH traffic accounting for 30% of the total on only 8% of the route-km. Its average AADT for 1993, a gargantuan 30,000, is five to six times higher than all provinces

except B.C. Since low-density rural highways are over-represented in this network, traffic growth since 1986 has been less strong compared with the larger networks, with the exception of Quebec which managed a compound growth rate of 3.8%.

Table 3: Traffic levels on important sub-networks, 1986 and 1993

Province/ territory	1993 route-kilometres ('000s)	Vehicle-kilometres (billions) 1993	1986	Average annual growth in veh-km	Average annual daily traffic (AADT) 1996	1986	Percentage distribution Route-km	Vehicle-km 1993	1986
National Highway System									
Nfld.	0.9	0.8	0.7	1.4%	2,400	2,100	3.6	1.0	1.2
P.E.I.	0.1	0.2	0.1	3.2%	4,200	3,300	0.5	0.2	0.2
N.S.	0.9	2.2	1.8	3.1%	7,100	5,700	3.6	3.0	3.1
N.B.	0.9	2.2	2.0	1.4%	6,400	5,800	4.0	3.0	3.4
Que.	2.8	18.5	13.9	4.1%	18,000	13,600	11.7	25.2	23.8
Ont.	5.0	27.1	21.5	3.3%	14,800	11,800	20.9	36.9	36.8
Man.	0.9	1.3	1.2	1.8%	4,300	3,800	3.6	1.8	2.0
Sask.	2.1	2.8	2.4	2.4%	3,700	3,100	8.8	3.8	4.1
Alta.	3.5	7.6	6.3	2.8%	5,900	4,900	14.7	10.4	10.7
B.C.	5.3	10.5	8.3	3.3%	5,400	4,300	22.3	14.2	14.3
Yuk.	1.0	0.2	0.2	1.5%	600	500	4.1	0.3	0.3
N.W.T.	0.6	0.04	0.04	0.3%	200	200	2.3	0.1	0.1
Total	24.0	73.5	58.5	3.3%	8,400	6,700	100.0	100.0	100.0
Trans-Canada Highway									
Nfld.	0.9	0.8	0.7	1.4%	2,400	2,100	11.9	3.3	3.6
P.E.I.	0.1	0.1	0.1	2.6%	3,700	3,100	1.4	0.6	0.6
N.S.	0.4	1.0	0.8	2.8%	6,600	5,500	5.7	4.4	4.4
N.B.	0.5	1.2	1.2	0.3%	6,100	6,000	7.5	5.3	6.4
Que.	0.6	6.7	5.2	3.8%	30,800	23,700	8.2	29.2	27.5
Ont.	2.1	4.2	3.4	2.8%	5,400	4,400	29.0	18.1	18.2
Man.	0.5	1.0	0.9	2.0%	5,600	4,900	6.7	4.4	4.7
Sask.	0.7	1.1	0.9	2.2%	4,400	3,800	8.9	4.6	4.8
Alta.	0.5	1.5	1.2	3.4%	7,700	6,100	7.3	6.5	6.3
B.C.	1.0	5.5	4.4	3.2%	15,400	12,400	13.3	23.7	23.3
Total	7.3	23.1	18.8	2.9%	8,600	7,000	100.0	100.0	100.0

Composition of the traffic stream

With the aid of the vehicle classifications also supplied, the pattern of car and truck travel on the various networks can be investigated. In Table 4, the estimates of truck vehicle-km are provided by province and territory for 1993.[4] Clearly, the NHS is the primary network for truck activity handling about 75% of the truck traffic moving over the IHS. On balance, about 15% of the traffic stream in Canada, regardless of network, is occupied by heavy trucks and other commercial vehicles.

---

[4] As the vehicle classifications were normally provided in the form of the percentage of commercial vehicles in the traffic stream, we show here the truck amounts. Car travel can be calculated as the difference between total vehicle-km and truck vehicle-km.

Ontario and Quebec are responsible for a larger share of total truck activity on the IHS and the NHS, about 65%, which is owing largely to the extensive industrial facilities located in the region, all heavy users of trucking services. Care should be exercised in using the B.C. figures which seem unreasonably low. Although similar definitions of commercial vehicles were sought, it seems likely that this discrepancy is more a measurement problem than a real phenomenon.

Table 4: Truck volumes by road network, 1993

Province/ territory	Truck vehicle-kilometres (bill.)			Share of trucks in traffic stream			Average annual daily truck traffic		
	IHS	NHS	TCH	IHS	NHS	TCH	IHS	NHS	TCH
Nfld.	0.1	0.1	0.1	17%	16%	16%	300	300	300
P.E.I.	0.0	0.0	0.0	10%	10%	11%	300	400	400
N.S.	0.3	0.3	0.2	13%	14%	19%	700	900	1,000
N.B.	0.5	0.4	0.3	19%	21%	28%	800	1,100	1,300
Que.	3.0	2.1	0.9	13%	13%	15%	1,100	2,000	4,100
Ont.	5.4	4.3	0.6	17%	19%	18%	1,600	2,300	800
Man.	0.3	0.2	0.1	14%	17%	17%	300	600	800
Sask.	0.6	0.4	0.2	17%	18%	18%	300	500	700
Alta.	1.3	1.0	0.2	16%	16%	15%	600	800	1,000
B.C.	1.3	0.9	0.4	9%	9%	8%	400	500	1,100
Yuk.	0.0	0.0	-	4%	5%	-	20	30	-
N.W.T.	0.0	0.0	-	61%	57%	-	100	100	-
Total	12.9	9.7	2.9	14%	15%	15%	800	1,100	1,100

Capacity utilisation

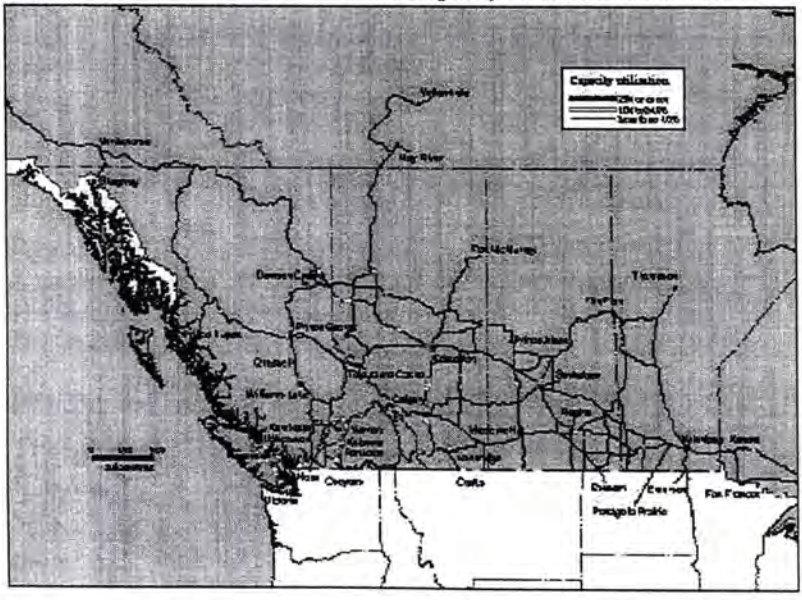
One of the benefits of using a GIS to collect road data is the ability to link traffic explicitly to infrastructure. This allows us to observe the pattern of capacity utilisation, an important qualitative dimension. Traffic engineers have developed convenient rules of thumb to assess highway capacity. For a multi-lane, controlled-access highway, such as the 401, capacity is normally 2,000 vehicles per lane per hour. For a two-lane rural highway without access control, such as Highway 17 in Northern Ontario, capacity is normally a little less at 1,400 vehicles per lane per hour.[5] Employing these facts, we can compute the capacity of our road system and determine its rate of utilisation. Figures 3A and 3B show these rates for the system as a whole, while Table 5 shows overall rates by province and traffic range. What is striking is the overall low

[5] Note that these general rules abstract from other physical characteristics (e.g. steep grades) which may also affect the capacity of a given facility.

level of utilisation for the vast majority of the network, particularly at the lower traffic ranges. Highways with traffic levels under 2,500 AADT use a little over 2% of available capacity.

As one moves up the traffic scale, the rate of capacity utilisation marches in step and accelerates once the traffic level gets above 50,000 AADT. At these ranges, one is again inside the major urban areas where daily commuting puts a tremendous pressure on road facilities. In many parts of the Big-3, daily traffic fills more than half the available capacity. This is remarkable given that commuting patterns cluster around rush hours and are not spread evenly throughout the day. This means that many urban highways are effectively busy for all the daylight hours. For example, the highest utilisation rate was observed along sections of the 401 through Toronto where traffic used up nearly 70% of road capacity. When we find that these road sections are twelve lane highways, we understand instantly the magnitude of the congestion problems faced daily in our big cities.

Figure 3A: Traffic as a share of road capacity, Western Canada, 1993

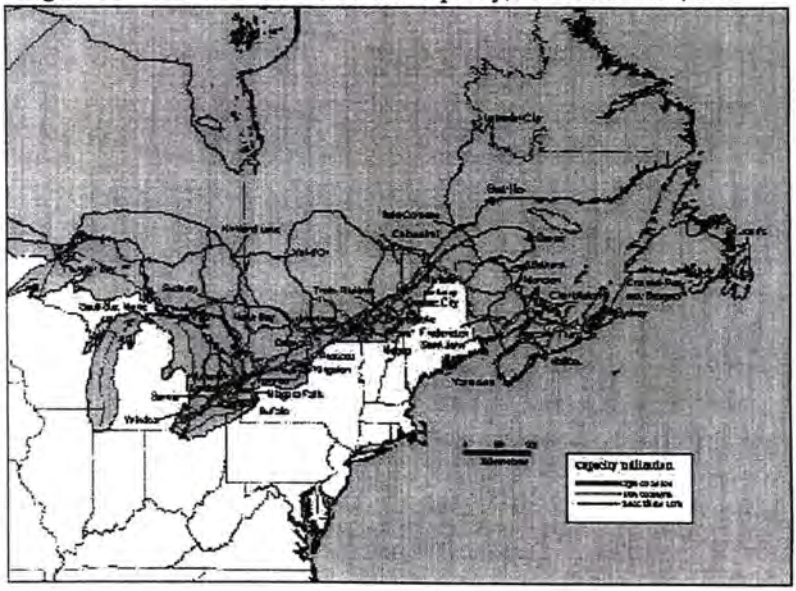


Ensuring adequate road capacity is, thus, largely an urban problem. For most rural highways there is more than adequate existing lane capacity to handle future traffic growth for many years to come.

Table 5: IHS capacity utilisation rates by province and traffic range, 1993

	AADT range						
	2,500 or less	2,501-5,000	5,001-10,000	10,001-50,000	50,001-100,000	100,000 or more	Total
Nfld.	2.7%	5.2%	5.2%	14.3%	-	-	
P.E.I.	3.8%		9.5%	20.7%		-	8.7%
	2.8%	5.2%		10.0%	-		7.4%
N.B.		6.1%	8.2%		-	-	
Que.	2.6%		9.2%	14.5%		50.0%	14.7%
	2.7%	6.3%		13.4%	31.1%		20.0%
Man.		3.5%	4.9%		-	-	
Sask.	2.1%		4.8%	14.1%		-	3.3%
	2.4%	4.0%		10.8%	-		5.3%
B.C.		5.4%	9.4%		33.9%	52.4%	
Yuk.	1.2%		-	-		-	1.2%
	0.4%	-		-	-		0.4%
Total		4.9%	8.2%		31.9%	48.6%	

Figure 3B: Traffic as a share of road capacity, Eastern Canada, 1993

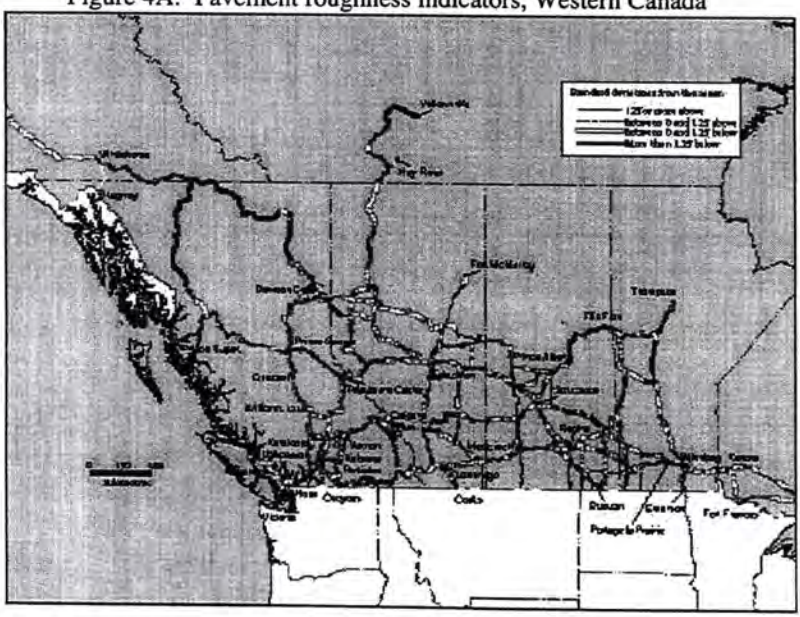


Pavement quality

The remaining part of the data request dealt another qualitative indicator of the road infrastructure pavement roughness. The smoothness of the surface over which vehicles drive is a valuable added piece of information which has so far received no national treatment. A variety of pavement quality rating systems are in use, but most jurisdictions employ the Riding Comfort Index (RCI), which measures roughness on a scale of 0 to 10 with values under 6 considered to be of poor ride quality.

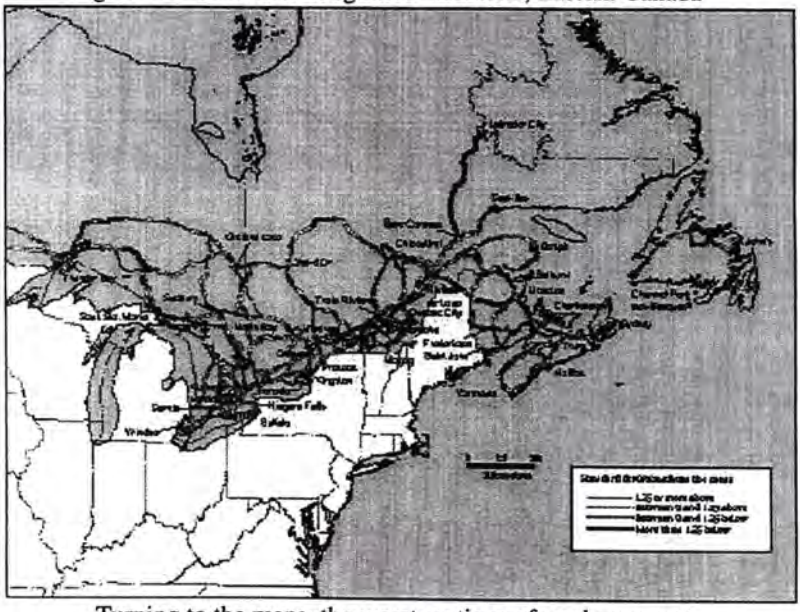
Although the RCI or close proxy was requested, some provinces were unable to supply data in this form. Manitoba, Ontario, and the N.W.T., for example, use a Pavement Condition Rating (PCR) with a scale running from 0 to 100. Since the PCR and RCI are not readily translated, all roughness data were converted into a standardised normal variable based on the mean and standard deviation of each jurisdiction s distribution of ride quality ratings. These values, expressed in units of standard deviation away from the provincial mean are mapped in Figures 4A and 4B.

Figure 4A: Pavement roughness indicators, Western Canada



Upon close inspection, it was found that 1.25 standard deviations below the mean was a convenient cut-off for determining an abnormally rough road. This threshold corresponded relatively well to an RCI of 5.5 to 6 in most provinces. As can be gleaned from Table 6, on the whole, the ride quality of the provincial networks was solidly average. Approximately 5,000 km or 12% of the intercity network had quality measures more than 1.25 standard deviations below the average, while about 4,000 km or 9% of the network had values in excess of 1.25 SDs above the mean. The rest of the road system, 35,000 km, had values fall somewhat closer around the averages which generally clustered in the mid-6 range on the RCI scale. New Brunswick, B.C., and Quebec appeared to have the best quality networks since about 60% of their route-km had RCI values above the mean. The poorest quality networks appeared to be Nova Scotia and Ontario which had only 44% and 47% of their route-km above average. In aggregate terms, the distribution was skewed to the right, with 54% of the IHS route-km above the mean values.

Figure 4B: Pavement roughness indicators, Eastern Canada



Turning to the maps, the worst sections of road were, not surprisingly, located in the more remote parts of the country such as

northern B.C., the two Territories, or northern Quebec, where the roads in these areas are either gravel or have very thin pavements. Farther south, the poor quality of significant portions of the NHS highways such as Highway 16 in Saskatchewan between Yorkton and Saskatoon, the Trans-Canada highway through Northern Ontario, and Highway 4 on Cape Breton Island in Nova Scotia was noted. In addition, much of the QEW and 400-series highways in Southern Ontario were in below-average condition.

Table 6: Summary recent pavement condition statistics by province

							IHS route-km less ferries ('000s)				
	Mean values			Standard deviation			More than 1.25 SD below	Within 1.25 SD below	Within 1.25 SD above	More than 1.25 SD above	Total route-km
	IHS	NHS	TCH	IHS	NHS	TCH					
Riding comfort index (RCI) or equivalent											
Nfld.	6.7	6.8	6.8	0.8	0.7	0.7	0.1	0.3	0.4	0.1	0.9
P.E.I.	6.7	6.7	6.7	0.7	0.8	0.7	0.0	0.0	0.1	0.0	0.2
N.S.	5.6	5.6	5.4	1.0	1.0	0.8	0.2	0.6	0.4	0.2	1.3
N.B.	6.3	7.0	7.0	1.3	0.5	0.5	0.3	0.3	1.2	0.0	1.8
Que.	6.7	6.9	7.0	0.8	0.7	0.7	1.0	1.9	4.0	0.5	7.5
Sask.	6.3	6.7	6.3	1.3	0.8	0.7	0.6	2.0	2.0	0.4	5.1
Alta.	6.5	6.3	6.4	0.7	0.7	0.4	0.4	2.7	2.5	0.6	6.2
B.C.	7.1	7.2	7.5	0.9	0.8	0.5	1.4	1.4	4.5	0.3	7.7
Yuk.	5.4	5.2	n.a.	0.4	1.1	n.a.	0.0	0.5	0.3	0.1	0.9
Pavement condition rating or index											
Ont.	70.0	69.9	67.0	14.8	15.7	16.5	1.0	4.0	3.0	1.5	9.5
Man.	74.4	78.7	80.0	11.3	8.6	7.7	0.3	1.1	1.3	0.2	3.0
N.W.T.	57.5	57.3	n.a.	19.5	18.7	n.a.	0.0	0.0	0.0	0.0	0.0
Total							5.4	14.9	19.7	4.0	43.9

Some of the best stretches in Western Canada were located in B.C., especially Highways 1 and 19 on Vancouver Island, the Coquihalla Highway, and the Trans-Canada highway between Kamloops and the Alberta Border. In Eastern Canada, the 401 near London, Highway 400 north of Toronto, Highway 40 near Trois-RiviŁres, the Trans-Canada highway in New Brunswick, and Highway 104 in Nova Scotia between Amherst and Wentworth Centre were in above-average condition.

Summary and recommendations for future action

Based on the information developed for this project, it is safe to say that Canada has erected a road system that, on balance, functions adequately. The population is served with a dense network of highways that permit virtually unfettered interaction; the majority of the roads are

of sufficient capacity to meet the demands placed on them for many years to come; and, the pavement surfaces over which much of the traffic moves is acceptably smooth. In many respects, this network is a great success story; yet, in other areas we face serious and, increasingly, pressing challenges that threaten its integrity. For one, ensuring an adequate level of maintenance will always be difficult task given the extent of the system and our harsh climate. There can really be no real relaxation on this front either since work delayed is more costly work later on. The most critical challenge of all, though, is how we collectively meet the growing demands for more and more road capacity in our major cities. Much of the urban highway facilities are literally bursting at the seams, clogged with traffic all day long, imposing substantial delays not only on commuters but truckers carrying goods to their markets. The policy question we will need to face is whether we can continue to stay on top of this inexorable growth by simply adding more lanes. I should say that the odds are stacked heavily against the planners. Even at 2% annual growth, traffic on the busiest parts of the 401 will be at 500,000 AADT by 2011. Will we also see 18, 20 or even 24 lane freeways in 20 years? Do we really want to? Clearly, new and imaginative solutions will have to be implemented or else the system will quite simply implode under its own weight.

In the short-term, it will be crucial, nay imperative, that we develop a much more elaborate and robust set of traffic statistics and other performance indicators. The first concrete step down the road to better statistical systems would be for all jurisdictions to advance with due diligence to geographically-referenced road data. The spatial dimension, so important to this project, must become the number one priority for all parties concerned. This advance should then be followed quickly with a serious attempt to make the collection of traffic counts, especially classified traffic counts, a more systematic and continuous process throughout the entire network. It is true that road authorities have been under enormous pressure lately to scale back their counting activities. This is an egregious and counter-productive move that risks undermining all the progress achieved in recent years. It must become policy at the highest-levels of government to preserve and enhance the collection of road information. Only in this way will we finally have a base of knowledge commensurate with the road mode's importance, and the confidence to make informed decisions about the future direction of road transport in Canada.

The Welfare Gain from Trucking Deregulation: Transborder Trucking and Cabotage

Darren Prokop

Ph.D Student
Department of Economics
The University of Manitoba
&
Research Associate
The Transport Institute
The University of Manitoba

Introduction

This paper represents a portion of my Ph.D thesis dealing with how to model the effect of cabotage in transborder trucking markets. It will be argued that: (1) deregulation in general represents a net welfare gain to a society as narrowly composed here of shippers and trucking firms; and (2) the current cabotage regulations, in place in Canada and the United States, are a non-tariff trade barrier that, even in the current deregulatory environment on both sides of the border, can lead to a further gain in societal welfare if they were removed either unilaterally or under a reciprocity agreement. In other words, cabotage reform is a positive sum game for North American transborder trucking markets.

Regulation vs. Deregulation in Trucking Markets

A stylized view of how the trucking industry responds to cyclical demand shocks under regulatory and deregulatory regimes has been given in **Prentice (1994)**. That analysis proposed industry supply curves which were

specially "kinked" to highlight the fact that the industry chose to either rate-adjust or quantity-adjust to an expansionary or contractionary demand shock as dependent upon which of the two regimes were in place. It is through comparisons of these temporary equilibria that estimates of the welfare gain from deregulation can be had. What is interesting is that the welfare gain exists independently of whether the industry is facing demand expansion or contraction.

Consider how the industry responds to a temporary demand shock while under a rate/entry regulatory regime. The long run/short run distinction of supply used in **Prentice (1994)** for the building of the relevant supply curves is not in the traditional microeconomic sense because time periods for adjustment, and presence of fixed factors of production, are not really the issue here; rather it is the firms' response to regulations that are important. The point is that the price elasticities with respect to industry supply differ depending on the context; that is, firms choose for example to quantity-adjust during a recession (i.e. a cyclical contraction of demand) when in a regulatory environment but, in a deregulatory environment, choose to price (or rate)-adjust instead.

Consider the regulatory environment and its effect on industry supply. During an expansion regulatory barriers to entry (i.e. operating licenses, geographical operation limits, weight and dimension limits) mean that industry size is fixed and rates expand faster than do the quantity of services supplied due to the ability of incumbent firms to co-ordinate prices in an oligopolistic sense due to regulatory protection.[1] During a contraction, the regulatory regime prevents competitive rate cutting which then allows firms to scale back their operations by letting a part of their fleets remain idle and by laying off drivers. This quantity-adjustment process on the firms' part allows them to ride out recessions and maintain limited competition by implicit cooperation serving to reduce capacity and keep prices higher than otherwise.

In the deregulatory environment the situations are reversed from those above. During an expansion there is a greater incentive for entry which serves to dampen the rate-adjusting process found in the regulatory environment. Not only does entry occur, the incumbent firms would compete by running their own fleets longer hours and perhaps limiting truck turnover

(which makes the case that fleets become overworked and less safe leading to an argument for, at least, some safety regulations to remain in place). To the extent that the industry relies to a large degree on owner-operators in a deregulatory environment, the expansion of fleets involves very short time lags. During a contraction, competitive rate cutting will indeed occur since owner-operators face the fixed costs of their vehicle forcing them to stay in business and "loss-minimize"; that is, while a firm may afford to idle one or a few trucks (as in the regulatory case), an owner-operator cannot do likewise so easily, owing to his having to personally bear the fixed cost of his vehicle.

Rates would be lower given the deregulated supply than they would be given the regulated supply which will turn out to be the theoretical source for the welfare gain found in a drive to deregulation in the trucking sector. In this regard, an industry demand curve can be joined with the industry supply curve.[2]

As can be seen in figure 1 below, a relative welfare gain occurs when supply characterized by regulation is discarded in favor of that allowing for deregulation, irrespective of whether demand is contracting or expanding. The welfare gain associated with a move to deregulate is given by the typical Harberger welfare triangle BCD. The deregulatory result gives a greater quantity at a lower rate (C as compared to B). Point A represents an initial long run equilibrium in the sense that the equilibrium rate and quantity are not affected by a change in regime because all firms are, by assumption, fully adjusted to the market. In the expansion and contraction cases, the demand curve shifts appropriately to point B which represents the initial regulatory equilibrium from which the welfare gain from deregulation is examined.[3]

Figure 1a: Demand Expansion

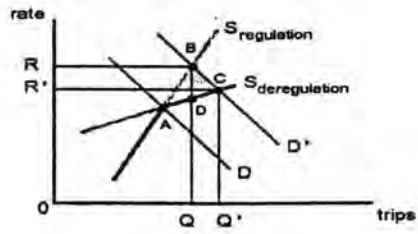
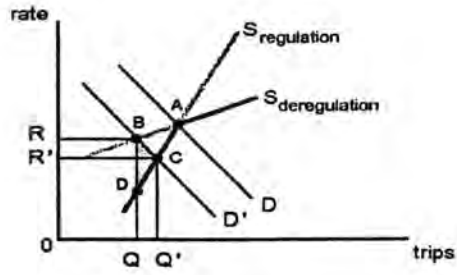


Figure 1b: Demand Contraction



The welfare gain of BCD represents the sum of the expanded consumer and producer surpluses obtained in achieving C instead of B. The breakdown of the welfare triangles is as follows. Point D represents the minimum cost of provision of quantity Q while point B represents the industry rate set under regulation. Thus, the line segment BD represent the incremental rent earned on the marginal output unit Q. At point C there are no rents since the number of firms in the industry has expanded to a competitive equilibrium. The incumbent firms do not exit the industry if they have regulatory protection and owner-operators do not exit so long as they can loss

minimize. At point C, there are more firms in the industry as compared to point B (likely, owner operators trying to thrive in the deregulatory environment) and the competitive rate-cutting expected during a contraction allows for no rents to obtain. Also it should be noted in the contraction case that point C is not a long run equilibrium; rather, it is a short run equilibrium. Since the incumbent firms are simply riding-out the recession, they are waiting for an expansion of demand to move the industry back to point A. This point highlights the cyclical nature of the trucking industry in that the movement back to point A is completely demand-determined and demand itself is, of course, a derived-demand.

Cabotage in Transborder Trucking: Definition and Regulatory Highlights

Cabotage regulations are an interesting phenomenon from an economic point of view. Economic theory predicts that regulations are put in place for the benefit of the firms being regulated since they are able, through their intimate knowledge of the industry, to "capture" control of the governmental bodies that regulate them **(Stigler (1971); Peltzman (1976)**. Many examples of such "capturing" occur in terms of associations for teachers, lawyers, accountants, physicians, etc. that advise, or even sit on, the regulatory boards that oversee their industry.[4] Indeed, in the days of jurisdictional licensure, the provincial motor carrier regulatory boards put the onus on new entrants to rebut, as necessary, any incumbent firms that felt that further entry would be a detriment to the industry as a whole (see **Kaplan (1989); Boucher (1991)**). Cabotage regulations, which oversee the operations of a foreign truck while on domestic soil, are a different matter. These regulations are international in scope and are, to a great extent, innocuous thus making compliance difficult. As a result, many firms that engage in transborder operations may find it safer to empty backhaul if transborder freight cannot be found rather than attempt a cabotage move only to face delays in terms of establishing legality or facing fines if the move were found to be in violation.

Cabotage is a french word meaning: "coasting" or "to coast" and highlights the fact that its origin was in merchant shipping. The limitation of foreign operators and their equipment in domestic jurisdictions came to be

applied to all modes of transport and the effects it has had on trucking in the United States and Canada have come to the fore since each country's deregulation in 1980 and 1987, respectively, opened up their borders to increased foreign trucking competition.

In both Canada and the U.S., cabotage regulations are divided into a set of regulations applying to the equipment used (which is treated as an import; albeit a temporary one) and another set applying to the driver (who is permitted only to move goods directly in and directly out of the country concerned). The equipment is covered under the respective Customs Acts while the driver is covered under the respective Immigration Acts. Furthermore, the specifics under each Act for Canada and the United States are not always reciprocal and, of course, the level of enforcement on each side of the border is always subject to variation (see **House (1993); Transmode (1991)**).

Two specific cabotage moves of note are known as: (1) incidental domestic moves; and (2) repositioning moves.[5] Highlights of the regulations applied to each will be examined in turn.

A foreign trucker may engage in domestic "incidental" operations such that domestic goods are delivered while carrying a less than full load of imports or exports. The domestic operation is incidental in the sense that it takes place totally within the country in question and that the route is consistent with the transborder route itself. Only minor detours off of the transborder route are acceptable. While "minor" is subject to interpretation, a maximum 20 km deviation might be considered acceptable by customs officials (see **House (1993)**). Still, the regulations are not specific and this subjectivity adds to compliance costs in terms of, say, a carrier's time cost in investigation of operational legality or in its turning down of technically legal operations out of risk-aversity due to lack of clarity. A further insight into these regulations comes from noting that for foreign trucks operating in Canada, the domestic "incidental goods" cannot exceed 30% by weight and value of the international goods carried. For example, a U.S. truck that was 99% loaded with furniture imported into Canada could not, as part of a Canadian incidental operation, pick-up and drop-off while still in Canada, an extremely valuable diamond that could fit in the 1% of space available because the 30% of value rule would be violated (**House (1993);p. 3**). Since

incidental movement regulations apply to exportables and importables, a U.S. trucker transporting U.S. items from a U.S. origin to a U.S. destination but happening to cross into Canada as part of that U.S. operation would not be eligible to engage in Canadian goods transport. As well, there are no limits to the actual number of incidental moves undertaken as part of a transborder operation so long as each conforms to the regulations. The foreign firm and its equipment must, however, be licensed by the particular province and the driver must meet the provisions of the Immigration Act. In the U.S. the provisions for incidental moves are such that they be part of a "regularly scheduled" transborder operation which implies, it seems, that only Canadian truckers with an established record of U.S. entry would receive the privilege to engage in incidental operations by U.S. customs officials.

During an incidental move, exportables or importables must carried simultaneously with the domestic items (and conforming with the 30% rule). If the domestic items are transported alone, it is not illegal; the only change is that the operation has changed from an incidental move to a repositioning move. Thus, straight domestic goods can be transported alone by foreign truckers so long as that operation puts them in position to make a transborder move. The key difference between Canada and the U.S. here is that, for the U.S., the international route must be a regularly scheduled part of the trucker's business no matter how much freight is actually carried each and every time; it is the travel frequency that is important. This provision ensures that the international move was not "artificially structured" so as to foster a cabotage move since cabotage must always be incidental to the overall purpose.

A similar form of movement while on foreign soil is thus the so-called repositioning move whereby one load of domestic goods can be transported anywhere between the drop-off point of the original transborder move and the pick-up point of an export load. That is the repositioning move must place the vehicle into a position such that a transborder operation out of the country will take place. To qualify as a repositioning move: (1) the export load for movement out of the country must be pre-arranged during or before pick-up of the domestic goods before the repositioning move is to take place; (2) this export load must be available for pick-up once the repositioning move has been completed meaning that there can be no stopping over; and

(3) the domestic load drop-off point must be in direct line of the export pick-up point, meaning that only a minor deviation off of the international route is admissible in a similar fashion to incidental movements. Between transborder drop-off point (A) and transborder pick-up point (B), the repositioning move must occur with only minor deviation off of the A to B direct line and the repositioning drop-off need not be anywhere near B; it just has to be in direct line (collinear) to it and the point of exit. One can also infer that doubling back is not allowed since the reposition drop-off must be reached with point B in-between the drop-off and the international border. Note that even though only one repositioning move is allowed while on foreign soil, the length of the move is unrestricted. This may work to Canada's advantage in a fully integrated North American free trade area in that, if a drop-off is made in the U.S., a repositioning move could be made into the deep south of the U.S. if a transborder move out of the U.S. and back to Canada cannot be arranged.[6] Perhaps a relatively short move out of the U.S. to Mexico is possible thus allowing a re-entry to the U.S. with the possibility of taking more domestic U.S. goods up north through another repositioning move that might be long enough to allow for the arrangement of transborder move back to Canada. Canadian immigration legislation is such that domestic repositioning moves require the use of domestic labour but U.S. carriers would be allowed to use their U.S. equipment. Any repositioning move in Canada has a time-limit involved as all foreign trucks, tractors and trailers must leave Canadian soil after 30 days from the date of entry. Since Canadian routes follow east-west patterns by and large, Canadian repositioning moves usually follow that pattern although there seems to be no restriction on north-south movements in Canada provided, again, that entry and exit consist of transborder moves with the repositioning move somewhere in between the natural route. But in the U.S., it is important to note, Customs rules are such that the repositioning move is undertaken so that one is moving in a return direction--- that is, northward--- to the drop-off point to the north and then picking-up a transborder load at least as far north so as to proceed the remaining distance out of the U.S.

One complication involved in U.S. Immigration rules that might affect the possibility of Canada invoking the NAFTA legislation so as to use the U.S. as an intermediate transport area while proceeding into Mexico is to note that the legislation specifies that the exit point out of the U.S. must be at the same border crossing from which entry took place, or at least close to it. Does this

eliminate a Canadian trucker's exiting of the U.S. from the south into Mexico first and then eventually making the way to the same Canada-U.S. border crossing some time later? The legislation is not clear. Furthermore, do the interests of the NAFTA override U.S. domestic Customs legislation?

Unlike the U.S., Canada does provide for switching procedures for U.S. carriers engaged a transborder movement. If the U.S. tractor and/or trailer needs to return to the U.S. along the way of its transborder trip while still in Canada, U.S. replacement equipment can be dispatched from a Canadian location to keep the load moving provided that such a movement of the equipment qualifies as a repositioning move. Of course, a different U.S. driver may be used as well. In terms of emergencies such as equipment breaking down while in Canada, any U.S. equipment and/or drivers may be used as replacements to keep the freight moving even if such a movement of replacements did not constitute a repositioning move. Such emergency procedures are, in fact, allowed in the U.S. as well. Only equipment switching is prohibited by the U.S. in regard to international movements.

The Cabotage Reform Effect on Industry Supply and Welfare

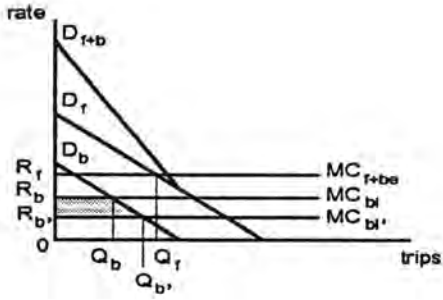
To build in the effect of cabotage reform on the deregulatory kinked supply curve, one must first analyze the effect of cabotage reform on the fronthaul and backhaul decisions in the industry. In the model presented below, only incidental move cabotage will be examined since it requires only an adjustment to the cost structure of the industry whereas a repositioning move would require, in addition, an expansion in backhaul demand since, after reform, other backhaul markets become legally accessible to the foreign carrier.[7]

For every fronthaul there is a corresponding backhaul (either empty or loaded) as part of the joint production aspect of a round trip. Cabotage regulations restrict the possibility of domestic activity on the part of a foreign driver and his equipment while engaging in transborder trucking. A relaxation of such regulations would make that equipment eligible to compete on foreign soil and thus increase the possibility of finding a backhaul. Furthermore, an increase in backhaul prospects allows for a deeper penetration of a Canadian truck into U.S. soil. (Currently about a 200 km

penetration is the norm while U.S. counterparts have access to all urban areas of Canada due to its population by and large straddling the U.S. border). The demand and supply model set out here will be taken to be the entire North American truck freight market. Since cabotage represents a further deregulatory initiative on the part of the two federal governments, there are expected to be further welfare gains from deregulation in the following forms: market rationalization in terms of increased competition in all traffic lanes due to a level playing field; and removal of regulatory uncertainty and lowering of compliance costs. First a fronthaul/backhaul demand model will be closely examined in terms of the cabotage effect and then, secondly, fitted into the regulation/deregulation supply curve analysis. Finally, the welfare gains from cabotage will be explored.

In the fronthaul/backhaul model employed by **Felton (1981)**, empty backhauls occur in the industry when the backhaul demand is low *vis a vis* the fronthaul demand and/or the marginal cost of a loaded backhaul is high *vis a vis* the summed marginal costs of a fronthaul and empty backhaul. To simplify the analysis it is assumed, following Felton, that all marginal costs are constant which will serve to more easily highlight the welfare effects of cabotage reform as specified below.

Figure 2: Backhaul Demand and the Cabotage Effect



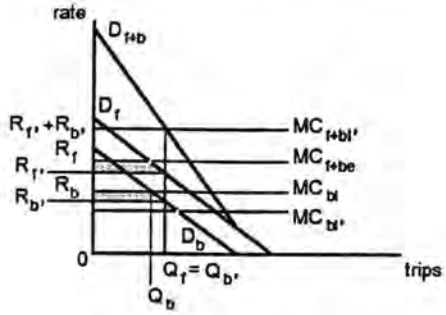
where:
$D_{(f+b)}$ = joint fronthaul/backhaul demand;
D_b = backhaul demand;
D_f = fronthaul demand;
$MC_{(f+be)}$ = marginal cost of loaded fronthaul and empty backhaul;
MC_{bl} = marginal cost of loaded backhaul.

To the extent that the equilibrium fronthaul quantity of trips exceeds the equilibrium backhaul quantity, that difference is the quantity of empty backhaul trips in equilibrium. When cabotage reform is introduced MC_{bl} will be reduced due to the regulatory restriction on domestic activity on the part of foreigners being removed thus increasing the potential for backhaul opportunities. The fall in MC_{bl} can be thought of in terms of, for example, lower administrative and search costs through red-tape elimination, and more direct routing. When the fall in MC_{bl} is not enough to totally remove the presence of empty backhauls in equilibrium, the welfare gain accrues totally to backhaul shippers alone as shown by the shaded trapezoid in figure 2. The simplifying assumption of constant marginal costs means that producers' surplus is always zero. At this point it can be said that if both Canada and the U.S. face situations where the backhaul demands in each of their respective transborder markets consists mainly of opportunities on the other's soil then the unilateral removal of cabotage restrictions by the respective government represents a welfare gain to its domestic shippers. In other words, the removal of a non-tariff barrier (NTB) will increase welfare in a positive-sum sense. It will be assumed for simplicity throughout that the backhaul demand curve consists of the demand on the domestic carriers by foreign shippers which, in a free-cabotage environment, means the demand to carry the foreign country's exportables as well as allowable cabotage moves.

A special equilibrium is proposed in figure 3 whereby MC_{bl} has fallen such that $Q_b = Q_f$ (at the hallowed out point on D_f). At this point the joint demand function $D_{(f+b)}$ becomes operative for the transborder trucking industry and there are no empty backhauls in this equilibrium. As will become apparent below, when $Q_b = Q_f$ a point has been reached whereby all the efficiency gains of lower cost will stop accruing only to backhaul shippers.[8] In any case, if MC_{bl} should fall any further the welfare analysis becomes slightly more complicated. There is no *a priori* reason to assume that the cabotage reform effect will not preclude MC_{bl} from falling below

this special point so its implications must be examined. In Felton's analysis, when MC_{bl} intersects D_b such that it seems that $Q_b > Q_f$ obtains (as can be seen by comparing the two hallowed out points), the true Q_b will have to equal Q_f as given by the joint demand curve as shown in figure 3.

Figure 3: Joint Demand and the Cabotage Effect



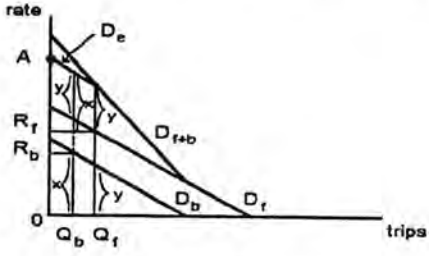
where:
$MC_{(f+bl)}$ = marginal cost of loaded fronthaul and loaded backhaul.

The result is that a portion of the fronthaul costs will be borne by the backhaul shippers due to the jointness of production leading to a welfare gain for fronthaul shippers and a welfare loss for backhaul shippers. It must be shown that the shaded welfare gain always exceeds the shaded welfare loss. This is best accomplished through the use of marginal welfare analysis whereby it is assumed that MC_{bl} falls incrementally (to $MC_{bl'}$) and the resulting additions to the welfare triangles are examined. It is appropriate to use marginal welfare analysis since, in the incidental movement cabotage framework, MC_{bl} cannot be expected to fall greatly under cabotage reform anyway because the transborder route is still in tact and all route-specific cabotage opportunities exist *ex ante*. It is just somewhat easier to locate and act on them. Figure 3 shows a visibly smaller welfare loss triangle (shaded next to the hallowed out point on D_b) as compared to the trapezoidal gain meaning that there is still a net welfare gain to backhaul shippers after a

further incremental fall in marginal backhaul cost.

The result of complete cabotage reform would be to rationalize transborder trucking networks and lead to the elimination of some empty backhauls as a consequence. Thus, in this model the relevant industry demand curve should be the joint demand curve reflecting the reform. In the industry supply framework of figure 1, demand shifts were taken as temporary while its permanent position was given by the kinked point on the supply curve where it was assumed that the number of firms currently in the market were fully adjusted to the current regulatory regime. If cabotage reform is taken as a permanent change in regulatory regime then, given that the higher-joint industry demand is now relevant, demand is thus seen to permanently increase *vis a vis* the pre-reform demand curve which is taken to be only a fronthaul/backhaul demand curve with empty backhauls occurring. The consequence of empty backhauls occurring in the pre-reform case is to generate an effective demand curve[9] (D_e) as shown in figure 4. Since empty backhauls are inevitable in the pre-reform case, this effective demand curve is obviously lower than the joint demand that would obtain under a zero empty backhaul situation after cabotage reform. Furthermore, the effective demand curve in figure 4 cannot, itself, be taken as permanent in structure since Q_b and Q_f are formed independently thus leaving Q_e endogenous and in a state of flux. But the main result is that when empty backhauls are present, the effective demand curve is unstable and always lower than the stable joint demand curve.

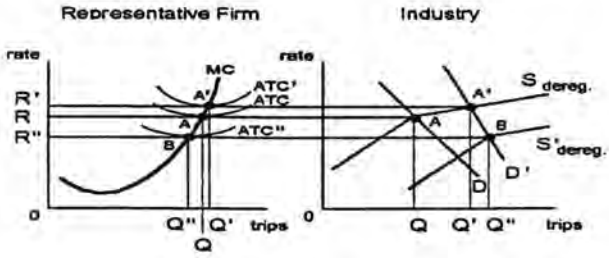
Figure 4: The Effective Demand Curve



Point A is the sum of the given fronthaul rate (R_f) and the R-intercept of the backhaul demand curve. The effective demand is the set of vertical sums of the stationary fronthaul rate and the set of possible backhaul rates.

With a permanent increase in demand along the deregulatory kinked supply curve, the industry can maintain a permanently larger number of firms. In figure 5 it can be seen that the long run equilibrium will move from point (A) to point (B). In effect, cabotage reform works like a permanent demand expansion in the industry with a permanently lower expected rate given that loaded backhaul costs permanently decline. Again, society gains on net because any welfare loss triangle applicable to backhaul shippers due to the rate being set above marginal cost will be outweighed by: (1) a portion of consumer surplus remaining due to the fall in marginal cost; and (2) the net gain accruing to fronthaul shippers in this regard.

Figure 5: Industry Supply and the Cabotage Effect



In the framework for figure 5, ATC contains the appropriate average joint costs of the fronthaul and empty backhaul at Q_f plus the average cost of the backhaul at $Q_b<Q_f$ while ATC' contains the effect of the typical assumption of increasing costs due to firm entry. The MC curve is drawn given the assumption that total industry costs increase equiproportionally meaning that MC in effect shifts up on to itself over the pre- and post-cabotage reform cases. In other words, cabotage reform is assumed for simplicity to affect

only the fixed costs of a firm's operation rather than its marginal (i.e. line-haul) cost.

In comparing the set of figures 2 through 4 with figure 5, it will be noticed that the marginal cost curves for the former were drawn horizontally for simplicity of analysis but can, in the present context, be considered to be the long run marginal costs for the industry indicative of passive supply. The marginal cost curve for the firm in figure 5 is short run in nature and drawn in the familiar fashion and, with each trucking firm approximated to have nearly constant returns to scale in the long run, their long run marginal costs would be relatively flat as depicted at the industry-level.[10]

The process of cabotage reform takes the firm and the industry from point (A) to A' and then to B. Each firm at B is producing less but there are more firms in the industry and the expected permanent rate is now lower after cabotage reform. The movement from (A) to A' highlights the effect of a permanent increase in demand in the post-reform environment. Referring to figure 4, the effect is that the joint demand becomes operational instead of the lower effective demand with empty backhauls present (D_e). At point A', the increasing cost nature of the industry as given by the supply curve leads to an increase in the freight rate when firms enter the industry attempting to capture the positive economic profits being earned by the incumbents. Point A' can also be considered to be the point where loaded backhauls have just increased to the level of loaded fronthauls. But, as mentioned above, there is no *a priori* reason to assume that loaded backhaul cost after cabotage reform might not decrease further. The move from A' to B highlights the effect that occurs when loaded backhaul costs become permanently lower. Using figure 3 as a reference, industry output increases further but with backhaul shippers paying a premium above their marginal loaded backhaul costs and fronthaul shippers paying less than their fronthaul (plus empty backhaul) costs. The sum of these two rates is lower after the fall in backhaul costs and this is reflected in point B as compared to A'. The supply curve (S) was drawn on the assumption of domestic market deregulation only while the new supply curve (S') reflects the permanently lower costs that come about due to transborder market rationalization. The final rate at B can be expected to be lower than that at (A): (1) because D', as the joint demand curve, is always more price-inelastic[11] than the demand curve D_e when the supply curve shifts; and (2) so long as increasing costs are not too pronounced.

Finally, at the firm level this rationalization is reflected by the fall in ATC' to ATC". Since the supply curve has shifted rightward the long run structure of the industry has been positively affected.

Conclusions

It has been shown that incidental-movement cabotage reform is a positive-sum game in that an integrated group of fronthaul and backhaul shippers receive net gains. The trucking industry itself receives efficiency gains and increases in size after reform has taken place, A shortcoming of the analysis is an inability to make precise welfare statements concerning repositioning cabotage moves; indeed, no such statements can be made even in the complete model given here because both the demand and supply curve shift simultaneously. However, this shortcoming exists only within the confines of the economic definition of consumer and producer surpluses. What is definitive, however, is that the trucking industry's efficiency gains will lead to an increase in corridor competition due to the removal of a barrier to entry that leads to complications in regulatory compliance.

Trucking cabotage reform is an exercise in removing uncertainty and lowering operational costs. As such, the welfare gains from such reform have been highlighted in this paper. If a domestic market is able to deregulate its trucking industry, there seems to be no reason why such effort cannot be extended to the transborder market; especially given the establishment of various free trade agreements worldwide. In this vein, negotiations leading to free *modes of trade* agreements are certainly called for.

Endnotes

[1] One of the theses in **Kaplan (1989)** is that firms under regulation were small in number and concentrated into micro-markets that, by their nature, would allow for collusive behavior.

[2] The industry demand curve itself will be defined more fully below. Suffice it here to say that it is an "effective" demand curve of fronthaul and backhaul movements that obtains.

[3] A technical point of note is that the welfare measure is accurate only when the demand curve is an income-compensated (Hicksian) demand curve. This is done because the goal is to examine the removal of the regulatory distortion in the context of the effect of a change in the rate while demanders, individually, remain at constant utility levels; that is, only the substitution effect induced by regulatory change is to be captured. The diagrams would retain their current form except to the extent that a compensated demand curve would be steeper than would a market (Marshallian) demand for a normal good.

[4] An application of the "capture thesis" to the area of teacher certification and university training of teachers can be found in **Dean and Prokop (1996)**.

[5] A third type of cabotage move, peculiar to the Canadian regulations, is the so-called "sufferance warehouse pick-up" sanctioned by Canada Customs in January of 1991. In this operation the delivery concerns international freight that had been originally dropped off at a sufferance warehouse in Canada pending a customs release. Any foreign trailer (including the original that dropped off the goods at the warehouse) may be used to pick up those goods when released so long as the conveyance equipment entered Canada in accordance with tariff item 9801.00.00 meaning that it must have entered Canada with a load of inbound foreign freight. As far as the tractor used in the pick-up is concerned, in a similar sense, if it is not the original tractor used in the transport of the newly released goods, the tractor is required to have entered Canada with an international load. More detail is provided in **Transmode (1991)**.

[6] It is interesting to note that neither the Canada-U.S. Free Trade Agreement (FTA) nor the North American Free Trade Agreement (NAFTA) explicitly discuss integration or coordination of transportation systems so as to facilitate the movement of goods. It is only recently that the European Economic Community began such an integration process given that the Treaty of Rome came into effect in 1958.

[7] **Mishan (1971)** notes that a shifting consumer demand curve provides no real insight into the welfare measurement problem.

[8] Strictly speaking, a portion of the gain in consumers' surplus to backhaul shippers is appropriated by fronthaul shippers leaving the former with a rate above the marginal loaded backhaul costs which will create a quasi-economic rent for the trucking firms obtaining backhauls.

[9] This effective demand curve is taken to be the sum of the marginal willingnesses to pay over the various potential backhaul points on D_b and the *stationary* fronthaul quantity demanded, Q_f.

[10] My thesis will provide more details concerning the relationship between the firm and the industry in terms of firm-level responses to industry-demand expansion and contraction that occur along the kinked supply curve.

[11] As long as a linear demand curve is: (1) contained completely under-neath another linear demand curve; or (2) intersects another linear demand curve along the quantity-axis then the higher demand will always be more price-inelastic than the lower demand.

Bibliography

Boucher, M. (1991). "Rent-Seeking and the Behavior of Regulators: An Empirical Analysis." *Public Choice*. vol. 69. pp. 51-67.

Dean, J. and D. Prokop. (1996). "Internal Threats to the University: Special Interest Groups and Rent-Seeking Behavior." *The Canadian University in the Twenty-First Century*. Centre for Higher Education Research and Development, Research Monograph no. 3. The University of Manitoba Press. pp. 67-82.

Felton, J. (1981). "Impact of ICC Rate Regulation upon Truck Back Hauls." *Journal of Transport Economics and Policy*. vol. 15. no. 3. pp. 253-67.

House, M. (1993). *A Canada-U.S.A. Cabotage Perspective*. Motor Carrier Branch, Policy Coordination Group. Transport Canada.

Kaplan, H. (1989). *Policy and Rationality: The Regulation of Canadian Trucking*. Toronto: The University of Toronto Press.

Mishan, E. J. (1971). *Cost-Benefit Analysis: An Introduction*. New York: Praeger Publishers Inc.

Peltzman, S. (1976). "Towards a More General Theory of Regulation?" *Journal of Law and Economics*. vol. 19. pp. 211-40.

Prentice, B. (1994). "The Stability/Efficiency Regulatory Trade-Off: Policy Implications for For-Hire Trucking." *Canadian Transportation Research Forum: Proceedings*. pp. 494-507.

Stigler, G. (1971). "The Theory of Economic Regulation." *Bell Journal of Economics and Management Science*. vol. 2. pp. 137-46.

Transmode Consultants Inc. (1991). *Implications of Alternative Cabotage Rules*.

TURNING THE NEW REALITY INTO REALITY: LESSONS FROM EUROPE

by Mario M. Bozzo and Lee S. Sims, IBI Group

INTRODUCTION

With greater European ties and a growing European transport policy, attitudes regarding transport in Britain have changed. After a decade of neglect, there has been an emergence in recent years of a new co-ordinated approach to transport planning within the UK which signifies a move toward the mainland European policy approach of integrating transport strategies into a comprehensive whole.

The importance that institutional structures have on the implementation of transport policy cannot be emphasized enough. To this end, as governments in Canada look to re-organize, the way in which transportation is delivered will inevitably change. It is now widely recognized that the "new realities" of transportation must involve the co-ordination of transportation strategies.

Against this background, this paper will discuss the relatively successful development of Integrated Transport Strategies in Europe, particularly recent developments in Britain. In the face of government change in Canada, there is great potential to establish the necessary institutional structures to implement truly integrated transport plans. With reference to European examples, this paper will illustrate the key ingredients in developing Integrated Transport Strategies.

THE MEANING OF INTEGRATION

By integration we mean developing individual transport measures together in a co-ordinated fashion, rather than implementing individual measures on their own. This requires a more holistic way of looking at transport, packaging measures that balance all modes of transport and groups of travellers. The Association of Metropolitan Authorities (1990), who first advocated Integrated Transport Strategies in the UK, defines integration as:

- integration between authorities;
- integration between measures involving different modes;
- integration between measures involving infrastructure provision, management of that infrastructure and pricing mechanisms; and
- integration between transport measures and land use planning measures

Integration at the operational level should not be understated, for it is at this level where transport is most visible to the traveller. And thus, co-ordinating pedestrian crossings with a local traffic management scheme, co-ordinating public transit and rail timetables or providing 'seamless' travel between modes through fare integration become equally important in the concept of integration.

There are several key elements required in developing an integrated transport strategy. These include:

- the Policy making framework and the underlying institutional structures which provide the context for which integrated plans are created;
- the range of measures considered and the emphasis placed on them;
- the analytical requirements, both in modelling, to test the range of measures against stated objectives, and in appraisal, to assess all modes included in a package on similar terms; and
- the mechanisms for implementing and financing integrated transport strategies

Each of these elements are discussed, in turn, below.

POLICY MAKING

A TRANSPORT POLICY FOR THE EUROPEAN UNION

One of Europe's assets is in its building of a trans-national policy consensus. European Transport Policy has its roots in 1974 with the signing of the Treaty of Rome. Here, a Common Transport Policy (CTP) was agreed; its major objectives included:

- re-inforcing the internal market;
- developing a coherent and integrated transport system with the best technology available;
- developing transport infrastructure to reduce disparities between regions;
- adopting programmes for environmental protection; and
- promoting safety, social measures, and links to third worlds.

There is one paramount principle in the delivery of these objectives - subsiduary (Miles, 1995). This means that the EC can make proposals only where the Member States are unable to achieve the objective and the EC is better able to do so. This principle has led the EC to focus on trans-frontier links.

Transport Policy in Europe gained strength since the signing of the Maastricht Treaty in 1992, where a clear framework for developing the Trans-European Network (TEN) emerged. Covering transport, telecommunications and energy, the Trans-European Network includes roads, railways and inland waterways. Recognizing the complexity of Trans-European projects, the EC establishes *ad hoc* 'project authorities' for each initiative in order:

- to ensure that projects are run efficiently so as to overcome regulatory obstacles;
- to instill flexibility so that local needs can be catered to; and
- to create a forum for representatives from Member States as well as from the private sector, who take some of the financial responsibility and accelerate implementation.

THE IMPACT OF EUROPEAN POLICY ON MEMBER STATES

The impact that this European Policy has on national policies is strong. In promoting sustainable mobility, the EC has provided generous political and financial support for multi-modal projects, particularly high speed and conventional rail. At the urban level, the most significant impact has been in Europe's commitment to research and development. In 1994, the EC established the Fourth Research Framework, an R&D programme to which transport alone will receive $6 billion over 5 years (1994-1998) (European Commission, 1994). In comparison, the overall budget of Transport Canada's Transportation Development Centre is less

than $20 million per year. Europe is currently initiating discussion for a Fifth Research Framework which will carry Europe's R&D into the next century.

Yet there is evidence to suggest that Europe is only able to fulfil its transport objectives at a rate dictated by the individual countries. This is more visible in Britain than on the mainland since, for example, Britain's privatization policies which have been successful in attracting private financing for capital projects is likely to have an impact elsewhere in Europe than vice versa.

EXAMPLES OF NATIONAL POLICY FRAMEWORKS

At the National level, one of the most clearest statements of policy comes from the Netherlands. Underpinning planning decisions, with strong links to the National Plan for the Environment (which identifies targets for external impacts of transport) and Regulations for Physical Planning, the Second Transport Structure Plan (SVV) sets out long term transport policy for the country. Linkages are also made to land use planning, as approval of new high density developments is closely related with the accessibility to public transit. The SVV, therefore, provides the context in which all projects have to be based, leading to a strict and sometimes lengthy approval process (De Jong, 1995).

Regions developing transport plans within this framework find it difficult to achieve the National objectives and targets without a considering package of measures or collaborating with other local/regional governments, for which Central government strongly encourages.

Transport Policy Integration in the UK

The UK government is less committed to providing a National policy framework; however, a common Central policy covering all modes and scales of planning is the drive to create private sector management and operating industries - evidenced by bus and rail privatization, and more recently, plans to privatize road network management and maintenance.

It was the completion of two comprehensive transportation studies in the early 1990's which initiated a fundamental shift in government policy. Studies in West Yorkshire and the West Midlands illustrated that objectives could be better achieved if strategies were combined in a co-ordinated fashion as opposed to implementing individual measures on their own. However, the findings of the studies could not be implemented, since government funding at that time was earmarked for individual measures. Given this, the Department of Transport initiated "Package Bids" under which funding for local projects was to be requested for a set of transport measures, designed to work together to achieve explicit objectives. This would be accompanied by a detailed analysis of how the chosen measures would meet these aims.

The government also stated their reduced emphasis on infrastructure provision, as they noted, "it will not be possible to solve transport problems in urban areas by a programme of road building and investment" (DoT, 1993). Rather, the DoT placed emphasis on low cost management measures, public transit to reduce a modal shift, and measures designed to reduce the need to travel (which acknowledged a link to land use).

In Ontario, a different and less technical approach to supporting and promoting integrated transport planning is being initiated. The Ministry of Transportation, Ontario has embarked on a series of long range, integrated, regional transport studies, the purpose of which is to create an updated provincial transportation master plan, projected to the year 2026. These studies, and the implementation of their recommendations, involve a significant degree of multi-modal partnerships and stakeholder consultations. However, unlike the European approach which ties the funding of local projects to the production of integrated plans, Ontario has backed away from the funding or 'carrot' approach to encouraging municipalities to adopt integrated plans.

THE RANGE OF MEASURES CONSIDERED

The range of transport measures that may be included in an integrated strategy is well documented, and can be broadly categorized into four main areas. These include:

- infrastructure provision - new roads/rail lines, Park and Ride facilities, etc.;
- management - Intelligent Transport Systems, UTC, car sharing, bus/rail service frequencies, etc.;
- pricing - parking charges, road pricing, fuel prices, fare levels and structure, etc.; and
- land use - density and mix of development, location relative to public transport, etc.

At the strategic level, May (1995) explains that it is the synergy between measures which is the key to success in the development of Integrated Transport Strategies, and the principal skill in policy development is the identification of measures which might achieve synergy.

As depicted in Exhibit 1, synergy is achieved by deciding which measures can:

- re-inforce or complement each other (C);
- make other elements of the strategy financially feasible (F); and
- be publicly acceptable (P). (May and Roberts,1995)

MODELLING

Within the range of measures to be included in a integrated strategy and the need for these measures to achieve a number of stated objectives, a transportation model is created to predict future travel patterns by modes that will result from implementing different transport measures; for each measure, different scales and intensities are also tested. Through this process, it can be estimated how well a certain package of measures (with various scales and intensities) can achieve the stated transport objectives.

Recognizing the need to evaluate the effects of a number of urban/regional policies together, and to identify more accurately the effects of strategic infrastructure investments, the UK DoT is currently developing a National Transport Model. Denmark and the Netherlands have alrready developed their respective National Transport Model, and these have been incorporated into their decision-making processes.

	Highways	PT Infra-structure	Park and Ride	Parking Supply	Traffic Mngmnt	Bus Priorities	Traffic Calming	Parking Control	PT Service Levels	Infor-mation Systems	Parking Charges	Road Pricing	Fuel Prices	PT Fares	Develop-ment Control
Highways						C	C					C			C
PT Infrastructure								C/P			C/P	C/P	C/P		C
Park and Ride		C						C			C	C			
Parking Supply							C	C							C
Traffic Management						C	C		C	C		C			C
Bus Priorities					C		C		C						
Traffic Calming								C/P			C/P	C/P	C/P		
Parking Control		C	C	C		C			C	C					C
PT Service Levels			C				C	C/P			C/P	C/P	C/P		C
Information Systems			C		C		C	C	C		C	C			
Parking Charges		C/F	C/F	C/F	C		C/F	C	C/F					C/F	
Road Pricing	F	C/F	C/F	F		C	C/F		C/F	C				C/F	
Fuel Prices	F	C/F	F	F			C/F		C/F						
PT Fares		C/F	C/F	C		C	C/P	C	C		C/P	C/P			
Development Control	C	C							C		C	C			

Exhibit 1: A Matrix of Interactions Between Transport Strategies at the Strategic Level *Source: May and Roberts (1995)*

Key: Measures in the left hand column can re-inforce the measure in the appropriate column by:
C: Complementing it.
F: Providing finance for it.
P: Making it more publicly acceptable.

The objectives of transport policy are often considered to be in conflict, with the need for the package of measures to achieve financial feasibility on one hand, and sustainability on the other. In an effort to find the package of measures with the correct intensity/scale which best meets both objectives, strategy combinations are tested against the stated objectives with respect to the do-minimum strategy. This process can involve many model runs and consume valuable time and resources Conceivably, for every strategy package, model runs would have to be conducted, focusing in turn on each objective. Even if the objectives were combined to form one objective function, a large number of runs would still have to be conducted to find, through a 'trial and error' approach, the most effective mix, scale and intensity of a preferred package of measures that best achieves that objective(s).

Optimizing Objectives

A process of finding optimal combinations of policies was developed by the Institute For Transport Studies, University of Leeds (Bonsall et al, 1994; May et al, 1995). This iterative process uses a regression model generated by GLIM (Generalized Linear Interactive Modelling) software to represent the behaviour of the strategic transport model. Using outputs from the transport model, the regression model includes the desired objective, i.e., a sustainability or an economic efficiency objective (the dependent variable) and several policy options (the independent variables). The independent variables are in the form of continuous (road pricing, public transit fares) and discrete variables (level of infrastructure investment represented as a dummy variable). Given this, the regression model is designed, by differentiating equations, to maximize the dependent variable.

Using the transport model for Edinburgh, the optimzation procedure was applied to both an economic efficiency (Bristow et al, 1994) and a sustainability objective (Bozzo, 1995). the work demonstrated that the strategy specification for Edinburgh's transport is sensitive to the objective being optimized. In addition, relative to a strategy optimized via an economic efficiency objective, a sustainable transport strategy for Edinburgh will lose half its NPV (Net Present Value) contribution, but gain about 50% improvements in sustainability.

The European Union is currently supporting a broader effort in which an integrated strategy for nine cities throughout Europe will be developed and the optimization procedure will be applied to draw conclusions about the differences between the most economically efficient strategy and the most sustainable strategy.

APPRAISAL

An important aspect of any evaluation of a transport strategy is to assess the monetary benefits and costs and any external effects associated with a strategy. Since measures and modes are combined in a package of measures, it is equally important that a common appraisal approach for all modes is adopted. In doing this, modes will be compared on a level playing field, giving little room to misallocate resources.

In the UK, the DoT has recommended municipalities use the Common Appraisal Framework, (MVA et al, 1994) to evaluate their integrated packages. The Common Appraisal Framework has been designed to appraise multi-modal transport packages in urban areas, taking into account both the direct effects and a wide range of externalities. These externalities are evaluated in the form of a checklist covering all aspects of environmental, safety and social impacts. There are still concerns, however, that the DoT does not require transport packages as a whole be evaluated, as well as their continued requirement of 'restricted cost-benefit appraisal' for public transit investments. The general criterion for this restricted appraisal requires that the cost of public transport investments should be re-couped by the user, which likely results in higher fare levels.

IMPLEMENTATION AND FINANCING

While mainland Europe has had more of a tradition in committing to integrated transport strategies, signs of implementation and financing of integrated transport strategies are less evident in the UK. The Birmingham study, for example, was over-optimistic regarding likely available levels of funding, resulting in a small percentage of Birmingham's preferred package being implemented to date. In general, many persons are critical of the UK DoT's continually decreasing

funding availability for the "package approach". In addition, local authorities are becoming frustrated with the barriers to implementation. These barriers include:

- bus and rail privatization - which have restricted the influence local authorities have on service levels, operational integration, and fares;
- private non-residential parking (PNR) - these PNR spaces account for as many as 50% of parking spaces in some cities. Local authorities cannot control the supply or pricing (if any) of these spaces. Control by local authorities would involve changes to legislation, and the government has not showed any signs of getting into urban transport policy. The effect is that urban parking policies become less effective;
- local government re-organization - there has been a major re-organization of local government which saw the abolition of "regional authorities" with smaller local authorities taking over responsibilities. As a result, there have been difficulties with formulating regional transport policies and inconsistencies in infrastructure investment, demand management, land use planning and economic development; and
- road/public transit budgets - while funding for municipalities is available as a "package", budgets for road and public transit infrastructure is still decided separately which means that biases in the ability to fund different types of measure are produced.

To ensure effective implementation, it is clear that there needs to be a long term commitment to financing, and any barriers that prevent effective implementation be removed.

CONCLUSION

As Canadians look to discover new ways of delivering transportation in a more effective and sustainable manner, the rich tradition of integrated transport planning in mainland Europe and the emerging commitment in the UK to a more co-ordinated approach offer possible solutions.

Our European partners have shown that with an institutional structure in place to foster integrated transport planning and with a commitment to eliminating barriers and financing projects, a range of measures can be considered and implemented to create synergy. Different combinations, scales and intensity of measures can be tested and optimized to see which

mix best meets policy objectives, and multi-modal appraisals can be developed to assess the monetary costs and benefits of a 'package' of measures.

Within this integrated framework, a common context for informed public debate on a wide range of issues that are typically dealt with individually can be created, making it possible to debate the relative merits of investment in system expansion, system management/optimization and maintenance/preservation on the same agenda.

REFERENCES

Association of Metropolitan Authorities (1990). Changing Gear: Urban Transport Policy into the Next Century.

Bonsall, P, A.L Bristrow, A.D. May, A.F. Fowkes (1994). A Tool for the Optimization of a Transport Strategy. Proc PRTC Summer Annual Meeting, Sept, PTRC London.

Bozzo, M.M. (1995). The Evaluation and Optimization of Integrated Transport Strategies: Tests Based on Edinburgh. MSc (Eng) Dissertation, Institute for Transport Studies and the Department of Civil Engineering, University of Leeds.

Bristrow, A.L., A.D. May, P. Bonsall, A.F. Fowkes and B.G. Still (1994). The Optimization of Integrated Transport Strategies: Tests Based on Edinburgh. Working Paper 425, Institute for Transport Studies, University of Leeds.

De Jong, M.A. National Transport Policy in the Netherlands (1995). Proceedings, Institution of Civil Engineers (Transport), Aug.

Department of Transport (1993). Transport Policies and Programmes, Submission for 1994-95. Local Authorities Circular 1/93, London DoT.

European Commission (1994). Fourth Research Framework: Information Package, Brussels.

May, A.D. (1995). The Development of the UK Package Approach and the Common Appraisal Framework. Draft, unpublished.

May A.D. and M. Roberts (1995). The Design of Integrated Transport Strategies. Transport Policy.

May, A.D., A.L. Bristrow, P. Bonsall and A.F. Fowkes. (1995). A Streamlined Approach for Package Approach Bids. Traffic Engineering and Control, Feb.

Miles, J.C (1995). Recent Developments in European Transport Infrastructure Policy. Proceedings, Institution of Civil Engineers (Transport), Aug.

MVA Consultancy, The Institute for Transport Studies and Oscar Faber. Common Appraisal Framwork for Urban Transport Projects. HMSO, London.

What is Canada's National Transportation Policy?

by Javier Caceres and Darrell Richards, Transport Concepts

Summary

In 1994, Transport Canada outlined a new vision for transportation in Canada. Regulation and subsidies would be phased out in favour of a more market-oriented approach based on deregulation, privatization, and commercialization.

Have distortions been eliminated in the freight marketplace? Not entirely. For example, Transport Canada has stated that the Government will fund operating costs of the Seaway when traffic volumes are low, as well as major Seaway asset rehabilitation. Trucking is subsidized through publicly funded infrastructure. Even if there were a level playing field in terms of direct government support, the marketplace currently does not price the full external costs of accidents and pollution. This lack of full cost-accounting favours the modes with the highest external cost.

An optimum freight transport system would minimize total costs, including external costs, by reflecting the total costs in market prices. In reality, this will be tempered by competitive conditions of the continental market. Canada needs a comparative analysis and harmonization of U.S. and Canadian transport legislation, regulation, policies, taxes, and subsidies. It needs to work together with our NAFTA partners for a continental transportation policy based on a level playing field for each mode among the countries and within each country.

Introduction

> A fundamental and persistent problem in the history of Canadian transportation is the interplay of two radically different concepts: straight business principles on the one hand and such matters as national unity, the movement of trade through Canadian ports, the opening up of new areas, defense, and avoiding the ruination of national credit on the other.
>
> A.W. Currie, Canadian Transportation Economics (Toronto: University of Toronto Press, 1967), p.1, our emphasis added

The 1994 Federal transportation policy and resulting Canada Transportation Act unequivocally chose the business principle side of Currie s dilema. Is the debate on Currie s question settled once and for all? Or have we merely put political intervention out of sight while making loud declarations of faith in market economics? Is transportation policy becoming the politics of transportation? Practical realities have led to a number of inconsistencies in the government s user-pay policy which this paper will explore.

Historical Development

Governments played an important role in developing Canada s infrastructure, starting with the Lachine Canal in the early 19th century. Governments provided loans, loan guarantees, subsidies, grants, and franchises or charters to encourage the development of private railways. Governments were also directly involved in building and operating road, marine, and airport infrastructure and the marine and air navigation support systems.

Development of the prairie wheat economy was linked to the creation of a branch plant economy in Ontario and QuØbec. Tariffs to support central Canadian manufacturing went hand in hand with regulated Crow rates for the export of prairie wheat. The Crow grain rate until 1983, and the Crow Benefit scheme from 1983 to 1995, regulated the relationship between grain production, transportation, handling, and marketing. It also pushed grain east-west when it more naturally would flow north-south.

Competition in the transcontinental air market was introduced in the late 1950 s. The expansion of trucking and road networks as competition to the railway oligopoly also led to a measure of deregulation in the 1967 National Transportation Act. The political-economic forces that created the national policy and infrastructure began to unravel in the 1970 s as the overall force of Canadian nationalism declined. This coincided with the transformation of the prairie grain cooperatives into commercial enterprises, and the entry of multinationals into the Canadian export grain industry.

The Conservative Government elected in 1984 published a new transportation policy, Freedom to Move. This steered Canada away from regulation, subsidies, and government involvement toward a free market approach. At the same time, Canadian transportation systems were facing stiff competition from U.S. carriers. The concept of globalization and the growing popularity of just-in-time deliveries drove this policy shift. This led to a new National Transportation Act in 1987 which implemented the Freedom to Move philosophy.

The Canada U.S. Free Trade Agreement is restructuring the Canadian economy and transportation system on a continental basis. However, there is no attempt to harmonize overall transportation policies and subsidies or to define the role of different levels of governments.

Lobbying by the U.S. marine sector kept domestic U.S. marine protection out of the agreement. Attempts were made to compare the marine subsidies available to U.S. inland grain movement with the Canadian Crow Benefit scheme. This is now irrelevant since Canada abolished the Crow subsidy, but the inland marine subsidies continue in the U.S. The potential deregulation of the Canadian grain industry will create a continental grain and transportation market.

In 1994 Transport Canada prepared a working paper on a Vision of Canadian Transportation in the 21st Century. In that vision the government would no longer own or operate the transportation system. Its role will be narrowed to safety regulation. The new policy called for solutions that are consistent across modes, and for treating freight rail, marine, road, and air systems equally in terms of self-imposed costs and government services. It also recommended making transportation more efficient and affordable by removing subsidies and cross subsidies, enhancing cost recovery and promoting a user pay/user say system based on market forces.

The Federal Government acted swiftly to implement the new policy in certain transportation markets and sectors. The February 1995 federal budget set in motion the elimination of the western grain transportation subsidies, the Maritime Freight Rate subsidies, and the Atlantic Region Freight Assistance program. Airports were transferred to local governments or to local airport authorities. The Canada Marine Act transfers major ports to port authorities with no resort to the Federal Treasury. The Canada Transportation Act furthered deregulation of the rail system. But.....

We Are Not in Transportation Heaven Yet

Relationship to the U.S. System

The transportation debate in Canada, as usual, is too inward-looking. The reality of NAFTA and the continental market is not being considered as a framework for policy discussion. Economic integration with the U.S. is a reality in patterns of trade. Political integration may not be too far behind. The fundamental question is how to make our production and shipping system more competitive while at the same time deciding if we want to maintain our identity and standard of living as Canadians in a North American market. A key factor in competitiveness is how transportation legislation, regulations, and policy in Canada compares with that of our major competitor, the United States.

a) Public Support for Infrastructure

Notwithstanding the impression of free markets, the Untied States continues to provide public support to transport infrastructure:

- Compared to shippers in Canada, shippers in the U.S. do not have access to competitive line rates. This tends to shield the market for the U.S. railroads from competition.

- The objectives of the U.S. Staggers Rail Act are to ensure adequate revenues for the railways and prevent railroad bankruptcies

- Under the Intermodal Surface Transportation Efficiencies Act, U.S. Highway funds can benefit the railroads, for example to support truck-rail intermodal facilities and rail access to ports.

- The inland waterway system in the United States is maintained by the U.S. government at public expense with no direct charges to users. [1] And under the Jones Act, U.S. domestic water commerce is reserved for American ship owners and workers and protected from outside competition.

- Several of the export grain terminals in the Pacific Northwest of the U.S. are publicly owned and leased at low cost to grain companies.

- Many U.S. ports receive substantial financial support from state and local governments. American ports are generally exempt from the payment of municipal taxes. Many pay grants-in-lieu of taxes, but these are set at levels well below the amount of property taxes which would otherwise be levied. Some U.S. ports have powers of direct taxation which provide them with significant revenues. [2]

---

[1]Manitoba Transport Institute, International Competitiveness of Western Canadian Transportation, prepared for Western Economic Diversification Canada, 1994, ch.5-p.34

[2]Transport Canada, briefing note on Bill C-44, the Canada Marine Act

- U.S. ports have access to credit over a longer term and at a lower interest rate than Canadian ports They also have access to the pooling of resources through the Army Corps of Engineers and the Harbor Maintenance Trust Fund.

- The U.S. administration has approved a $400 million loan for the Alameda rail corridor project serving the Ports of Los Angeles and Long Beach, which will compete with B.C. ports

Canada has put its railway and port infrastructure on a user pay self sufficiency basis. Now the Canadian rail-port export routes will have to compete with rail-port export routes in the U.S. where the ports receive public support. And the Canadian railways may seek access to U.S. ports as an alternative to Canadian ports. CP has already done this on the Atlantic Coast. CN is requesting access to New York.

b) Objectives of a Transportation Policy

The objective of the United States surface transportation policy is to develop a National Intermodal Transportation System that is economically efficient and environmentally sound, provides the foundation for the Nation to compete in the global economy and will move people and goods in an energy efficient manner. [3] The principle U.S. policy document is Toward a National Intermodal Transportation System by the National Commission on Intermodal Transportation. The National Highway System includes links to ports and rail terminals.

[3]Intermodal Surface Transportation Efficiencies Act

The concept intermodal does not appear in the declaration of National Transportation Policy in the Canada Transportation Act. Yet intermodalism is a key tool in reducing overall costs, by making the best use of each mode, and by enhancing productivity by exploiting available infrastructure capacities. In summary, the Canadian Act declares that a safe, economic, efficient, and adequate network of viable and effective transportation services accessible to persons with disabilities and that makes the best use of all available modes of transportation at the lowest total cost is essential to serve the transportation needs of shippers and travelers, including persons with a disabilities, and to maintain the economic well-being and growth of Canada and its regions.

The purpose of the Staggers Rail Act in the U.S. is to provide for the restoration, maintenance, and improvement of the physical facilities and financial stability of the rail system of the United States. One of the goals is to assist the railroads to meet the demands of interstate commerce and national defense.

c) Jurisdictions

The U.S. Intermodal Surface Transportation Efficiencies Act integrates the planning of Federal, State, and Metropolitan governments. Each State and large Metropolitan area must develop an integrated transportation plan through a public particpation process. Highway funds can be spent on intermodal projects, and there are set asides for projects that improve air quality and reduce congestion. The U.S. Federal Government regulates interstate rail commerce.

In Canada, the Federal Government is responsible for rail, air, and marine. Provinces and municipalities are responsible for road. The Federal Government has vacated most of its infrastructure responsibilties, keeping only the safety function. Rail is a private responsibility. Major airports and ports are being transferred to local authorities. Smaller airports and ports, except in remote areas, are being transferred to lower governments. Transportation policy at the federal level, except for safety, is non existent. It is up to private companies to foster transportation policies while maximizing revenues in the short term.

The U.S. Federal Government retains safety regulation of short lines within a State, if the line generates traffic that moves interstate. It also regulates truck weights on the Interstate system and truck lengths on the National Network which includes the Interstates. ISTEA limits longer combination vehicles with gross weights in excess of 80,000 pounds to the States in which they already operated on June 1, 1991. In Canada inter-provincial and international truck regulation has been delegated to the provinces. The Federal Government does, however, act as a broker in trying to standardize truck safety, size and weight regulations across the country.

Inconsistencies Between Modes

Transport economists state that resources will be allocated under optimum conditions when each mode of transport can compete in a level playing field. We do not have a free market in transportation. The Federal budget of 1995 eliminated $700 million in annual rail freight subsidies. The rail freight system and ports are now on a user-pay basis. Although the direct cash payment to trucks under the Maritime Freight Rates system was abolished, the subsidies to trucking operations through road costs continues.

IBI Group estimates the annual infrastructure subsidy to trucking in Canada at $450 million. This is the heavy truck (>5 tonne) share of road costs in excess of fuel taxes and license fees paid by trucks. The Royal Commission on National Passenger Transportation estimates the average semi-trailer truck is subsidized $10,000 annually through the road system costs. The deterioration of highways in Canada is being debated as if highways were a government responsibility. This is an anomaly in the free market policy. The Canada Marine Act enables future federal funding of Seaway operating costs and major asset rehabilitation costs.

Safety regulation adds to costs. Because the level of regulation varies by mode, the cost of safety compliance varies. The Federal Government is responsible for the safety of federal railways, airlines, and marine. Accident investigation is separated from safety regulation in order to avoid conflict of interest. Interprovincial truck safety has been delegated to the provinces, although the Federal Government has tried as a broker to get the provinces to sign on to a National Safety Code. This has not been completely successful. This has resulted in different standards for different modes.

Railways must report all accidents, including near collisions. Accidents are classified by type so that trends can be tracked and problems identified. Rail accident data is available to within two-three months, while commercial vehicle accident data is two-three years old.

Truck accident data is unreliable. It is difficult to calculate accident rates by mile driven. The B.C. Auditor General s report on truck safety has identified serious problems with overweight trucks, poorly maintained trucks, and drivers exceeding hours of work limits. The Canadian Council of Motor Transport Administrators reports that 32.6% of trucks in Canada randomly inspected in June 1996 were ordered out of service due to poor mechanical fitness.

There are no market signals for indirect or external costs. Allocation of resources based on direct costs alone may not lead to the least cost system because it does not include the consumption of public and social resources.

The following table illustrates the different levels of indirect costs for freight modes of transportation that are not reflected in prices:

Subsidies (1994 cents) per freight tonne kilometre

Mode	Basic Subsidies (largely infrastructure)	External Costs	Total Subsidies
Intercity Truck	.4	2.12	2.52
Intercity Rail	-.04	.41	.37
Intercity Marine	.1	.16	.26

Source: IBI Group, in association with Boons, Jones, and Associates, 1995

These indirect costs for freight transportation are a substantial burden on the Canadian economy, a total $4.6 billion a year in freight costs that are absorbed by taxpayers, accident victims, and society at large. There are market signals for the external benefits of just-in-time logistics through premium freight rates for punctual and rapid delivery. Because the hidden costs are not reflected in market prices, the market is distorted toward the mode with the highest hidden costs per tonne-kilometre. In effect, our highways have become rolling warehouses because the true costs of just-in-time logistics are not reflected in the marketplace.

There are other inconsistencies creeping into the Federal transportation policy. The Prince Rupert rail line and port have been put on a completely user-pay basis. Now the Federal Government is considering funding improvements to the Churchill route. It has also passed legislation enabling future subsidies to the Seaway route which would give it a competitive advantage over Prince Rupert.

Political interference is hidden, but still there. Provinces and grain companies still consider grain transportation a political rather than market issue. For example, they are asking the federal government to investigate grain transportation delays that occurred in the winter of 96/97. This is political intervention in the market. The Federal Government is providing financial backing to a short line railway in the GaspØ through a regional development program, while prairie branch lines are on their own.

Policy development is still a gradualist process because we do not want to confront the full impact of user-pay on the political environment. For example, the Crow Benefit was canceled, but rate regulation was maintained for at least another four years.

Where We Are Going

The macro framework driving the restructuring of the transportation sector is set by globalization and NAFTA. The Federal Government is no longer involved in national transportation policy, except safety regulation. Canada does not have national goals and objectives to enhance Canadian competitiveness within the continental market.

Under these conditions, we can expect the following developments in Canadian transportation:

- Increased investment in north-south highways, and increased diversion of exports away from the east-west Canadian system to U.S. rail lines and ports, undermining the viability of the Canadian system.

- Accelerated path of economic integration into a continental market, with investment in north-south links such as NAFTA highway corridors

- Deregulation of the Canadian grain industry, absorption of the Canadian grain companies and Coops by transnational companies, and the creation of a continental grain market.

- Pressure on the viability of Canadian ports and railway infrastructure.

- Pressure from Canadian shippers to allow U.S. railroads to serve traffic on Canadian rail lines through more open access rules.

- As the east-west Canadian system deteriorates, economic activity will migrate from areas more than say 200 kilometres from the U.S. border to southern areas with close highway access to the U.S. transportation system.

Options

On the assumption that we need options as a country, and since that means a certain amount of government intervention, there may be need for further analysis and discussion of transportation policy in the context of a Canadian identify in a continental market. If that is the case, we need at least a long-term planning process and harmonization of our policies with the U.S. system.

Swedish Model

The 1963 transportation policy was based on direct business costs. The 1977 policy included tangible indirect costs (e.g. material cost of accidents). The 1988 policy included intangible indirect costs (e.g. value of life). Indirect costs are included both in investment analysis and pricing (in practice full infrastructure costs are not recovered from freight).

Both rail and road infrastructure are put on a cost/benefit basis in order to level the playing field. This contrasts with the North American situation where taxpayer assistance to road construction is justified on social benefits such as time saved by drivers (unpriced consumer surplus) whereas rail investment must be paid entirely by rail users.

Sensible Transportation Policy - Maine

The Sensible Transportation Policy was imposed on the Maine Department of Transportation by a referendum on the future expansion plans of the Maine Turnpike. This policy gives preference to alternatives to highway expansion. Faced with a policy that froze highway development, the Department convened stakeholders from trucking, railways, bicyclists, environmentalists, construction, economic development, and municipalities to agree on a regulation to implement the policy. After about 9 months there was unanimous agreement for a multi-stakeholder, multi-modal, multi-jurisdiction transportation planning process that operates both at the State and regional levels.

Intermodal Surface Transportation Efficiencies Act (ISTEA)

The philosophy of ISTEA parallels the multi-stakeholder, multi-modal, multi-jurisdiction philosophy of Maine s Sensible Transportation Policy. It does not go as far as the Swedish model to explicitly price external costs such as accidents and pollution, but surface transportation plans must consider the overall social, economic, energy, and environmental impacts. It also works in conjunction with the Clean Air Act Amendment (1990) to require transportation plans to reduce emissions in regions with poor air quality.

Challenges

The stated policy of the Canadian Government is to get out of the transportation business, except for safety regulation of some but not all modes, and to let the market decide transportation prices and investment. We now operate in a continental rather than a domestic transportation market. The market is distorted by public support to marine infrastructure in the U.S. which competes with Canadian export routes. It is distorted by subsidies to trucking and the Seaway. It is distorted by different taxes and safety regulations for different modes.

An optimum transport system would minimize total costs, including external costs, by reflecting the total costs in market prices. In reality, this will be tempered by conditions of the continental market. What needs to be done?

- Comparative analysis and harmonization of U.S. and Canadian transport legislation, regulations, policies, taxes, and subsidies.

- Develop an integrated transportation plan that is multi-modal, multi-jurisdiction, and multi-stakeholder. The goal would be enhancing the competitiveness of the Canadian system within a continental market while sustaining our standard of living and quality of life and environment.

- Work with our NAFTA partners to develop a continental transportation policy based on a level playing field for each mode among the countries and within each country.

Emissions From Light-Duty Vehicles in Hamilton-Wentworth Under Policy Scenarios: Applications of an Integrated Model

Pavlos S. Kanaroglou, William P. Anderson, and Robert South

Department of Geography and
McMaster Institute for Energy Studies
McMaster University

Background

This paper reports on simulations conducted by the authors as members of the Transportation and Land-Use Workgroup of the Hamilton-Wentworth Air Quality Initiative (HAQI). HAQI is a co-operative project of the Regional Municipality of Hamilton-Wentworth, the Ontario Ministry of Environment and Energy, The Ontario Ministry of Transportation, Environment Canada, public interest groups, and a number of researchers from McMaster University. Its objective is to assess air quality problems and the factors that underlie them in the Hamilton metropolitan area. The mandate of the Transportation and Land-Use Workgroup is to assemble information on the impact of transportation on air quality, explore the link between evolving land-use patterns and pollution from transportation, and envision likely future trends in transportation emissions.

As part of the committee's work, the authors conducted computer simulations of automobile emissions at ten-year intervals from 1991 to 2021 under a number of different policy scenarios. The scenarios, which were developed in consultation with the Workgroup, incorporate spatial projections of population, employment, household structure, and transportation infrastructure that were provided by the Region of Hamilton-Wentworth. They were executed using an integrated urban simulation model known as IMULATE (Integrated Model of Urban Land-use, Transportation, and Environmental analysis) that was developed by researchers at McMaster University and the University of Toronto.[1]

The pollutants for which emissions estimates are generated are

hydrocarbons (sometimes called volatile organic compounds), carbon monoxide, and nitrogen oxides. They were chosen for three reasons. First, they are all known to have air quality impacts: carbon monoxide has direct health impacts and hydrocarbons and nitrogen oxides are precursors to ozone. Second, automotive emissions make up a large proportion of the total emissions of these pollutants in most cities. Finally, they were chosen on pragmatic grounds because they can all be estimated using a well established emissions model. It should be stressed that the emissions estimates are for light duty vehicles (cars, vans, and small trucks) only. No attempt was made to estimate emissions from heavy duty vehicles (buses and trucks) or from non-transportation sources. Also, particulate emissions from transportation are not provided because the modelling technology needed to make such estimates was not sufficiently well-developed. Finally, the study deals with emissions only -- it does not attempt to estimate how changes in emissions translate into changes in ambient air quality.

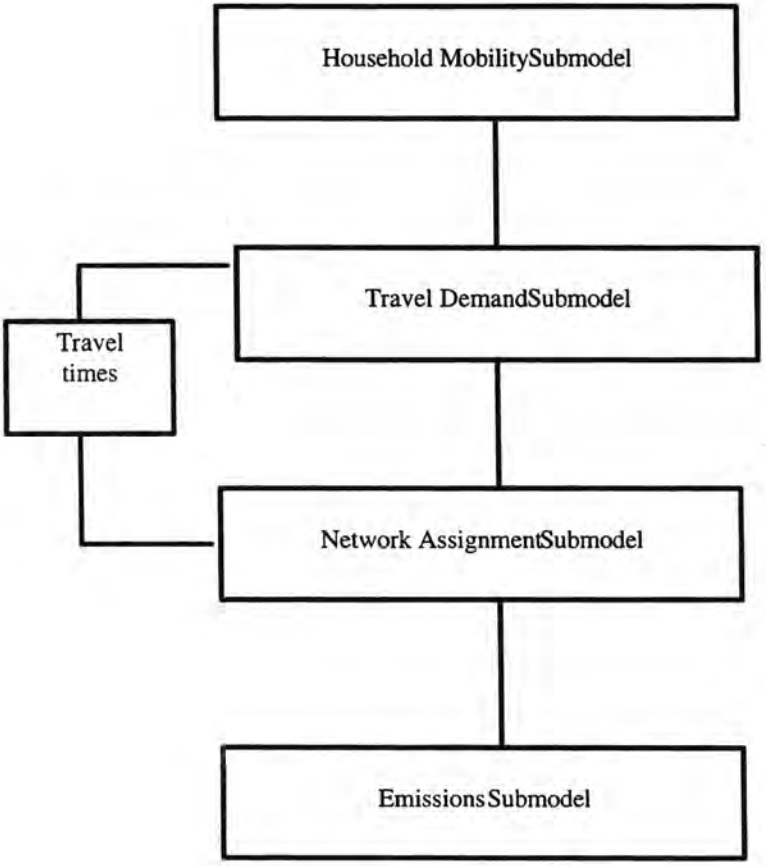
The procedure for estimating automotive emissions is as follows. IMULATE is used to simulate the flow of automobile traffic along the Hamilton-Wentworth road network in a typical morning rush hour. The model produces estimates of flow and congestion level on each link in the network. This information is passed to an emissions model to estimate emission of the three pollutants on a link-by-link basis. Summing across links provides region-wide estimates.

Emissions estimates were made for the years 1991, 2001, 2011, and 2021. The 1991 results provide our best estimates of contemporary emissions. Future emissions will vary depending upon developments in road infrastructure, vehicle technologies, public transportation ridership, and land use patterns. To reflect this, estimates for the years 2001, 2011, and 2021 were made for a number of scenarios comprising different assumptions about those developments.

IMULATE

The IMULATE model starts with three types of information: 1) information on dwellings, households, and business establishments in 151 zones of the Hamilton metropolitan area; 2) an initial "place of residence / place of work

Figure 1: IMULATE Model Structure


Household MobilitySubmodel
Travel DemandSubmodel
Travel
times
Network AssignmentSubmodel
Emissions Submodel

matrix which matches residents in each zone with their zones of employment; and 3) information on transportation options including a digital representation of the road network.

The model has four main components, as illustrated in Figure 1. The first is the Household Mobility submodel, which assigns population to zones and allocates all workers to zones of employment by means of a residential choice logit model system. This submodel also simulates changes in the number of households due to migration and household formation and dissolution.

The Travel Demand Submodel estimates the number of work, school, and discretionary trips occurring between each pair of zones. (Discretionary trips include shopping, recreation, etc.) Each category of trips is broken down into four modes -- auto driver, auto passenger, public transit, and walk/bicycle – by logit models based on relative travel time and cost and socio-economic characteristics of the population.

The Network Assignment Submodel translates the zone-to-zone automobile flows from the Travel Demand Submodel into flows along specific links in the road networks by means of a *user equilibrium* model. At the same time, it estimates the effect of congestion on the speed of travel along individual links based on the assigned traffic. In order to simulate traffic in a realistic way, it takes account of the fact that many travellers will chose indirect routes in order to avoid traffic. Based on the link speeds, this submodel calculates zone-to-zone travel times. (Since travel times are used to determine modal shares, there is a feedback of information from the Network Assignment Model to the Transportation Demand Submodel. See Figure 1.)

The Emissions Submodel takes information on flows and average speeds from the Network Assignment Submodel to estimate emissions of carbon monoxide, hydrocarbons, and nitrogen oxides on each link. Average link speeds are used as proxies for driving conditions, with low speeds representing stop-and go conditions experienced in urban traffic. At these speeds, emissions per kilometre travelled are considerably higher than under the more continuous driving cycle associated with higher speeds. Emissions of nitrogen oxides also increase at very high speeds. The technical parameters in this model are generated by MOBILE5.C which is the Canadian version of an

automobile emissions model that was developed by the U.S. Environmental Protection Agency.[2]

A typical application of IMULATE begins with a simulation of transportation patterns and emissions using the best available data for a base year. Projections for later years are conducted under "scenarios" that include assumed or expected changes in the distribution of households and workplaces, the extent and capacity of the transportation network, the relative costs of transportation options, consumer preferences, and the technical characteristics of the vehicle fleet. In the application described here, simulations were made for the base year 1991, and projections for 2001, 2011, and 2021 were made under six different scenarios.

The IMULATE data set for Hamilton is based on 151 zones, corresponding to census tracts, and a network of over 1500 links. Data defined at the zone level, such as population, socio-economic characteristics, and employment locations, were obtained from the Census and from surveys conducted by the Region of Hamilton-Wentworth. Zone-to-zone data on travel flows and mode splits are from the Transportation Tomorrow Survey, and network link characteristics were provided by the Joint Program in Transportation at the University of Toronto.

The model is run under six *scenarios*. Each scenario incorporates a set of suppositions about the future. Like any model, IMULATE incorporates many assumptions, and can only simulate behaviour that is consistent with what has been observed in the past. Since modelling is an imprecise science, and since human behaviour and preferences can change in ways that are impossible to anticipate, the model's outputs should not be taken as accurate predictions of the future. Rather, they are our best estimates of what might happen under certain sets of circumstances.

Scenarios

The six scenarios were designed to reflect possible future trends in road infrastructure, vehicle technology, transit ridership, and spatial development patterns. Figure 2 demonstrates the relationships among the scenarios. The nesting of scenarios in this diagrams indicates the degree to which assumptions are consistent from one scenario to another. For example,

the New Roads scenario preserves the basic assumption of the Base scenario but adds some additional assumptions. Both the Technological Improvement and 25% Transit scenarios preserve the main assumptions of the New Roads scenario and add some of their own. However, the new assumptions in the Technological Improvement scenario are not included in the 25% Transit scenario. The scenarios are designed in this way to make it easy to see the impacts of different changes in assumptions about the future of projected emissions.

The Base Scenario

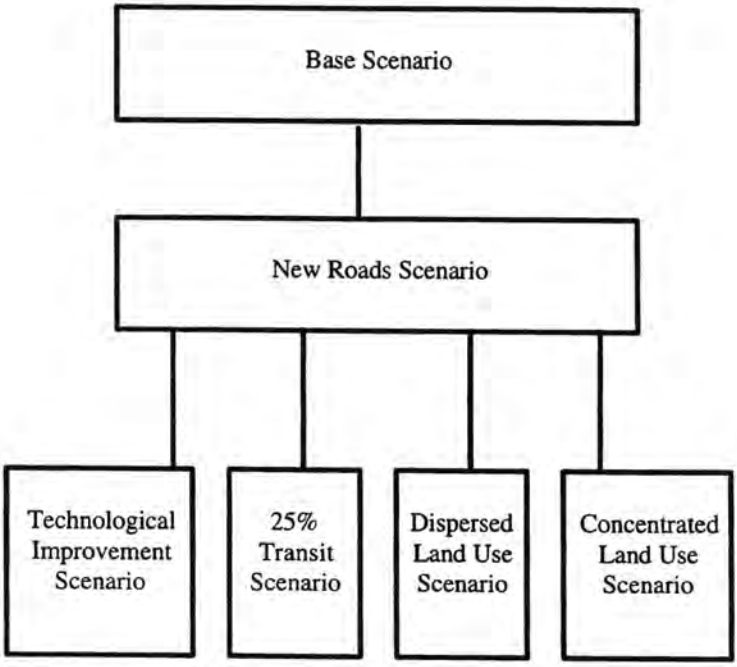
The first of the six scenarios is called the Base Scenario. Its purpose is to provide our best estimates of transportation activity and emissions levels in 1991, and "business as usual" projections for the later years. These future projections are not intended as "most likely" outcomes, since many factors are likely to change in the future. Rather, they serve as bases for comparison to see how changes in various factors are likely to affect emissions in the future. The 1991 input data is drawn from the sources described above. These data are updated for the later years based on zonal projections of population and employment provided by the Region. Adjustments were also made to reflect expected changes in household size and the age structure of the population.

The main trends in the Base and all other scenario runs are summarised in Tables 1 and 2 which shows the absolute values and percentage changes (respectively) in key variables over the entire H-W region between the 1991[3] and 2021 estimates. The increase in vehicle kilometres travelled (VKT) reflects two things: the increasing number of trips due to more people with more cars and the increasing average length of trips. The reduction in average speed indicates more congestion on the regional road network. An important result is that the emissions of all three pollutants increases by a greater percentage than does VKT. This is because of the increase in congestion which results in higher emissions rates. (This effect is less pronounced for Nitrogen Oxides.)

The New Roads Scenario

The Base Scenario assumes that the extent and capacity of the road network will remain unchanged between 1991 and 2021. A more realistic assumption is that some existing roads will be expanded and some new roads will be added. The New Roads Scenarios incorporates all improvements and

Figure : Scenario Structure



additions to the network that are currently envisioned by the region, including the Red Hill Creek Expressway and a bypass of Highway 5 around the town of Waterdown.

Changes in key variables between 1991 and 2021 (Table 2) under the New Roads Scenario show some interesting differences when compared with the Base Scenario. Growth in VKT is greater, but the increase in congestion (as indicated by the reduction in average speed) is smaller. Despite faster growth in driving, emissions grow more slowly than in the Base Scenario due to the strong positive link between emissions and congestion.

These results should be interpreted with some caution since they refer only to the morning rush hour when most trips are commuting trips. Discretionary trips which occur at other times of the day may be stimulated by the increase in road capacity and the reduction in congestion. Also, expanding the network may lead to more spread-out land use patterns. Both of these effects would lead to a greater increase in VKT. Therefore the emissions benefits of road expansion would probably be much smaller -- or possibly even negative -- if the analysis were done over a twenty-four hour period and reflected the impacts of infrastructure on land use.

Since the network changes in the New Roads Scenario are based on existing plans, the scenario is considered more likely than the base scenario in which the network will remain unchanged. For this reason, New Roads network assumptions are carried forward into all the remaining scenarios (see Figure 2.)

Technological Improvements Scenario

Both the Base and New Roads Scenarios assume that the technological characteristics of cars remain constant from 1991 to 2021. However federal legislation in both the U.S. and Canada will require significant improvements in vehicle emissions performance over this period. To account for this, the Technological Improvement Scenario makes reductions in the vehicle emissions factors over the period to represent the expected impacts of these improvements.

Since this scenario also includes the expanded road network, it has the same values for changes in VKT and average speed as the New Roads

Scenario. But there are striking differences in emissions. Instead of increases, significant reductions in emissions of all three pollutants are projected.

These results should also be treated with some caution. Recent research suggests that expected reductions in emissions are not always achieved due to in-service degradation in the performance of emissions control devices.[4] Inspection and maintenance procedures can improve this performance, but they are not currently required in Ontario. Thus, these results are best regarded as optimistic. Still the ability of technology to reduce emissions is beyond question.

The reductions in emissions due to technological improvements should apply to any future scenario. In the last two scenarios, however, the assumption that vehicle technology remains at 1991 levels is retained. This is done in order to make it easier to distinguish the independent effects of travel behaviour and land use changes from the effects of technological changes.

25% Transit Scenario

Public Transit has the capacity to reduce emissions by reducing the number of cars on the road. At present, about 11% of trips in the morning rush hour are made by transit. This scenario increases that value to 25% – a number more typical of European cities. (The 25% figure is aggregate. Different types of individuals living in different parts of the region are assumed to have different transit propensities.) Such an increase could probably not be achieved by minor reductions in fares or other incentives. Rather, it would take significant shift in commuters' attitudes toward public transportation. Still, it is not an impossible goal given the existing patterns of residences, workplaces, and transit routes.

The results in Tables 1 and 2 indicate that this change would cut the growth in VKT in half and lead to an increase in average speed due to reduced congestion. It would also greatly reduce the growth in emissions when compared to either the Base or New Roads Scenarios. These results confirm the important role of public transportation in addressing air quality problems. (Note, however, that the model does not estimate emissions from buses, so it does not take account of the increased emissions from transit vehicles. However, assuming reasonable load factors the net effect would certainly be a reduction in emissions growth.)

Dispersed Land Use Scenario

One of the main reasons for the rapid growth in automotive VKT is changes in land uses that lead to increased automobile dependency and longer commuting trips. To illustrate this, the Dispersed Land Use Scenario alters the projections of population growth in the Base Scenario to limit density in growing zones to 3500 households per square kilometre, which is roughly the current density of Westdale, an "old suburb" neighbourhood of Hamilton. (Zones that already had higher density in 1991 were assigned no growth in households, and zones with lower density in 1991 were assigned new households only up to the density limit.) This scenario also assumes the expanded road network.

This scenario produces a growth pattern that is more dispersed and is almost completely diverted away from the Lower City. The result is a greater increase in VKT and a slightly greater reduction in average speed than in the New Roads Scenario. Growth in emissions is also somewhat greater.

Concentrated Land Use Scenario

This scenario assumes a more concentrated pattern of growth. This is achieved by transferring all the growth assigned to the peripheral ring of zones in the regional projections to zones closer to the central city -- thus it simulates the establishment of an urban growth boundary. This scenario also assumes the expanded road network.

Not surprisingly, growth in VKT is about 20% lower in this scenario than in the New Roads scenario. What may be surprising, however, is that the higher concentration of population growth leads to slower growth in congestion when compared with the New Roads scenario. This is because less traffic is assigned to peripheral roads with low capacities and because with shorter average commutes, fewer vehicle kilometres are assigned to the network. Fewer VKT and less congestion naturally produce lower emissions.

While the results of the two land use scenarios indicate relatively modest impacts, it should be noted that they are based on relatively modest changes in development practices. More radical changes would have larger impacts. Also, many land use changes that are "transit supportive," such as concentrating residential development along transit corridors, occur at a smaller

level of resolution than the IMULATE model is currently able to simulate, so their impacts are not reflected in our results.

Conclusions

The results of these simulations are not intended to be accurate predictions of the future. Rather, they provide our best available indication of the relative importance of various factors in determining future emissions levels. Viewed in this way, they allow us to make a number of conclusions.

The first such conclusion is that congestion is a critical factor in emissions determination, so expansion of the road network will reduce future emissions as long as transportation flows are relatively stable. However, this result may not hold true if expanding the network has a positive impact on the volume of travel.

A second conclusion is that technological improvement, especially in the emissions control systems of vehicles, has the potential to reduce emissions significantly, even in the face of increased driving. This is not a reason for complacency, however. There is some question as to whether such improvements will reach their full potential. Even if they do, these improvements may be offset if there is sufficient growth in driving.

A very interesting conclusion is that an increase in public transit ridership could substantially offset projected increases in driving and congestion. A similar increase in the use of non-motorised modes such as walking and cycling could make these benefits even greater. Changes in land use produce measurable, but relatively small, changes in our simulations. Further simulations based on more detailed land use scenarios are needed, however, before this option can be dismissed.

Table 1: Key Variables for Morning Rush Hour Period, 1991 and Six scenarios for 2021

Scenario	Vehicle Kilometres Travelled	Average Congested Speed	Emissions of VOC's (kilograms)	Emissions of Nitrogen Oxides (kg)	Emissions of Carbon Monoxide
1991 Base	797085	33.28	7480	6577	80983
2021 Base	972619	30.78	12136	8611	132527
2021 New Roads	995324	32.68	10134	8308	111022
2021 Technological Improvements	995324	32.68	5119	4759	58052
2021 25% Transit Scenario	891151	33.71	7982	7309	87573
2021 Dispersed Land Use	1016085	32.66	10157	8324	111242
2021 Concentrated Land Use	961180	32.75	9952	8199	108983

Table 2: Percentage Change in Key Variables for Morning Rush Hour Period, 1991 - 2021

Scenario	Vehicle Kilometres Travelled	Average Congested Speed	Emissions of Hydrocarbons	Emissions of Nitrogen Oxides	Emissions of Carbon Monoxide
Base	22.0	-7.5	62.2	30.9	63.6
New Roads	24.9	-1.8	35.5	26.3	37.1
Technological Improvements	24.9	-1.8	-31.6	-27.6	-28.3
25% Transit Scenario	11.8	1.3	6.7	11.1	8.1
Dispersed Land Use	27.5	-1.9	35.7	26.6	37.3
Concentrated Land Use	20.6	-1.6	33.0	24.7	34.6

* Change over Base Scenario 1991

[1] The development team included Anderson and Kanaroglou at McMaster, and Eric Miller of the Department of Civil Engineering at the University of Toronto.

[2] Additional technical details are furnished in William P. Anderson, Pavlos S. Kanaroglou, RonaldBuliung, and Eric J. Miller, (1996) "Estimating Urban Emissions in an Integrated Model"*Transportation Research Record* in press.

[3]1991 is chosen as the starting year because it is the most recent year for which census data are available. It was assumed that trip making and mode choice rates observed in the Transportation Tomorrow Survey did not change between 1986 and 1991.

[4] See M. Ross et al (1995) *Real World Emissions from Model Year 1993, 2000, and 2010 Passenger Cars*(Washington DC: American Council for an Energy Efficient Economy.)

THE RELATIONSHIP OF FUEL CONSUMPTION AND ROAD SURFACE QUALITY

by Jerry Fruin and Dan Halbach*

University of Minnesota
Department of Applied Economics
University of Minnesota
1994 Buford Avenue
St. Paul, MN 55108-6040
612-625-8720
Fax 612-625-2729
jfruin@dept.agecon.umn.edu
dhalbach@dept.agecon.umn.edu

THE RELATIONSHIP OF FUEL CONSUMPTION AND ROAD SURFACE QUALITY

by Jerry Fruin and Dan Halbach*

ABSTRACT

The purpose of this article is to determine the difference in fuel consumption of a vehicle due to road surface type and quality. Rural road sections in three Minnesota counties (Anoka, Dakota and Chisago) were studied. The road surfaces included new bituminous, gravel surfaces of varying condition and aggregate composition, and one dirt surface. Photos and video of the test sections help to identify and record road conditions. Test sections did have measurable differences in energy requirements.

Key words: surface, energy, gravel, dynamometer, road

INTRODUCTION

An adequate road system is essential for the economic and social well-being of rural people. Although rural traffic is not heavy by urban standards, travel on rural roads is a necessity. This necessary travel takes place on diverse surface-type road networks. The development of the automobiles and truck industries during the 1920s and 1930s created the need to get rural areas "out of the mud." By the 1950s a majority of the local roads were improved with all-weather gravel or paved surfaces. Widths, grades, bases, surface designs and capacities of many of today's local rural roads are based on traffic needs of the 1940s and 1950s. Roads that were adequate in their time often do not meet standard of today's uses due to technological advancement in vehicles. Based on the importance of the rural roads system the Minnesota Local Road Research Board and the University of Minnesota Center for Transportation Studies have supported a program to determine the best way to apply limited funds to rural road infrastructure problems.

BACKGROUND

Accurate knowledge of increased cost due to the quality of road surface is very important if we are to determine the most cost effective methods and procedures for deciding when, where, and to what extent rural roads should be maintained or upgraded. The existing information often used by various studies including Baumel et al.[1] is based on data from Robley Winfrey's[2] work at the Iowa Engineering Experiment Station in the 1920s and 30s.[3] More recent World Bank studies of roads in developing countries are not comparable to U.S. rural conditions because all the roads are poor to extremely poor by U.S. standards. In addition, traffic levels are higher and include extremely heavy trucks on earth and gravel surfaces that are unknown in the U.S. Current vehicles (and rural roads) in the U.S. are not comparable to those that existed when the Winfrey studies were done. The authors' previous work[4] on a computer simulation of a 1500 square kilometer area with three categories of traffic (agricultural marketing, personal travel, overhead) made clear that surface type and the associated variable cost assumptions are important. Variable costs are those associated with kilometers driven. Vehicle operating costs included fuel, oil, maintenance and other expenses.

Working with the Minnesota Local Road Research Board (comprised of county highway engineers) a project was developed to measure fuel consumption using replicated field tests over measured courses. It was determined that the fuel component of vehicle operating cost would be the best variable to try to measure in field tests on gravel and earth surfaces as opposed to laboratory tests and highway metered driving. While attempting to design such a study, we discovered recent developments in on-board dynamometers. While developed and used for racing applications, we proposed to evaluate and use this technology to determine actual road surface energy relationships. The resulting cost estimates and methodology suing these new technologies should be useful to researchers and officials evaluating alternative investment, disinvestment, and maintenance strategies for funding our rural road system.

A DYANAFACT© on-board dynamometer and computer developed by Banks Technologies of Azusa, CA was installed in a 1994 GMC ½ ton, 4 wheel drive extended cab pickup truck equipped with a 5.7 liter engine and automatic transmission. Baseline chassis dynamometer

laboratory runs were done to determine rear wheel horsepower output to check and calibrate with on-board dynamometer. the DYANAFACT© system computers horsepower based on a vehicle's velocity (kph from a calibrated speed sensor), acceleration (g forces monitored by means of an accelerometer), mass (vehicle weight) and vehicle drag characteristics. Given these parameters, a microprocessor can mathematically calculate horsepower. Data logging to a portable computer twenty times per second records: time stamp, horsepower, g force, mph, and distance traveled.

Both the aerodynamic and frictional drag horsepower acting on the vehicle can now be measured. Aerodynamic drag is the amount of horsepower for a vehicle to move through the air. Frictional drag is the amount of horsepower required to overcome the combination of mechanical losses through the tires, gear bearings, etc.

The road surfaces included new bituminous, gravel surfaces of varying condition and aggregate composition, and one earth surface. All sites were at least three kilometers in length and in general very flat to reduce the influence of elevation changes. The paved site was in Anoka County at 149th Avenue near the city of Ham Lake. The three gravel sites were in Chisago County. The highest quality segment was on Chisago Co. 70. This is a wide, well-maintained surface of very good gravel, a "picture perfect" gravel road. Chisago Co. 69 site has "less gravel" and covers two distinct underlying soil types. The southern section is found in an area with peat soils and sod farms while the northern half is on sand and gravel. The last gravel site was hard packed and had very little gravel on the surface on 310th Street near Taylors Falls. The earth surface site located at 190th Avenue in Dakota County is a posted minimum maintenance road signed with "travel at your own risk."

A number of calibration parameters were carried out and set before each set of runs on a daily basis. Vehicle weight was checked with four tire scales (Intercomp© load cells) to check distribution and adjust ballast to maintain consistent weight distribution and total weight. Tire pressure was set at 50 psi cold. Vehicle wheel alignment (caster, camber, toe in) was set and monitored so changes did not occur when recording data. A driver and equipment operator made at least five (two directional) runs to calibrate the DYANAFACT© for both frictional drag and

aerodynamic drag at each site on each day. Data was collected only on low wind speed days on dry surfaces in fair weather. No correction for temperature or barometric pressure changes were made. Speed was held constant with the vehicles electronic speed control. Data was logged in both directions at various speeds over a distance of 1.6 kilometers.

RESULTS

Fuel consumption is approximately proportional to horsepower expended. Horsepower is a measure of *power*, which has a technical definition in physics. It measures the amount of *work* that a force does as it acts over time. Work measures the actual effect of a force in moving an object over a distance. If you move an object one foot by applying a force of one pound, you are doing one foot-pound of work. If it takes one second to move the object, you have exerted one foot-pound per second of power. A horsepower is 550 foot-pounds per second. Horsepower-hour can be converted to Btu's by multiplying by 2544.4336. Gasoline varies but is considered to have a Btu content of about 125,000 per gallon so with assumptions about mechanical/engine efficiency factors, fuel consumption can be estimated.

Table 1 reports the horsepower required from speeds of 42 to 108 kph over the various surfaces at a tire pressure of 50 psi. Note that both gravel surfaces require substantially more horsepower than new pavement and earth.

Table 1
Horsepower at Various Speeds by Surface

KPH	MPH	New Paved	Earth	Excellent Gravel	Good Gravel	Fair Gravel
42	25	4.91	8.09	8.06	8.30	7.06
50	30	6.90	10.58	10.89	11.19	9.61
58	35	9.40	13.49	14.33	14.91	12.74
66	40	12.48	16.89	18.49	19.31	16.53
75	45	16.20	20.84	23.45	24.57	21.06
83	50	20.65	25.39	29.30	30.79	26.42
91	55	25.90	30.61	36.13	38.06	32.69
100	60	32.01	36.56	44.04	46.49	39.96
108	65	39.00	43.29	53.12	56.17	48.31

Table 2 reports relative fuel consumption with consumption on new pavement equal to 100. Assuming a 20 mpg fuel consumption rate on new bituminous at 55 mph, we estimated earth surface would obtain 16.9 mpg, and excellent gravel, 14.3 mpg.

Table 2

Relative Fuel Consumption in Relation to Surface

KPH	MPH	New Paved	Earth	Excellent Gravel	Good Gravel	Fair Gravel
42	25	100.0	164.8	164.1	169.0	143.8
50	30	100.0	153.3	157.8	162.2	139.3
58	35	100.0	143.5	152.5	158.6	135.5
66	40	100.0	135.3	148.1	154.7	132.4
75	45	100.0	128.6	144.7	151.7	130.0
83	50	100.0	123.0	141.9	149.9	127.9
91	55	100.0	118.2	139.5	146.9	126.2
100	60	100.0	114.2	137.6	145.2	124.8
108	65	100.0	111.0	136.2	144.0	123.9

During a return trip to the earth-surfaced road for additional runs, it was discovered that a surface change had occurred. A thin layer of ground bituminous from a county road improvement project had been applied. Table 3 shows the effect of this surface change. Horsepower requirements increase from over 3 percent at slow speeds to over 10 percent at higher speeds.

Table 3

HP Effect of Gravel Change on Earth Surface

KPH	MPH	Earth	After Improvement	% Change
42	25	8.09	8.36	3.3
50	30	10.58	11.48	8.5
58	35	13.49	14.24	5.6
66	40	16.89	18.01	6.6
75	45	20.84	22.46	7.8
83	50	25.39	27.58	8.6
91	55	30.61	33.52	9.5
100	60	36.56	40.34	10.3
108	65	43.29	NA	NA

The relative impact of different vehicle factors was also of interest. Table 4 reports the effects of lowering the tire pressure from 50 psi to 35 psi. This change was large enough to make the vehicle "feel, ride or handle" differently.

Table 4

Change in Required Horsepower from Tire Pressure Change on Gravel Surfaces

KPH	MPH	Best Gravel			Good Gravel		
		50 psi	35 psi	% Change	50 psi	35 psi	% Change
42	25	8.06	8.83	9.6	8.30	9.09	9.5
50	30	10.89	11.88	9.1	11.19	12.23	9.3
58	35	14.33	15.43	7.7	14.91	16.04	7.6
66	40	18.49	19.76	6.9	19.31	20.64	6.9
75	45	23.45	25.04	6.8	24.57	26.12	6.3
83	50	29.30	30.97	5.7	30.79	32.57	5.8
91	55	36.13	38.11	5.5	38.06	40.11	5.4
100	60	44.04	46.26	5.0	46.49	48.83	5.0
108	65	53.12	55.48	4.4	56.17	58.60	4.3

The impact of additional vehicle weight is reported in Table 5. This was done on County 69 or the "good" gravel section. A two-axle trailer with electric brakes was loaded at two weight levels. The number of runs was limited due to safety considerations. Runs were made at 80 psi, the maximum loading capacity of the tires.

Table 5

Horsepower Requirements of Various Weight on a Gravel Surface

Speed		Weight		
KPH	MPH	2449 kg	3900 kg	5262 kg
50	30	10.90	24.34	31.70
66	40	19.40	41.70	59.10
83	50	31.60	64.90	82.80
100	60	47.92	92.50	NA

OBSERVATIONS

Fuel consumption and other operating costs can not be the primary consideration for rural road infrastructure decisions. Both the earth and poor gravel surface had safety problems at moderate speeds and could not be considered "all weather" roads. Other operating costs, while not as measurable as fuel use, are impacted by surface quality. For example, the vehicle used for this study needed alignment after every set of runs to stay in specification. The gravel dust dramatically decreased the service life of air filter and motor oil. The percent of total miles driven on poor quality surfaces will vary by region and locality. In our test area only a small percentage of the total miles would have occurred on these types of roads. When possible, drivers go quickly to hard surfaced roads.

CONCLUSIONS

Significant differences in fuel consumption exist by surface type. Major components of vehicle operating costs can and need to be accurately measured. Technology continues to improve and should be applied in innovative ways to road and infrastructure issues.

ACKNOWLEDGMENTS

This research is supported by funding from the Minnesota Local Road Research Board, Minnesota Department of Transportation, Center for Transportation Studies at the University of Minnesota, Minnesota Agricultural Experiment Station.

REFERENCES

1. Baumel, C.P., S. Hanson, C. Hamlett, G. Pautsch. "Vehicle Travel Cost on Paved, Granular and Earth Surfaced County Roads." *Proceedings of the Transportation Research Forum Twenty-sixth Annual Meeting,* Washington, DC 1985, pp. 401-406.
2. Winfrey, R. *Economic Analysis for Highways.* Scranton, PA: International Textbook Company, 1969.
3. Winfrey, R. "Automobile Operating Cost and Mileage Studies, " Iowa State College of Agriculture and Mechanic Arts, Official Publication, Vol. 30, No. 8, 1931.
4. Fruin, J and D. Halbach. "Rural Roads, Investment and Disinvestment in a Minnesota County." University of Minnesota, Department of Applied Economics, Staff Paper P92-25, St. Paul, MN, 1992.
5. Lawlor, J. *Auto Math Handbook: Mathematical Calculations, Theory, and Formulas for the Automotive Enthusiast.* Los Angeles, CA: HP Books, 1992.

ENDNOTE

*Jerry Fruin is Associate Professor of Marketing and Transportation in the College of Agricultural, Food and Environmental Sciences at the University of Minnesota. Daniel Halbach is a Research Fellow in the Department of Applied Economics at the University of Minnesota.

Overblown: Examining the Costs and Benefits of Automotive Air Bags

W.G.Waters II
The University of British Columbia

One of the major safety innovations on automobiles in recent years are "air bags," rapidly-inflated air cushions which deploy in a collision to help reduce injury or prevent death. Air bags are credited with saving a number of lives. "More than 1700 people are alive today because of their airbags." (Insurance Institute..., 1997). However, air bags are not a panacea for accident/injury reduction. They are effective primarily for frontal collisions (although side impact air bags are also starting to emerge). Air bags are intended to be a "supplementary restraint system" (SRS) and not an alternative to seat-belts.[1] However, controversy is emerging about some technical characteristics of current air bag technology. North American air bag technology is dominated by United States' regulatory requirements, which in turn reflect U.S. test criteria. U.S. drivers are less likely than Canadians to use seat-belts; hence U.S. air bag testing requirements (using "crash test dummies") are designed for non-belted car occupants, specifically the fiftieth-percentile-size male driver. The air bags must deploy at relatively low impact speeds *and* inflate very rapidly. These aggressive air bags have become the source of injury and even deaths. "Sixty one deaths have been reportedly have been caused by airbags inflating in low severity crashes." (Insurance Institute..., 1997). Many were infants in rear-facing infant carriers or unbelted young children in the front passenger seat. Particularly for Canada with our high use of seat-belts, there is at least some question about whether or not current air bag technology is cost-effective.

This paper reviews the issues and some of the available evidence regarding the costs and effectiveness of automotive air bags and the

---

[1] The term "seat-belts" is understood here to include shoulder restraints as well as the lap belt.

implications for Canada. It begins with a brief summary of how air bags work. Part II summarizes evidence on the effectiveness of air bags. Part III addresses the costs and benefits of air bags; despite the wide publicity of lives saved, doubts are raised about the overall net benefits of air bags. Some research needs are identified and brief conclusions follow.

1.0. Automotive Air Bags

An air bag consists of the air bag itself, folded in a module on the steering wheel or dashboard; a propellant (usually nitrogen released by sodium azide pellets despite some toxicity) to inflate the bag; and a sensing mechanism to detect sharp acceleration or deceleration which triggers the firing mechanism. (See Breed, 1993, for an explanation of how air bags function).

What Takes Place during a Collision:

By way of illustration, at a speed of 48 kph (30 mph), the driver and passengers continue to move forward at 48 kph at the instant of impact. If the car stopped instantly, it would take the driver about .12 seconds to hit the steering wheel (Evans, 1991, p.221); similarly for the front passenger. The energy-absorbing design of the automobile, progressively collapsing upon impact, adds a bit of time (for a total of about .15 seconds, *Ibid.*) before the driver and passengers hit the steering wheel or dashboard, respectively (the collapsing car also absorbs some of the energy on behalf of the driver and passengers but this effect may be partially offset by the vehicle re-bounding). Seat-belts restrain the driver and passenger, absorbing some of the sudden deceleration, and reduce injury from hitting the steering wheel or dashboard. If the driver (or passenger) is not belted, the design criterion is that the air bag must inflate before the driver travels more than 13 cms (Breed, 1993, p.S70), which would take about .05 seconds. The air bag can inflate in about .03 seconds hence the sensor mechanism must function in less than .02 seconds to deploy the air bag (*Ibid.*)[2] All this happens, literally, quicker than the blink of an eye. Figures

---

[2] The sensor mechanism has a complex task because early in a collision the passenger compartment is not yet decelerating due to the impact. Hence most sensor mechanisms include sensors located near the front of the car which detect the crash impact first, supplemented by a more sensitive motion-change sensor in

1 and 2 illustrate the air bag deployment sequence and initial position of the driver.

Potential dangers of air bags

The air bag needs to inflate very rapidly hence it is potentially very powerful. A second and partly related consideration is the threshold speed of impact required to trigger the bag. As noted in the introduction, the present air bag technology conforming to U.S. regulations triggers at a relatively low threshold, many as low as 12 to 15 kph, because unbelted drivers could be injured at this speed. But the power of the inflating air bag can be a source of injury itself. The outer edge of the air bag is inflating at a speed of about 300 kph. If a person is hit by the air bag during the inflation process, substantial injury and even death can result. If drivers or passengers are out of position, or sit closer to the steering wheel/dashboard (which is the

FIGURE 1
Air bag deployment sequence

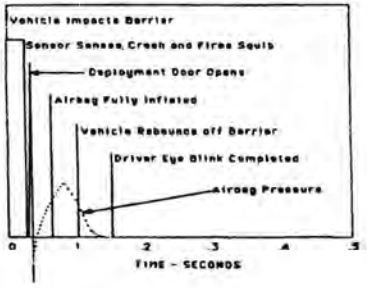


FIGURE 2



---

the passenger compartment. Both sensors must detect the crash for the air bag to deploy.

case for many females and smaller-stature individuals), the air bag deployment is potentially harmful.[3] In time, "smart bags" may be developed which could sense the nearness of the driver/passenger and avoid deploying the bag or alter the deployment rate as required, but these are not available at present. As a result, there are many documented instances of harm and even death resulting from air bags (54 deaths attributed to air bags in the U.S. and three in Canada, reported in the *Vancouver Sun*, November 28, 1996; the Insurance Institute for Highway Safety, 1997, reports a total of 61 deaths attributed to air bags deploying in low-severity crashes). This issue of threshold speed to trigger the air bag is raised later after reviewing the general evidence on the effectiveness of air bags at reducing fatalities and injuries.

2.0 Air Bag Effectiveness

There is no doubt that air bag technology can save lives and reduce injuries, although there are different estimates of its effectiveness. Early predictions were based on crash test simulations and subjective assessments of the survivability of different types of accidents. Now that air bags have been in use for some years, one can compare actual experience from accidents with and without air bags. The latter might seem the most useful guide but there are many variables to control for and biases in the data collected from actual crashes. There are known sources of error in recording data (e.g., seat-belt use is over stated in non-fatal accidents) (see Evans, 1991, Chapter 9 for a discussion). Crashes do not provide a random sample of drivers and vehicles.

> "For example, air bags originally were offered in the larger, more expensive and the sporty, high-performance model lines. These vehicles may be driven much differently than, for example, station wagons or family-size sedans.... some make/models will experience more single-vehicle crashes on rural, higher speed roads, which tend to be more severe than the average two-vehicle crash in an urban, lower speed environment.... heavier cars, on average, offer greater protection to their occupants than do lighter

---

[3] "Both the U.S. and Canadian data agree that female drivers are at higher risk of sustaining bag-induced injuries." Dalmotas (1996), p.7.

> cars (which were generally not included in the early introduction of air bags)."
>
> NHTSA (1996) p.7-8.

Certain groups are more crash-prone than others. For example, "There is copious evidence that belt wearers are more careful drivers than nonwearers [citations follow]" (Evans, 1991, p.226).

Nonetheless, analyzing actual crash results is a valuable data source for analyzing the effectiveness of air bags. "As of July 1996, NHTSA's Fatal Accident Reporting System (FARS) contained records of nearly 10,000 fatally injured front-seat occupants of cars and light trucks ... equipped with air bags." (NHTSA, 1996, p.9). This enables examination of several driver/passenger characteristics (age, size) and fatality rates with and without seat belt use and air bags.

Air bags and fatality rates.

There is no doubt that air bags can and do reduce fatalities. Air bags are most effective for direct, frontal collisions; they are less effective for "partially frontal" crashes, and generally assumed to have little or no benefit for side collisions or roll-overs. In a study of crashes with and without air bags including situations where the driver had air bags but front passengers did or did not have them, the data show a 34 percent reduction in the fatality rate for head-on collisions compared to no air bags present. This drops to an 18 percent reduction for all frontal collisions, and a 10 percent reduction for all crashes (NHSTA, 1996, p.11). However, these statistics include both those using and not using seat belts. Examining crashes where seat belts were in use, the NHSTA data still show a 21 percent improvement in fatality rates for belted drivers in full frontal ('head-on') crashes. That is, air bags help both belted and unbelted drivers in head-on collisions.

The NHSTA results are similar to their own previous analysis and those found by other researchers. Examining crashes for air-bag-equipped cars, 1985-91, O'Neill and Lund (1993) report a substantial reduction in fatalities for air-bag-equipped cars (29 percent reduction compared to seat-belt-equipped cars for frontal collisions). In a similar study incorporating additional years' data, Lund and Ferguson (1995) report a 24 percent

reduction in fatalities in frontal crashes, and 16 percent reduction in all crashes, for air bags compared to seat-belt only automobiles.

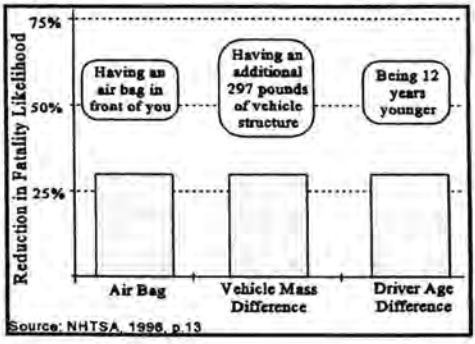
Whereas air bags offer protection primarily for frontal collisions, seat belts offer protection in all manner of crashes. **Table 1** summarizes NHTSA's estimates of the effectiveness of air bags, seat belts and their combination averaged over all accident types. These are "when used" estimates, i.e., these are the probabilities assuming that seat belts are actually used. In practice, many drivers do not use seat belts which makes air bags more valuable than the figures in Table 1. The combination of belts and air bags amounts to a 50 percent reduction in fatality rates compared to 45 percent reduction obtained by proper use of seat belts (including shoulder strap). The findings in Table 1 correspond closely to the figures used by Lawson (1993) in his analysis of whether or not Canada should adopt mandatory air bag requirements (he recommended against it; more later).

Table 1
Estimated Effectiveness of Occupant Protection Systems in Reducing Fatality Risk for Passenger Car Drivers

System Used	Fatality Reduction
Air bag alone	13%
Manual lap-shoulder belt	45%
Air bag plus shoulder belt	50%

source: NHTSA (1996) p.14.

The NHTSA data permit analysis of some other driver and vehicle characteristics including age of driver/passenger and size of vehicle. Elderly drivers have higher fatality rates, which is not too surprising. More worrisome is that: "Every analysis that includes frontal crashes shows a higher fatality risk for the children in cars with dual air bags than for children in comparable cars without air bags." (NHTSA, 1996, p.15-16.) Larger, heavier vehicles are also associated with lower fatality rates, i.e.,



the larger mass absorbs more of the energy from the collision. The forgoing is summarized in **Figure 3**. Particularly striking is the significance of vehicle weight, although one must be cautious about drawing conclusions from data with a high variance.

Air bags and injury rates.

Usually, safety features which reduce fatalities also reduce injuries; more specifically, there is a cascading effect of former fatalities who now are in "critical" or "severe" injured category, but correspondingly more people previously injured "severely" are reduced to a "moderate" category, with corresponding displacement and reductions to the "minor" injury category.

> "Historically, if an occupant protection system was effective in reducing the risk of fatal injury, it was also effective in preventing serious non-fatal injuries. The present findings indicate that this is not likely to prove true in the case of air bag systems, at least as presently designed."
>
> Dalmotas (1996) p.7

The NHTSA (1996) report on their examination of the NASS (National Accident Sampling System) data on crashes of all types, using the standard Abbreviated Injury Scale (AIS) rating of the severity of injuries (AIS 6 is an injury so severe it cannot be treated, down to AIS 1 or "minor" injuries; note that fatalities are included in this classification and can occur with injury levels 4 and 5 as well as 6). NHTSA examined the effect of air bags with and without belts for two injury categories, "moderate" or worse (AIS 2+) and "serious" or worse (AIS 3+). They also examined the data on specific types of injury: head, chest, upper and lower extremities, as well as by age and sex of the driver. They included car crashes with frontal damage only as well as all damage areas (the former are more relevant for air bags). Only some results are summarized here.

Table 2 shows the percentage reduction in injury probability for three occupant restraint alternatives based on car damage areas (columns 2 and 3) and for specific bodily-injury area (right four columns), for "moderate" and greater injury (AIS 2+) and "serious" or greater injury (AIS 3+).

Table 2
Estimated Effectiveness of Occupant Protection Systems
(percentage reduction in injury probability relative to no protection)

	vehicle damage		injury area			
System used: **Moderate injury+**	all damage areas	front damage	head	chest	upper	lower
air bag + belt	60	61	83	59	45	37
air bag alone	18	6	46	-31	25	21
manual lap-shoulder belt	49	56	59	59	14	53
Serious injury +						
air bag + belt	59	69	75	66	-40	78
air bag alone	7	- 8	16	18	-14	- 5
manual lap-shoulder belt	60	74	38	54	28	79

source: NHTSA (1996), pp.17-20

The air bag plus seat belts show slightly greater injury prevention than seat belts alone, but the air bag alone is much inferior to seat belts. Of particular interest are the figures for frontal damage accidents; these are ones where air bags are more likely to have deployed as these are the type of accidents they are designed for. The air bag alone shows only a 6.0 percent likelihood of preventing this injury level. The results are more striking for the "serious" injury accidents (bottom three rows of columns 2 and 3). Seat belts and air bags combined are not better and possibly worse than seat belts alone, as indicated by the negative impact of air bags alone. Although this is not significantly different from zero, it does reveal a minimal if not adverse effect of air bags on injuries.

The rightmost four columns in Table 2 show the impact of air bags and seat belts on specific types of injuries. Air bags perform well at reducing head injuries, and substantially improve the performance compared to seat belts alone. They have some effect on chest injuries, but keep in mind that seat belts are likely to 'cause' some "moderate" chest injuries because they effectively substitute these injuries for more severe ones sustained if the driver is not restrained. But for injuries to extremities, air bags have a low and even negative effect, i.e., aggravating injuries particularly for upper extremities.

Dalmotas (1996) has also examined the NASS data base focusing particularly on the impact of air bags on lower level injury categories (AIS 1, AIS 2 and AIS 3, "minor," "moderate" and "serious," respectively), including a calculated "harm index" which weights various injury categories according to a dollar valuation relative to the AIS 6 category (essentially equivalent to a fatality). The paper focuses on frontal or partial frontal collisions. The total sample of crashes was 7,336. This included 2600 cases of unbelted occupants without air bags; 3683 cases of seat belts used without air bags; 276 cases of unbelted drivers but with air bags; and 777 incidences where both air bags and belts were used. The data were analyzed in various ways, e.g., excluding cases where vehicle speeds were not known (the "delta v" or estimated impact speed, an index of accident severity), adjusting for delta v's to correct for potential distortions (e.g., unbelted drivers are over-represented in high speed crashes). We only report the initial table of results here as **Table 3**.

Dalmotas' analysis shows that the presence of air bags generally aggravates rather than reduces injury probabilities for "minor" to "serious" categories of injuries. For example, over all crashes in the sample, the probability of an unbelted driver without air bags receiving minor injuries or worse is 69.8 percent. An air bag without wearing belts increases the probability of injury to 75.6 percent, whereas use of seat belts without air bags reduces the probability of injury to 46.4 percent. Note that for all three injury categories, air bags are associated with higher injury rates on average. An analysis by specific injury type (head, upper extremity, etc.) confirmed that air bags have some positive effects besides preventing fatalities in frontal collisions:

> "The present findings suggest that these fatality reductions are being achieved through the decreased likelihood of AIS $\geq$3 head injuries in the case of the belted and the decreased likelihood of AIS $\geq$3 chest/abdomen injuries in the case of the unbelted. The findings ... also suggest that these reductions are likely to be accompanied by a substantial increase in the overall incidence of non-fatal serious injuries, particularly among belted occupants." (Dalmotas, 1996, p.7).

> "The problems associated with overly aggressive air bags are compounded by deployment practices that, in low-speed collisions, can expose a belted occupant to a greater risk of injury from the air bag than from the collision itself.... In jurisdictions such as

Canada, that have achieved seat belt use rates approaching 95%, it is counterproductive to seek to increase the safety of the unbelted occupant at the expense of the belted." (*Ibid.*)

TABLE 3
Injury Probability with and without Air Bags and Seat Belts

Injury Probability	no air bag, belts not worn	air bag, belts not worn	no air bag, belts worn	air bags and belts worn
AIS >= 1	69.80%	75.60%	46.36%	46.04%
AIS >= 2	19.26%	19.70%	7.36%	7.65%
AIS >= 3	5.33%	8.17%	1.62%	2.90%

source: Dalmotas (1996) Table 1A.

In sum, the crash data reveal safety tradeoffs accompanying air bag requirements. North American cars employ air bags which meet U.S. requirements. These are designed to protect the average unbelted male driver. Air bags must deploy at relatively low speeds (many as low as 12-15 kph) and deploy at a very aggressive rate. But these air bags are causing some deaths and a rise in injuries compared to using seat belts. These increases in injury rates associated with current air bag technology must be weighed against reductions in fatalities and/or "severe" and "critical" injuries brought about by air bags.

3.0 Considering Costs and Benefits of Air Bags

The foregoing concentrate on estimating the effectiveness of air bags at reducing fatalities and their effect on injury rates. But the fact that some lives are saved is not a sufficient economic justification for the adoption of air bags. A crude calculation of costs and benefits is to compare the estimated number of fatalities avoided thus far, times a value of life saved, and compare this with the expenditures on air bags over the same period.

"More than 1700 people are alive today because of their air bags." (Insurance Institute, 1997). Using a value of life of $2.0 million (and ignoring discounting), this is an economic benefit measure of $3.4 billion.

"About 55 million ... cars now on U.S. roads have driver airbags. About 28 million also have passenger airbags." (Insurance Institute..., 1997). Because the proportion of airbag-equipped cars is rising, this figure will overstate the costs to compare with the number lives saved. Nonetheless, multiplying the 55 million cars times $375 average cost per car (Lawson, 1993 used $500 as a resource cost of two air bags, and about half of the existing cars have two air bags),[4] equals $20.6 billion. This crude calculation shows a benefit cost ratio of .165. The overstatement of costs because the now-expanded total fleet cost is used could be corrected by using only half of the fleet (the average between close to zero in 1987 to the present 55 million cars). This doubles the benefit cost ratio but it is still only .33. This calculation is crude and based on fatalities only, but it serves notice that more careful examination of benefits and costs of air bags are called for.

One indicator of the valuation of safety benefits is consumer acceptance of air bags. Many people have demonstrated a willingness to pay for air bags as optional equipment. Assuming that buyers are well-informed about the safety benefits (a big assumption but one commonly made in other market analyses), this reveals that many people value air bags highly. An analysis yields an implied value of life of $3.0 million (Mannering and Winston, 1995). But not everyone purchases air bags; this leaves a large segment of the population who apparently value it less highly. Analysis of the market for air bags is greatly complicated by (U.S.) regulations which are requiring air bags as standard equipment, i.e., eliminating the consumer choice. (However, Mannering and Winston, 1995, argue that consumer preferences would have led to widespread use of air bags even without the regulatory requirements).

But safety decisions usually are not left to individual valuation. Buyers might not be well informed about safety features, and there are externalities (careless drivers impose risks on others), hence government intervenes on safety.

---

[4] This probably is a low estimate of air bag costs. Manufacturers differ in the price for air bags where they are optional equipment, but most are standard equipment now.

Graham (1993) summarized a range of cost effectiveness calculations for air bags and showed that the magnitudes were consistent with other expenditure programs and regulations regarding road safety, although this begs the question of whether or not air bags truly are economically justifiable.

Fildes, et al. (1994) evaluated various size driver air bags for Australia, a country with very high seat belt use. They constructed a "harm index" which assigns a monetary value for various type injuries; then the benefits are the sum of monetized changes in the probability of injury associated with the different air bag designs. They concluded that only the largest driver air bag showed a benefit cost ratio greater than one. This was predicated on the assumption that air bags brought a net reduction in injuries, but the North American data show mixed results on this.

The costs and benefits of air bags in Canada (Lawson, 1993)

A careful analysis of costs and benefits of air bags was carried out by Lawson (1993). He reviewed whether or not Canada should adopt mandatory air bag requirements consistent with those in the U.S. He combined estimates of reduced fatalities as well as reduced injuries based on data available at the time. **Table 4** shows his assumptions about the effectiveness of air bags relative to seat belts, both for reducing fatalities and injuries. The assumed reduction of fatalities is quite consistent with the estimates cited above (Table 1) from recent NHTSA studies, although Lawson may have over-stated the effectiveness of air bags alone at reducing fatalities. Lawson's figures for injury categories cannot be compared exactly to those cited earlier in this paper. Dalmotas' figures on injuries (Table 3) are for frontal collisions and not all crashes. The closest figures would be those in column 1 of Table 2 which shows the estimated effectiveness at "moderate" and "serious" (AIS 2 and AIS 3) injuries. Those results suggest that air bags may actually cause net harm, or at least offer very low injury reduction. If so, Lawson's (1993) analysis based on Table 4 would overstate this aspect of the benefits of air bags.

TABLE 4
Estimated Effectiveness of Restraint Systems for Front Seat Occupants
(percentage reduction compared to unrestrained occupants, by injury class)

Restraint type	AIS 1	AIS 2-5	AIS 1-5	Fatal
3 point seat belt	10-20%	45-55%	15-25%	45-50%
air bag, no seat belt	5-10%	20-40%	7-15%	15-30%
air bag plus seat belt	15-25%	55-60%	20-30%	45-55%

source: Lawson (1993) p.S94.

Because most Canadians will purchase cars meeting U.S. standards which include air bags, Lawson (1993) estimated the incremental reduction in fatalities and injuries if Canada made air bags mandatory, and compared this to the incremental costs of adding air bags to cars that would have been purchased without them. Using a value of life of $1.5 million, and an average cost per accident of $11,000, he concluded that the benefits did not justify the costs. He carried out various sensitivity analyses but in virtually no case did he show a positive net benefit. He included lower assumed seat belt use (as low as 60 percent). If one combined low seat belt use with a higher value of life ($2.5 million), this might justify the decision to require air bags. This is relevant to the U.S. with their lower seat belt usage, but not to Canada.

The costs and benefits of drivers' side only air bags are more attractive. The passenger seat often is not occupied therefore the cost per passenger fatality avoided is much higher than that for drivers. Examining Lawson's (1993) calculations, using the higher value of life might justify driver air bags, but marginally so. But if air bags are causing injury costs rather than benefits, this would weaken the case for air bags even for drivers.

An important consideration in the U.S. air bag requirements is the design criteria which assumes that the vehicle occupants do not use seat belts. Despite many years of campaigning and seat belt laws in all but one U.S. state, average seat belt use is estimated at 68 percent (NHTSA, 1996, p.4). Seat belt use is much higher in Canada, 86 percent by 1992 and was then projected to reach 95 percent by 1995 (Lawson, 1993, p.S94); it was at 88 percent in late 1993 (Transport Canada, 1994) and is reportedly is now over 90 percent. The overall or average effectiveness of air bags

depend crucially on whether or not seat belts are used. What may be appropriate for the U.S. may not be appropriate for Canada.

Complications and information needs in evaluating air bag costs and benefits

There are more issues involved than just establishing that there is an overall average improvement due to air bags. There are a number of additional questions about air bag costs and effectiveness. Existing studies show the effectiveness of air bags *given* that a crash has occurred, and given the population of vehicles and drivers in the data base. But there are subtle additional considerations, such as whether or not the presence of air bags (or other safety features) are modifying people's driving behaviour. If so, it may be embodied in the frequency and severity of crashes, but it requires separate study to identify these behavioural influences.

Air bags give rise to some familiar dilemmas in safety regulation: viz., the problem of moral hazard and externalities (drivers may not recognize risks imposed on others). By making cars safer, some safety benefits may be nullified because motorists may compensate by driving faster and more aggressively (the moral hazard problem). Insofar as consumers choose to consume potential safety benefits in the form of time savings, the economists' interpretation is that this is their prerogative and it shows that people value time savings more than safety (e.g., Lave and Weber, 1970). That is, the safety benefits are real and valuable even if they are not consumed as safety. But if externalities are involved, as they often are in safety issues, the "offsetting behaviour hypothesis" does at least partially undermine the benefits of safety measures. That is, offsetting behaviour means the external benefits of safety improvements (reduced risks to others on the road) are reduced. These aspects of driver behaviour must be taken into account in evaluating costs and benefits of air bags or other safety features.

In the case of air bags, evidence suggests that the offsetting behaviour hypothesis is at work. Peterson, Hoffer and Millner (1995) review crash records from two sources and find that:

> "... risk to drivers of cars equipped with air bags in single car crashes *is not* diminished, that the percentage of occupants killed in single car air bag crashes *is* unusually high, and that drivers of air-bag-equipped cars initiate an unusually large percentage of

such crashes. We conclude that drivers of air-bag-equipped cars tend to be more aggressive than drivers of cars not so equipped, that their added aggressiveness diminishes the protection afforded drivers of cars equipped with air bags and imposes additional risks on occupants and passengers in other vehicles." (Peterson, Hoffer and Millner, 1995, p.262) (see also Peterson and Hoffer, 1994).

A further complication involves the interrelationship between seat-belts and air bags. As noted in the introduction, air bags are intended to supplement seat-belts. But there is the possibility that some drivers might rely on air bags instead of using seat-belts. Because seat-belts themselves are highly effective at reducing deaths and injuries, a shift to air bags from seat belts could actually reduce safety and associated net benefits. Lawson (1993) included an illustrative calculation in his sensitivity analysis of the costs and benefits of air bags. If air bags induce even a modest reduction in seat belt use, this may be sufficient to undermine the safety contribution of air bags.

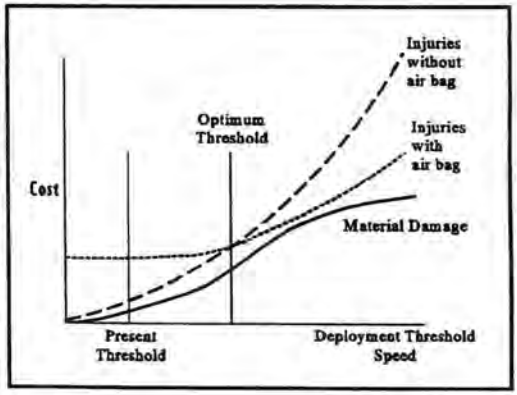
More is needed on the costs of installation and replacement costs of air bags. The actual production costs of air bags is not well-known, but manufacturers were charging on the order of $500-$1000 for air bags as optional equipment when the car is first built (the passenger-side air bag is slightly larger and more costly). The replacement costs are much higher. The air bags often cause damage to the interior of the car (windshields, dashboards) and the sensors may have to be replaced. The Insurance Corporation of British Columbia (ICBC) reports an average cost of $3400 to replace 1995 model air bags (driver and passenger side) (personal interview). The costs sometimes are more than twice this amount.

Another cost component of air bags has received little recognition thus far. As vehicles age and their market value declines, insurance companies will be reluctant to replace the air bags and will scrap the car instead (this presumes that cars will not be placed back on the road without the air bags). That is, the presence of air bags may cause a de facto reduction in the economic life of some cars which have been in a crash. Ordinarily, cars will be scrapped if repair costs are greater than the wholesale value N. The wholesale value N is a function of age and the initial value. Suppose air bag replacements costs are $2000. Recognizing the replacement cost of air bags means that the same damaged car will be scrapped at (N+$2000). Expressed another way, the full additional costs of

air bags are not only the initial costs of the bags, but also their replacement costs after they have been deployed, or the lost value due to premature scrapping of damaged vehicles not worth the air bag replacement costs.

Another characteristic of most safety interventions is that they have an effect over a whole range of injury severities. Safety measures usually have a cascading effect whereby some deaths are prevented and replaced by some level of injury, otherwise-"severe" injuries are now less-severe, and otherwise-"minor" injuries may be avoided altogether. There are additional complications in that different types of injuries may be substituted within the same AIS classification. Evaluation of safety interventions requires incorporating the gamut of injury and damage severity in accidents. The current North American air bags do not have a simple cascading effect on injuries. While some deaths are prevented and the people injured instead, the current aggressive air bag technology in North America is actually worsening some injury rates. More analysis is needed to clarify the injury rates and shifts among injury categories resulting from the various combinations of air bags and seat belts and their use or non-use.

The foregoing raises the need to study the optimal threshold speed for deploying air bags. Part of the costs of present air bag systems arise because deployment at relatively low speeds may be causing as many or more injuries as they prevent, especially for belted drivers, not to mention the costs of air bag replacement in otherwise relatively minor accidents. There is a tradeoff which needs to be studied. To put it simply: if air bags deployed at very low impact speeds, the injuries from the air bag would greatly exceed injuries which would be sustained otherwise. As impact speeds increase, there are progressively greater injuries. Material damages also increase progressively with impact speeds, but taper off once substantial damage has been done to the vehicle. Theoretically, there is some optimal deployment threshold speed where injury



probabilities from the accident exceed those which may be inflicted by the air bag. This is illustrated in Figure 4. Injury rates rise progressively without air bags, whereas low speed air bag deployment would cause injuries at first but these would be overshadowed by crash injuries as impact speeds increase, indicating the optimal threshold speed. Adding the costs of air bags (not shown) would raise the optimal threshold speed still more.

4.0 Conclusion

There has been substantial media attention recently to some deaths caused by low speed deployment of air bags, and the injuries sustained by many others due to the aggressive rate of expansion of air bags. Official warnings have been issued about children in the presence of air bags. There is also concern for people of small stature who may sit closer to the air bags (e.g., many women). There are steps underway to correct for some of these problems. Raising the air bag deployment threshold speed is the most important thing for belted drivers. Reducing the expansion rate of the bag might assist unbelted drivers and low speeds, but this will tend to be offset by reduced effectiveness of the bag for higher speed crashes.

But this recent debate over the potential harm of air bags in low speed impacts has not stimulated much debate about the overall merits of air bags. Indeed, spirited defence of air bags has been the norm. The exploratory calculations here suggest that the benefits of air bags might not warrant their costs, particularly for those who use seat belts properly. There is a great deal of effort going in to monitoring air bag experience and improving the technology. But there is a case for more economic analysis of air bags to go along with the extensive engineering analysis which is already taking place.

ACKNOWLEDGEMENTS

The author wishes to thank Trevor D. Heaver, John Lawson and Mike Macnabb for helpful conversations and advice, and Tony Kan and Shawn Tay for very capable research assistance. The author remains responsible for the interpretations and viewpoints expressed.

REFERENCES

Breed, D.S. (1993) "How Air Bags Work (Design, Deployment Criteria, Costs, Perspectives," *Chronic Diseases in Canada* 14:4 (supp.) S70-S78.

Dalmotas, D.J., (1996) "Performance of the Combination Manual Three-Point Seat Belt and Supplementary Restraint System Based on U.S. Field Accident Data," Discussion Paper, Transport Canada, Road Safety and Motor Vehicle Regulation Directorate, Ottawa, (January).

Evans, L. (1991) *Traffic Safety and the Driver* (New York: Van Nostrand Reinhold).

Fildes, B.N., M.H. Cameron, A.P. Vulcan, K.H. Digges, and D. Taylor (1994) "Airbag and Facebag Benefits and Costs," *Accid. Anal. and Prev.* Vol.26 No.3, 339-346.

Graham, J.D. (1993) "An Economic Perspective on Air Bag Regulation for Canada," *Chronic Diseases in Canada* 14:4 (supp.) S125-S128.

Insurance Institute for Highway Safety (1997) "Airbag Statistics," *www.hwysafety.org/PASSVEH/ABstates.htm*, 28 February.

Lave, L.B. and W.E. Weber (1970) "A Benefit Cost Analysis of Auto Safety Features," *Applied Econ.* Vol.2, 265-75.

Lawson, John (1993) "Cost-Benefit and Cost-Effectiveness Assessments of Potential Regulation Requiring Air Bags in Passenger Cars in Canada," *Chronic Diseases in Canada* 14:4 (supp.) S93-S100.

Lund, A.K. and S.A.Ferguson (1995) "Driver Fatalities in 1985-93 Cars with Airbags," *Journal of Trauma* 38:4, 469-475.

Mannering, F. and C. Winston (1995) "Automobile Air Bags in the 1990s: Market Failure or Market Efficiency?" *J of Law and Econ.* Vol.38, 265-279.

NHTSA (National Highway Traffic Safety Administration) (1996) *Effectiveness of Occupant Production Systems and Their Use*, Third Report to Congress, U.S. DoT, Washington, D.C. (December); available from *www.nhtsa.dot.gov/cars/rules/rulings/208con2e.html.*

NHTSA (National Highway Traffic Safety Administration) (1997) "Questions and Answers Regarding Air Bags" *www.nhtsa.dot.gov/cars/rules/rulings/airbagqa.html* March.

O'Neill,B. and A.K.Lund (1993) "The Effectiveness of Air Bags in Driver Fatalities in the United States," *Chronic Diseases in Canada* 14:4 (Autumn) S53-S59.

Peterson, S., G.Hoffer (1994) "The Impact of Air Bag Adoption on Relative Personal Injury and Absolute Collision insurance Claims," *J.Consumer Res.* Vol.20 (March) 657-662.

Peterson, S., G.Hoffer and E.Millner (1995) "Are Drivers of Air-Bag-Equipped Cars More Aggressive? A Test of the Offsetting Behavior Hypothesis," *J. of Law and Econ* Vol.38, 251-264.

Transport Canada, (1994) Road Safety Directorate, Leaflet TP 2436, Table 1, "Results of October 1993 Survey of Seat Belt Use in Canada."

The Safety of Sea Transport: A Crew Injury Perspective

by

Wayne K. Talley

Frederick W. Beazley Professor of Economics
Old Dominion University
Norfolk, Virginia 23529

Thomas W. Rivers Visiting Distinguished Professor of International Studies
East Carolina University
Greenville, North Carolina 27858

I. Introduction

The focus of the safety regulation of sea transport has been the ship rather than the crew, i.e., regulation has focused on the condition of the ship rather than human actions on the ship. However, the focus is shifting: For example, the International Safety Management Code which will become mandatory for many ships in 1998 requires shipping lines to document their management procedures for detecting and eliminating unsafe human behavior.[1] "This code is at the heart of the industry's plan to switch toward regulating human factors instead of physical ones" (Abrams, May 2, 1996, p. 8B). Also, the ship classification societies[2], United Kingdom's (U.K.) Lloyd's Register of Shipping and the Norweigian Det Norske Veritas, have instituted quality assurance systems applied to the management of ships. The motivation for the shift has been attributed to the following facts: 1) most ship accidents are caused by human error[3], 2) ship accident insurance claims are often attributed to human error[4] and 3) it is less expensive to change human behavior than it is to redesign ships for safety.

Supporters argue that human-action regulations will not only reduce ship accidents but also the number of crew injuries in accidents. A necessary condition for these regulations to be effective in reducing crew injuries is that human causes (the targets of human-action regulations) of ship accidents are positive determinants of crew injuries. The purpose of this study is to

investigate whether this positive relationship exists as well as other determinants of the number of crew injuries in ship accidents. In doing so, the study fills a void in the econometric transport safety literature, since no such investigation heretofore has appeared in the literature. Further, international shipping associations, governments, shipping lines and insurance firms will find the results of the investigation useful information in their debate: on the effectiveness of human-action regulations in improving ship safety; in establishing specific human-action regulations; and in reducing ship accident insurance claims attributed to human error.

A ship accident crew injury model is constructed and estimated, using a unique microdata set of tanker, container and bulk ship accidents.[5] The estimation results provide strong evidence of a positive relationship between crew injuries and human (as opposed to environmental or ship) causes of ship accidents as well as strong evidence of fire/explosion accidents resulting in greater numbers of crew injuries than other types of ship accidents.

The study is structured as follows: The availability of ship injury statistics are discussed in Section II, followed by a model of ship accident crew injuries in Section III. Data are discussed in Section IV. Estimation results are detailed in Section V, followed by conclusions in Section VI.

II. Ship Injury Statistics

Statistics of worldwide ship accident injuries are sparse, published neither by the International Maritime Organization (IMO) nor by the International Labour Organization (ILO). Some countries publish statistics (as used in this study) of injuries in ship accidents that they investigate, but these statistics are not standardized across countries as to definitions and coverage. The ship classification society, Lloyd's Register, publishes the number of fatal injuries of total loss ship accidents (reported to its agents), but not those of ships that are not total losses; further, no non-fatal injury statistics are published. Other ship classification societies publish even less injury data.

While ship accident injury data collected by ship classification societies are sparse, ship (and cargo) damage data are extensive. The reason is clear: Ship classification societies are primarily concerned with safety of property as opposed to safety of crew at sea. Ships and their cargo are usually insured (requiring property damage data for processing insurance claims), while the lives of people on board are often not.

Recent statistics of worldwide ship fatalities (compiled by the Institute of London Underwriters) show a large increase in ship accident fatalities: 730 people were either killed or reported missing in the first half of 1996, compared with 225 during the same period for 1995 and 316 for all of 1995 (Porter, 1996c, p. 8A). Goss, Nicholls and Pettit (1991) argue that accidental

injuries on commercial ships are under reported by a factor of approximately 10; if so, there are 10,000 annual fatalities at sea worldwide and some 550,000 non-fatal injuries of sufficient severity to be reportable.

III. A Model of Ship Accident Crew Injuries

I model the number of crew that are injured in a ship accident (NUMINJ) as a function of the number of crew on board the ship (COBD), ship damage severity (SDS) and injury-prevention investment and effort (IPVE):

$$NUMINJ = f(COBD,SDS,IPVE). \quad (1)$$

The a priori sign of the relationship between COBD and NUMINJ is positive, i.e., as the number of crew on board a ship increase greater the likelihood that a member of the crew will be injured in a ship accident. Ship damage severity (SDS) should have a non-negative effect on NUMINJ, given that a damaged ship does not necessarily result in injuries. IPVE is a vector of investment and effort activities undertaken by both government and shipping lines to prevent ship-accident injuries and should have a negative effect on injuries.

I also model COBD, SDS and IPVE as endogenous variables, i.e., as functions of other variables. The number of crew on board a ship (COBD) varies with ship size (SSIZE) and ship age (SAGE), i.e.,

$$COBD = g_1(SSIZE,SAGE). \quad (2)$$

SSIZE should have a non-negative effect on COBD, since a larger sized ship does not necessarily require a larger crew. A positive relationship is expected between SAGE and COBD, since older ships tend to be more crew labor intensive than newer ships.

Ship damage severity (SDS) is expected to be affected by the type of accident (ATYPE), cause of accident (ACAU), type of ship (STYPE), ship operating conditions (SOC), ship size (SSIZE), ship age (SAGE) and ship damage-prevention investment and effort (SDPVE), i.e.,[6]

$$SDS = g_2(ATYPE,ACAU,STYPE,SOC,SSIZE,SAGE,SDPVE). \quad (3)$$

The type of accident includes collision (COLLISION), a fire/explosion (FIRE-EXPLOSE), a material/equipment failure (MAEQ-FAILURE), or a grounding accident.[7] The a priori relationship between type of accident and SDS is indeterminate. The cause of an accident may be due to either a human[8] (HUMAN), an environmental[9] (ENVIRON), or a ship cause[10]. It is unclear, however, which cause will result in a greater ship damage severity. The type of ship includes tanker (TANKER), container (CONTAINER), or bulk ship. Whether a tanker or a container ship will incur greater ship-accident damage than a bulk ship is unclear.

Ship operating conditions describe the environment in which a ship was operating at the time of an accident. This environment includes

weather/visibility characteristics and phase of operation. Weather characteristics include foggy weather (FOG) and wind speed (WIND). Visibility is differentiated by time of day, i.e., by whether the accident occurred at night (NIGHT) versus day. Although adverse weather and visibility are likely to increase the risk of an accident, their impact on accident ship damage severity is unclear.

The phase of operation is described by whether the ship was underway (UNDERWAY) or adrift or docked or moored at the time of an accident. In groundings and collisions, ship damage should be greater for ships that are underway (given the speed of impact) than for those that are adrift or docked or moored. However, no such prediction can be confidently made for fire/explosion and material/equipment failure accidents.

Returning to the variables in equation (3), we come to ship size (SSIZE) and ship age (SAGE). The a priori sign of the relationship between SSIZE and SDS is indeterminate. A positive relationship is expected between SAGE and SDS, since vessel structural failure is expected to increase with age.[11] The last determinant of SDS is ship damage-prevention investment and effort (SDPVE) and is expected to have a negative effect on SDS.

I also model SDPVE as an endogenous variable -- dependent on the extent of government ship safety regulation (SSR), ship safety regulation enforcement (SSRE), ship size (SSIZE), ship age (SAGE), and price of shipbuilding and repair services (PRICE-BLDREP):

$$SDPVE = h(SSR,SSRE,SSIZE,SAGE,PRICE\text{-}BLDREP). \quad (4)$$

Safety regulation of ships varies among countries. The United States (U.S.), however, is a nation among nations with the highest maritime safety standards (Gracey, 1985). I measure the extent of the variation in the safety regulation of ships by their flags of registry, i.e., by whether they are of U.S. (US-FLAG) or of non-U.S. registry. A positive relationship is expected between US-FLAG and SDPVE. If US-FLAG reflects factors other than regulation (e.g., subsidy practices), the directional relationship between US-FLAG and SDPVE is indeterminate.

While a country may have high ship safety standards, the enforcement of these standards may be lax, thereby diluting the positive impact of the former on ship safety.[12] I measure the extent of the variation in the safety regulation enforcement of ships by their classification society, i.e., by whether they are classified (and therefore inspected for adherence to safety regulations) by the American Bureau of Shipping (ABS-SOCIETY) or by an alternative classification society. The relationship between ABS-SOCIETY and SDPVE is indeterminate. Although the American Bureau of Shipping is one of the four largest classification societies in the world (the other three

being the Lloyd's Register of Shipping, the Bureau Veritas of France and the Det Norske Veritas based in Norway), both the smallest and the largest classification societies have been criticized for lax safety inspections.

The directional relationship between SSIZE and SDPVE is indeterminate. The a priori sign of the relationship between ship age (SAGE) and SDPVE is negative. The effect of the price of shipbuilding and repair (PRICE-BLDREP) on SDPVE is not predictable: On the one hand, the more expensive repair services are, the greater the incentive to reduce ship damage (i.e., a positive safety effect); on the other hand, higher shipbuilding and repair prices may make some damage-prevention investments to ships more expensive, thereby reducing them (i.e., a negative safety effect).

The remaining endogenous variable in equation (1) is injury-prevention investment and effort. I model IPVE as a function of ship safety regulation (SSR), ship safety regulation enforcement (SSRE), ship size (SSIZE) and ship age (SAGE):

$$IPVE = g_3(SSR,SSRE,SSIZE,SAGE). \quad (5)$$

As for the SDPVE equation, SSR and SSRE are measured by US-FLAG and ABS-SOCIETY, respectively. The expected directional relationships between IPVE and its explanatory variables are the same as those for these variables in the SDPVE equation, i.e., the directional relationships for IPVE with respect to US-FLAG, ABS-SOCIETY, SSIZE and SAGE are positive, indeterminate, indeterminate and negative, respectively.

Substituting equation (4) into equation (3), and then equations (2), (3) and (5) into equation (1) gives:

$$NUMINJ = F(SAGE,SSIZE,SSR,SSRE,STYPE, ACAU,ATYPE,SOC,PRICE\text{-}BLDREP). \quad (6)$$

Replacing variables SSR, SSRE, STYPE, ACAU, ATYPE and SOC in equation (6) by the variables used to classify or measure them, I obtain the following equation to be estimated:

$$NUMINJ = F^*(SAGE, SSIZE, US\text{-}FLAG, ABS\text{-}SOCIETY, TANKER, CONTAINER, HUMAN, ENVIRON, COLLISON, FIRE\text{-}EXPLOSE, MAEQ\text{-}FAILURE, FOG, WIND, NIGHT, UNDERWAY, PRICE\text{-}BLDREP). \quad (7)$$

I obtain two estimates of equation (7) -- an estimate for the number of crew fatalities (NUMFATAL) in a ship accident and an estimate for the number of crew members (NUM-NONFATAL) that incurred non-fatal injuries in a ship accident.

IV. Data

I estimate equation (7) with microdata of individual commercial ship (U.S.

and foreign flag) accidents (investigated by the U.S. Coast Guard) for the eleven-year time period 1981-91. The foreign flag ship accidents occurred in U.S. waters; the U.S. flag accidents are not restricted to any body of water. Data (except for that of PRICE-BLDREP) were obtained from a computer tape of marine casualty information (i.e., the CASMAIN database) provided by the Coast Guard. Price index data for the variable PRICE-BLDREP were taken from various issues of Producer Prices and Price Indexes (Bureau of Labor Statistics, U.S. Department of Labor). The real price of shipbuilding and repair was determined by dividing PRICE-BLDREP by the U.S. Producers Price Index for all commodities (divided by 100) found in various issues of Producer Prices and Price Indexes (Bureau of Labor Statistics, U.S. Department of Labor).[13] Variables used in the estimations of equation (7) and their specific measurements appear in Table 1.

Descriptive statistics (mean and standard deviation) of the microdata via variable appear in Table 2. The mean statistics reveal that 64.6 percent of the ships were U.S. flag; 44.7 percent were classified by the American Bureau of Shipping; 42.1, 21.5 and 36.4 percent were tanker, container and bulk ships, respectively; and 25.0, 4.8, 43.1 and 27.1 percent were involved in collision, fire/explosion, material/equipment failure and grounding accidents, respectively. Human, environmental and ship factors caused 42.5, 16.6 and 40.9 percent of the accidents, respectively; 43.8 percent occurred at night; and 82.9 percent of the ships were underway.[14]

V. Estimation Results

Equation (7) could be estimated by ordinary least squares. However, ship accidents with crew injuries are infrequent; also, if injuries occur, they are few in number. The preponderance of zeros and the small valves and clearly discrete nature of the dependent variable (number of crew injuries in a ship accident) suggest that the estimation could be improved (relative to least squares) by using the statistical technique, Poisson regression analysis, which accounts for these characteristics.[15] Estimation bias from possible omitted relevant explanatory variables that are correlated with time is addressed by including a yearly time trend (YEAR) variable (having integer observations 1981 through 1991) in the estimation.

Separate Poisson regression estimates of equation (7) for number of fatal crew injuries (NUMFATAL) and number of non-fatal crew injuries (NUM-NONFATAL) in a ship accident appear in Table 3. The fatal injury estimate reveals that none of the ship characteristic variables are statistically significant. For the non-fatal injury estimate, ship age is significant with a positive coefficient, suggesting that the number of non-fatal injuries in ship accidents increase with ship age. For the type-of-ship variables, TANKER

is significant while CONTAINER is insignificant in the fatal injury estimate, suggesting with their positive coefficients that tanker accidents have a larger number of crew fatalities than container or bulk ship accidents. However, the insignificance of both variables in the non-fatal estimate suggests that neither type of ship has more non-fatal crew injuries than another.

The estimation results for cause of accident are particularly important: They predict that ship human-action regulations will reduce (if they are effective in reducing human causes of ship accidents) the number of both fatal and non-fatal crew injuries. HUMAN is highly significant with a positive coefficient in both estimates, while ENVIRONMENT is insignificant -- suggesting that if the cause of a ship accident is human related, the number of crew injuries (both fatal and nonfatal) will be greater than if the cause is either environmental or ship related. Further, HUMAN's positive coefficient is supportive evidence of a positive relationship between crew injuries and human causes of ship accidents.

In both estimates, FIRE-EXPLOSE has a positive coefficient and is highly significant; further, its coefficient is more than twice the size of either positive coefficient of the remaining type-of-accident variables. These strong results suggest that fire/explosion ship accidents result in a greater number of crew injuries (fatal and non-fatal) than collision, material/equipment failure or grounding accidents. COLLISION is significant in both estimates; MAEQ-FAILURE is significant in the non-fatal injury estimate but not in the fatal injury estimate. These results taken together suggest that: 1) collision or fire/explosion accidents result in a greater number of crew fatalities than material/equipment failure or grounding accidents and 2) collision, fire/explosion or material/equipment failure accidents result in a greater number of non-fatal injuries than grounding accidents.

Among operating condition variables, WIND has a positive coefficient and is significant in both estimates, suggesting a positive relationship between the number of crew injuries (fatal and non-fatal) in a ship accident and wind speed. FOG and NIGHT are significant (insignificant) in the fatal (non-fatal) injury estimate; the signs of their coefficients are positive and negative, respectively. The results suggest that there are a greater number of fatal injuries if a ship accident occurs in foggy weather and less at night. The latter may reflect a more safety conscious crew at night given the reduced visibility. Price of ship building and repair (PRICE-BLDREP) has a negative coefficient in both estimates. It is significant at the 15 percent level (not reported in Table 3), a low (and often not reported) level of significance, in the fatal injury estimate and insignificant in the non-fatal estimate. However, these results may suggest that rising prices for ship building and

repair provide an incentive for shipping lines to reduce costs elsewhere -- e.g., by becoming more safety conscious by taking actions to reduce ship accidents and crew injuries and therefore their associated costs.

The coefficient of the yearly time trend variable (YEAR) is negative in both estimates and significant at the 10 and 14 (not reported in Table 3) percent levels for the fatal and non-fatal estimates, respectively. These results suggest a declining trend in the number of fatal and non-fatal crew injuries per ship accident.

VI. Conclusion

This study has filled a void in the econometric transport safety literature by investigating determinants of crew injuries (both fatal and non-fatal) in ship accidents. Poisson regression estimates suggest that the number of fatal crew injuries are greater: 1) for tanker than for container or bulk ships; 2) if the accident cause is human rather than environmental or ship related; 3) for fire/explosion (in particular) or collision accidents than for material/ equipment failure or grounding accidents; 4) at higher wind speeds; and 5) when the weather is foggy. Poisson estimates also suggest that the number of non-fatal crew injuries are greater: 1) for older ships; 2) if the accident cause is human related rather than environmental or ship related; 3) for fire/explosion (in particular), collision or material/equipment failure accidents than for grounding accidents; and 4) at higher wind speeds.

The strong and robust results for human (as opposed to environmental or ship) cause of accidents have particularly important safety regulation policy implications for sea transport: They are strong evidence of a positive relationship between crew injuries and human causes of ship accidents, thereby providing support for the current shift in safety focus toward regulating human actions on the ship as opposed to the condition of the ship. Further, they predict that ship human-action regulations will reduce (if they are effective in reducing human causes of ship accidents) the number of both fatal and non-fatal crew injuries. It is left for future research to investigate which human-action regulations will be effective in reducing human causes of ship accidents.

Notes

1. An in-depth discussion of this code is found in Beetham (1995).

2. Ship classification societies (some dating back more than 200 years) inspect ships to ensure that: vessels are seaworthy, they meet national-flag

requirements and conform to international safety standards. No underwriter will insure a ship if it is not classed. For a discussion of current issues related to classification societies, see Boisson (1994) and Brooks (1996).

3. A common assertion by those associated with shipping is that nearly 80 percent of ship accidents are caused by human error (Goss, 1994). Human error is defined by U.K.'s P&I Club as any human action or omission identifiable as the immediate cause of the event from which the liability arises, including blame worthy behavior from simple mistakes in arithmetic, judgement and deliberate risk taking. For further discussion of human error in shipping, see Millar (1980), Goulielmos and Tzannatos (1995), Porter (1996a) and Porter (1996b).

4. For the period 1987-91 sixty percent of the claims of the U.K. P&I Club, the world's largest marine liability insurer, were attributed to human error (Tzannatos and Goulielmos, 1996).

5. For a general discussion of commercial ship accidents, see Ch. 7 of Loeb, Talley and Zlatoper (1994).

6. A similar ship damage severity function is found in studies by Talley (1995,1996) investigating determinants of the ship damage severity of tanker and container ship accidents.

7. A ship accident may be described as a: 1) collision -- vessel struck or was struck by another vessel on the water surface, or struck a stationary object, not another ship (an allision); 2) grounding -- vessel is in contact with the sea bottom or a bottom obstacle, struck object on the sea floor, or struck or touched the bottom; 3) fire and explosion -- the fire and/or explosion is the initiating event reported, except where the first event is a hull/machinery failure leading to the fire/explosion; or 4) a material and equipment failure (or structured-machinery-other) -- hull/machinery damage, missing, and miscellaneous non-classified reasons (e.g., vessels sunk due to either weather or break-up related to causes not covered by other casualty categories).

8. Examples of human causes of ship accidents include stress, fatigue, carelessness, operator error, calculated risk, improper loading, lack of training, error in judgement, lack of knowledge, physical impairment, improper cargo stowage, inadequate supervision, improper mooring/towing, design criteria exceeded, psychological impairment, intoxication, failed to

yield right of way, improper safety precautions, failed to keep proper lookout and failed to proceed at safe speed.

9. Examples of environmental causes of ship accidents include debris, shoaling, lightning, adverse weather, submerged object, channel not maintained, unmarked channel hazard, hazardous bridge/dock/pier and adverse current/sea condition.

10. Examples of ship causes of ship accidents include corrosion, cargo shift, dragging anchor, stress fracture, brittle fracture, fouled propeller, improper welding, steering failure, propulsion failure, static electricity, temperature stress, inadequate controls/displays/lighting, inadequate horsepower, inadequate lubrication and auxiliary power failure.

11. Meek, Brown and Fulford (1985) conclude that for ships in general a positive relationship exist between ship accident rates and ship age.

12. Although there is a worldwide parity trend in vessel safety standards, there are significant differences in their enforcement across nations. "Disparities in enforcement do much more than economic damage -- they serve to undermine the safety system" (Gracey, 1985, p. 10C).

13. Since the accidents of non-U.S. flag ships (in the data) occur in U.S. waters, these vessels may have ship repairs done (especially for relatively major ship damage) in U.S. shipyards; if so, the U.S. price index of shipbuilding and repair is an appropriate ship-repair price index for these ships.

14. Our data consist of 29 ship accidents with one or more crew fatalities; the largest number of fatalities in one accident is 31. Seventy-six accidents incurred one or more non-fatal crew injuries; the largest number in one accident is nine.

15. Poisson regression analysis has been used extensively in investigating determinants of (infrequent) air carrier accidents (see Rose, 1990 and the references therein). An application of Poisson regression to investigating determinants of motor-vehicle fatalities is found in Keeler (1994). For a discussion of the Poisson regression model, see Greene (1997).

References

Abrams, A. "New Rules Put Focus on Human Factors." Journal of Commerce, May 2, 1996, p. 8B.

Beetham, E. H. "The ISM Code - The Human Factors." Proceedings of the International Symposium on Human Factors on Board, Bremen, 1995, pp. 1-9.

Boisson, P. "Classification Societies and Safety at Sea." Marine Policy, vol. 18, no. 5 (1994), pp. 363-377.

Brooks, M. R. "The Privatization of Ship Safety." Selected Papers from the 7th World Conference on Transport Research, Editor, H. Haralambides, a special issue of the journal Maritime Policy and Management, vol. 23, no. 3 (July-September, 1996), pp. 271-288.

Bureau of Labor Statistics, U.S. Department of Labor. Producer Prices and Price Indexes. Washington, D.C.: U.S. Government Printing Office, various issues.

Goss, R., C. Nicholls and S. Pettit. "Seamen's Accidental Deaths and Injuries Worldwide: A Methodology and Some Estimates." Journal of the Royal Institute of Navigation, vol. 44, no. 2, 1991.

Goss, R. "Safety in Sea Transport." Journal of Transport Economics and Policy, vol. 28, no. 1 (January, 1994), pp. 99-110.

Goulielmos, A. M. and E. Tzannatos. "The Man-Machine Interface and Its Impact on Shipping Safety." Proceedings of the Conference on Critical Issues in Maritime Economics, Massachusetts Institute of Technology, Cambridge, Massachusetts, December, 1995.

Gracey, J. S. "Many Question Value of Drive for Vessel Safety Standards." Journal of Commerce, May, 1985, p. 10C.

Greene, W. H. Econometric Analysis, 3rd ed. Upper Saddle River, New Jersey: Prentice Hall. 1997.

Keeler, T. E. "Highway Safety, Economic Behavior, and Driving Environment." American Economic Review, vol. 84, no. 3 (June, 1994), pp. 684-693.

Loeb, P. D., W. K. Talley and T. J. Zlatoper. Causes and Deterrents of Transportation Accidents: An Analysis by Mode. Westport, Connecticut: Quorum Books, 1994.

Meek, M., W. R. Brown and K. G. Fulford (1985): "A Shipbuilder's View of Safety." Maritime Policy and Management, vol. 12, no. 4 (October-December, 1985), pp. 251-262.

Millar, I. C. "The Need for a Structured Policy Towards Reducing Human-Factor Errors in Marine Accidents." Maritime Policy and Management, vol. 6 (1980), pp. 9-15.

Porter, J. "Ship Crew Training Blasted." Journal of Commerce, July 2, 1996a, p. 8A.

Porter, J. "Videos Teach Ship Safety." Journal of Commerce, July 10, 1996b, p. 6A.

Porter, J. "Many More Killed at Sea than in Air Crashes." Journal of Commerce, August 6, 1996c, p. 8A.

Rose, N. L. "Profitability and Product Quality: Economic Determinants of Airline Safety Performance." Journal ofPolitical Economy, vol. 98, no. 5 (October, 1990), pp. 944 -964.

Talley, W. K. "Vessel Damage Severity of Tanker Accidents." The Logistics and Transportation Review, vol. 31, no. 3 (September, 1995), pp. 191-207.

Talley, W. K. "Determinants of the Ship Damage Severity of Containership Accidents." Selected Papers from the 7th World Conference on Transport Research, Editor, H. Haralambides, a special issue of the journal Maritime Policy and Management, vol. 23, no. 3 (July-September, 1996), pp. 239-247.

Tzannatos, E. and A. M. Goulielmos. "Management Information System for the Promotion of Safety in Shipping." A paper presented at the 1996 International Conference: International Association of Maritime Economicsts, Vancouver, Canada, June, 1996.

Table 1

Variables and Their Measurement

Variable	Measurement
Age of Ship (SAGE)	Years
Size of Ship (SSIZE)	Gross Tons
U.S. Flag Ship (US-FLAG)	1 if a U.S. flag ship; 0 otherwise
American Bureau of Shipping (ABS-SOCIETY)	1 if ship classified by the American Bureau of Shipping; 0 otherwise
Tanker Ship (TANKER)	1 if a tanker ship; 0 if a container or bulk ship
Container Ship (CONTAINER)	1 if a container ship; 0 if a tanker or bulk ship
Human Cause of Accident (HUMAN)	1 if ship accident caused by a human factor; 0 if caused by an environmental or a ship factor
Environmental Cause of Accident (ENVIRON)	1 if ship accident caused by an environmental factor; 0 if caused by a human or ship factor

Table 1 cont'd.

Variables and Their Measurement

Variable	Measurement
Collision Accident (COLLISION)	1 if a collision; 0 if a fire/explosion or material/equipment failure, or if a grounding
Fire/Explosion Accident (FIRE-EXPLOSE)	1 if a fire and/or explosion; 0 if a collision or material/equipment failure, or if a grounding
Material/Equipment Failure Accident (MAEQ-FAILURE)	1 if a material and/or equipment failure; 0 if a collision or fire/explosion, or if a grounding
Fog (FOG)	1 if foggy weather; 0 otherwise
Wind Speed (WIND)	Knots
Night (NIGHT)	1 if nighttime; 0 if daytime
Underway (UNDERWAY)	1 if ship is underway; 0 if adrift or if docked or moored
Price of Ship Building and Repair (PRICE-BLDREP)	Price Index
Crew Fatalities (NUM-FATAL)	Number of crew fatalities in a ship accident
Crew Non-Fatal Injuries (NUM-NONFATAL)	Number of crew non-fatal injuries in a ship accident

Table 2
Descriptive Statistics: Mean and Standard Deviation

Variable	Mean	Standard Deviation
SAGE	15.442	12.183
SSIZE	24,271	24,229
US-FLAG	0.646	0.478
ABS-SOCIETY	0.447	0.497
TANKER	0.412	0.494
CONTAINER	0.215	0.411
HUMAN	0.425	0.494
ENVIRON	0.166	0.372
COLLISION	0.250	0.433
FIRE-EXPLOSE	0.048	0.213
MAEQ-FAILURE	0.431	0.495
FOG	0.055	0.229
WIND	12.482	13.685
NIGHT	0.438	0.496
UNDERWAY	0.889	0.377
Real PRICE-BLDREP	94.466	7.889
NUM-FATAL	0.027	0.603
NUM-NONFATAL	0.040	0.343

Table 3

Ship Accident Crew Injuries: Poisson Regression Estimates

Variable	NUMFATAL	NUM- NONFATAL
Ship Characteristic		
SAGE	0.019	0.017*
	-0.015	-0.01
SSIZE	0	0
	0	0
US-FLAG	-0.278	-0.223
	-0.56	-0.357
ABS-SOCIETY	0.72	0.443
	-0.476	-0.331
Type of Ship		
TANKER	0.832*	-0.016
	-0.476	-0.309
CONTAINER	0.406	-0.215
	-0.601	-0.382
Cause of Accident		
HUMAN	3.608***	1.299***
	-1.051	-0.32
ENVIRON	1.511	-1.125
	-1.432	-1.04
Type of Accident		
COLLISION	2.232**	2.728***
	-1.045	-1.03
FIRE-EXPLOSE	5.812***	5.665***
	-1.052	-1.041
MAEQ-FAILURE	1.316	3.655***
	-1.438	-1.051
Operating Condition		
FOG	1.326**	0.824
	-0.675	-0.544
WIND	0.038**	0.027***
	-0.016	-0.01

Table 3

Ship Accident Crew Injuries: Poisson Regression Estimates

Variable	NUMFATAL	NUM- NONFATAL
NIGHT	-0.717*	-0.417
	-0.426	-0.266
UNDERWAY	0.34	0.048
	-0.464	-0.32
Real Price of Ship Building and Repair and Time Trend		
Real PRICE-BLDREP	-0.163	-0.073
	-0.113	-0.068
YEAR	-0.494*	-0.24
	-0.279	-0.162
Constant	983.61	475.477
	-563.816	-326.342
N	1,957	1,957
-2 x log Likelihood Ratio	169.200***	447.442***

aStandard errors are in parentheses; N represents sample size.
***(**,*)Significant at the 1(5,10) percent level.

Pavement Deterioration and Truck Loads in Winnipeg

Mohammed M. Alam[1], Alan M. Clayton[1], Fred P. Nix[1]

Truck traffic is believed to be important in the process of pavement deterioration. However, there is much that is either not known or, at least, imperfectly understood about the relationship between axle loads and pavement deterioration. These unknowns create difficulties for pavement designers ("what is optimal pavement strength?"), for those who determine truck weight and dimension regulations ("what is the maximum load for a given axle group?"), and for those who attempt to allocate highway costs ("how much of pavement costs are attributable to truck traffic?")

This paper describes research[2] that develops pavement performance equations for truck routes within Winnipeg. The equations relate the change in condition of pavements over time to differences in axle loads. It uses historical pavement condition data collected between 1985 and 1994 and traffic data—converted to axle loadings—developed from a recent study.[3]

The key finding is that pavement deterioration in Winnipeg is relatively insensitive to truck axle loads. Environmental factors (aging) are far more important. Strictly, it is not possible to extend these findings to other pavements in Canada. Nevertheless, there are reasons to believe that other pavements on low-volume roads are similar. The paper concludes with speculations on the implications.

Background

Pavements are designed to accommodate a projected number of axle load repetitions of a specified magnitude, for a projected service life. Axle loads are measured with a load equivalency factor, the most common one being the 18,000-pound equivalent-single-axle-load (ESAL) developed by the AASHO (now "AASHTO") road test of the early 1960s. The

performance of pavements over their service life is measured by an index that is estimated based on measured pavement distresses. These occur under the stresses imposed by heavy axle loads, freeze-thaw cycles, and subgrade moisture changes. The quantitative measure of pavement deterioration for a given pavement section is the change of this performance index over time. A commonly-used index is the Pavement Condition Index (PCI).

Pavement performance—that is, the change in pavement deterioration (PCI) over time—is generally measured in one of a number of models. These explain pavement deterioration through mathematical relationships between a pavement performance index and other explanatory variables (loads, climates, moisture, construction practice etc.). The most commonly used model was developed through the AASHTO road test. This relationship was further extended in 1986 and 1992, to include empirical factors for loss of support, drainage, climate, and reliability. A pavement-performance index (not the PCI) was used as a measure of the pavement deterioration. Pavement damage was explained by the accumulation of ESALs throughout the analysis period.

The AASHTO road test and subsequent research did not address all the complex range of considerations required for a contemporary understanding of pavement performance. Of particular importance is the fact that these tests were conducted in an accelerated manner, accumulating large pavement loadings over a short, two-year period. For this reason, the interactive effects of aging and loading were suppressed. Further, certain axle group arrangements—such as tridems and split tandem axles, which are now common—were not considered. Nor were the tests able to address changes in suspension systems, types and sizes and tires, and tire pressure which have occurred since the AASHTO test. In the early 1990s, the Canadian long-term-pavement-performance (C-LTPP) part of the Strategic Highway Research Programm (SHRP) was undertaken to address these and other limitations of the AASHTO road tests.

Winnipeg's regional streets provide a unique opportunity to examine the relative contribution of axle loads and aging. This is because Winnipeg has a significant number of pavement sections where design features, construction and maintenance histories, and subgrade conditions are similar, but where loading histories, and on-going loading levels differ

significantly. Winnipeg's Streets and Transportation Department has been monitoring and recording the different distresses in these pavement sections since 1985. These distresses, along with measures of truck traffic, provide an opportunity to investigate the effects of loading under otherwise quite constant infrastructure circumstances.

Model of pavement deterioration for Winnipeg

The development of a pavement deterioration model was carried out in the following stages:

1. Identify and group Portland Cement Concrete (PCC) pavement sections into categories defined by similar design specifications, construction, and maintenance histories.

2. Identify the distresses for the pavement sections in (1) using the data on pavement performance. This data is produced using the PAVER pavement management system.

3. Develop a measure to estimate the relative truck loading and classify the sections under (1) based on the historical truck loading.

4. Develop and compare the pavement deterioration model using the pavement condition of (2) for the loading of (3).

The PCI is established by using a trained observer who records a variety of distresses for randomly selected sample units within each section. The PCI estimation methodology uses Deduct Values (DVs) for the individual distress from a series of prior-established curves that are function of the type, severity, and density of each distress. The DVs of all distresses due to load, climate and subgrade conditions in one sample unit are then combined in a nonadditive, nonlinear manner to arrive at a total DVs (ranging 0 to 100) that is then subtracted from the maximum PCI of 100. The value of PCI ranges from 0 to 100, where 100 is excellent. The methodology also assumes that pavement distresses after "n" years are a result of: (1) climate-related effects; (2) traffic-related effects (axle loads); and (3) other factors. Consequently, pavement deterioration or the change in PCI (ΔPCI) for a given year can be expressed as:

$$\Delta PCI_{for\ a\ given\ year} = \Delta PCI_e + \Delta PCI_l + \Delta PCI_o \ ;$$

where:
ΔPCI = total loss in PCI,
ΔPCI_e = loss of PCI due to climate,
ΔPCI_t = loss of PCI due to traffic load,
ΔPCI_o = loss of PCI due to other factors.

PCI for a specific pavement section in year "n" would be,

$$PCI_n = PCI_{n-1} - \Delta PCI$$

For a specified region such as Winnipeg, climatic factors such as extreme temperature gradient and freeze-thaw effects are assumed uniform for a given year thus causing the same ΔPCI_e. This assumption is consistent with AASHTO's "climatic zones" concept. The ΔPCI_o, mainly due to subgrade conditions, may vary from one pavement section to another. No information on subgrade conditions was available. The analysis simply assumes that subgrade was designed based on the specification as outlined in Winnipeg's design manual for rigid pavements and, therefore, is mostly uniform over the truck route network.

Data and quality analysis

The quality and consistency of the data used in the research had to be assured before proceeding with model development. From the standpoint of the pavement performance data, three sources of potential anomalies were assessed: inconsistent maintenance, sampling variability, and measurement variance.

The PCI rating for an individual section can increase (pavement improves over time) or decrease at a faster rate compared with the last-rated PCI because of unrecorded maintenance activity.[4] To ensure that pavement sections analysed had, in fact, been maintained in a similar manner through the period in which their PCIs had been measured, only those sections receiving routine maintenance (low level maintenance) through the observation period were retained for analysis. Pavement sections which received a major maintenance effort between two observations of PCI were rejected. This is necessary to ensure compliance with the assumption of similar design, construction, subgrade conditions, and maintenance of pavement sections considered.

Another potential source of variability in a PCI estimate is the sampling process. That is, the person collecting the data uses random samples from each pavement section. However, an analysis of variability in random sampling showed that the variance across the PCI estimates due to different combinations of random sampling was not statistically significant.

A third potential source of variability is the experience, knowledge, and judgement of the individual inspector collecting the data. There is an element of judgement used in labelling pavement distresses and the resulting PCI data may reflect the views of different people. The possibility that this variability is present in the Winnipeg PCI data was investigated through an analysis of variance. For the newer pavements, with a PCI of 50 or higher, there is little variation in the ratings assigned by different inspectors. Pavements with a rating of less than 50 are labelled "failed" pavements. And, with these failed (and older) pavements , the differences in the ratings of different inspectors increases rapidly. This is because, for badly deteriorated pavements, there are many distresses to measure each with some element of judgement. The mean distribution of PCI used in the analysis ranges between 75-77, indicating low measurement variation.

Some PCI data were removed from the analysis if one or more of the following criteria were met:

- pavement sections not part of Winnipeg's truck route system (the analysis is only interested in the streets on the truck route network);
- sections showing a different construction time during the observation periods (in other words, pavement rehabilitation took place during the analysis period and the inclusion of such data would only "confuse" the analysis of deterioration of original pavements);
- sections with a PCI lower than 100 at the year of construction (that is, for some reason, a new pavement was constructed that already—in terms of the PCI data—shows some deterioration);
- sections involved with unknown maintenance activities;
- outlier PCI data points whose corresponding t-statistics have a 5% probability of being exceeded in an observed number of trials.

In total, 337 PCC sections were identified in the database, with 696 observations. An "observation" is one record of a PCI at a point in time.

After removing observations based on the above criteria, the remaining 298 sections with 405 observations were used.

Total ESALs, based on the truck traffic obtained from a recent study [3], were estimated using a methodology proposed by Fekpe .[5] Truck weight data collected at provincial scales , in August of 1994, were used to determine the mean axle weight of each axle group of different truck classes (loaded trucks only), as well as the equivalent truck load factor (TLF) value for those classes. In other words, the assumption is that the characteristics of trucks just west (Headingley) and east (West Hawk) of Winnipeg are valid for truck traffic within the city. No truck weight data were available for light trucks (LT) and heavy single unit trucks (HT). Therefore, maximum axle weights (maximum under current weight and dimension regulations) are used to calculate the TLF value for these two categories. This introduces a slight bias to the results (ie, these trucks are assumed to be heavier than trucks in the actual traffic flow). Table 1 summarizes the average TLF value per truck, for each class.

Table 1 Mean Axle Group Weight, GVW, & Truck Load Factors

Truck Class[6]	Axle 1 kg	Axle 2 kg	Axle 3 kg	Axle 4 kg	Axle 5 kg	GVW kg	TLF
3-S2	4,835	12,736	11,773	-	-	29,334	1.8865
3-S3	5,053	14,740	18,017	-	-	37,810	2.4229
7-axle A	4,995	13,951	13,138	7,709	8,029	47,823	4.5277
7-axle B	4,954	15,845	15,034	15,156	-	50,989	4.4750
8-axle B	5,068	15,834	20,835	14,974	-	56,711	4.5745
LT truck	4,500	9,100	-	-	-	-	1.5877
HT truck	5,500	16,000	-	-	-	-	1.6728

Note: mean values exclude empty trucks.

Total ESALs per year on an individual pavement section is the sum of ESALs per year for all truck classes. This information is then used to separate PCI data into three load groups:

Group-1: Pavement sections with less than 50,000 ESALs per year
Group-2 : Pavement sections with 50,000 to 200,000 ESALs per year
Group-3: Pavement sections with more than 200,000 ESALs per year.

Model development

The influence of truck traffic on pavement deterioration is investigated by establishing the relationship between the PCI and AGE for the pavement sections for each of the three load levels. Alternative specifications of the model—where both AGE and ESALs are used as explanatory variables—were tried. However, the addition of ESAL as an independent variable does not significantly improve the results (the R^2 increases from 0.86 to 0.87). For this reason, a model using pavement sections grouped by annual axle loads is preferable. The pavement performance curve of each load group is determined using nonlinear regression analysis. The following simplified nonlinear equation is used to quantify the pavement deterioration for each load group.

$$PCI = \alpha \; e^{-\left(\frac{T}{\rho}\right)^{\beta}}$$

Where:

T the age of the pavement in years;
ρ controls the position of the first inflection point on the curve along the age axis;
β the coefficient which controls the shape of the curve (convex or concave);
α an asymptote (intercept) that controls the maximum value of PCI at the first year of pavement construction or overlay; and
e the base of the natural logarithm.

Applying the initial value of PCI at T=0, the previous equation can be expressed as:

$$PCI = 100 \; e^{-\left(\frac{T}{\rho}\right)^{\beta}}$$

The following statistical measures were investigated to measure the degree of relationship between the variables, model validity and usefulness;

- estimation of the coefficient of determination, ie, R^2;

- the F-test;
- plots of studentized residuals versus predicted values of PCI;
- plots of studentized residuals versus pavement ages; and
- plot of predicted PCI versus observed PCI.

The mean and standard deviation of PCI and AGE in each load group are summarized in Table 2. These are an indication of an unbiased data set that follows a similar distribution of PCI and AGE. The outlier data points in each group were determined based on the t-value that has a 5% probability of being exceeded in the observed trials. Table 3 summarizes the calculated R^2 values, shape coefficients, and standard errors of the estimated coefficients.

Table 2 PCI and AGE Population, Mean, and Standard Deviation

		AGE		PCI	
Pavement Sections	Population	Mean	Standard Deviation	Mean	Standard Deviation
Group-1	85	14	9	77	18
Group-2	228	14	9	75	17
Group-3	102	15	9	75	18

Table 3 Shape Coefficient and Standard Error

Pavement Sections	R^2	ρ	Standard Error	β	Standard Error
Group-1	0.86	37.000	1.765	1.310	0.087
Group-2	0.85	49.360	2.085	1.001	0.040
Group-3	0.89	41.151	1.827	1.235	0.071

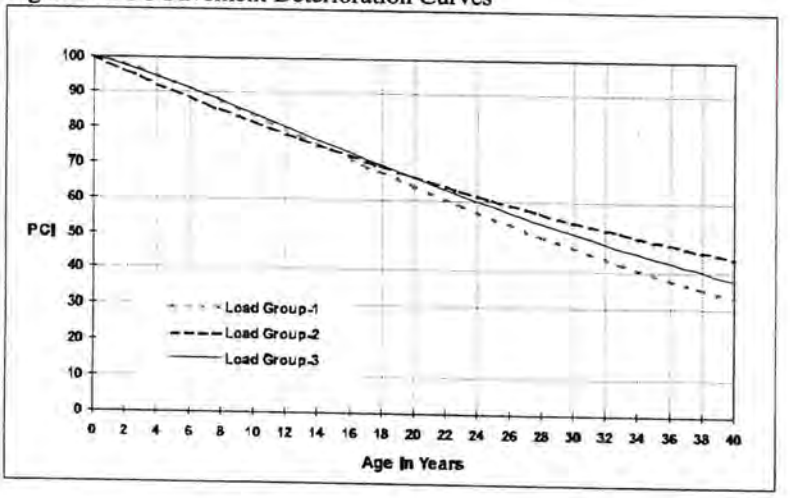
Table 4 F-test Parameters

Pavement Sections	n-2	F (estimated)	$F_{.05,\,2,\,n-2}$ (critical)
Group-1	83	5666	3.10
Group-2	202	20377	3.41
Group-3	87	9160	3.11

The high R^2 and low standard error confirm the visual validity of the developed performance curve in each group. The F-test statistics of each relationship are summarized in Table 4. The null hypothesis, that there is no positive relationship between PCI and AGE, is rejected at the 95% confidence level if $F_{estimated} \geq F_{.05,2,n-2}$. Since the estimated F-value is significantly larger than the critical F-value for all groups, the null hypothesis is rejected. In other words, the mathematical relationship obtained in the analysis is highly significant. Additional tests were carried out by plotting the studentized residual variables against the predicted PCI and observed pavement age. The residual plots in each group showed no unusual pattern, and are normally distributed about the axis of predicted PCI and observed AGE. The model developed is statistically adequate in predicting the pavement performance over time, across the load groups.

All deterioration curves, across the load groups, are then combined to compare their deterioration over time. Figure 1 illustrates the deterioration curves for the three load groups.

Figure 1 PCC Pavement Deterioration Curves



The most significant observation is that, although Load Group-3 has accumulated more than four times the ESALs over the years than Load Group-1, the deterioration of the pavement under both load groups is

similar. The model indicates that, for the light traffic loads and harsh climates in Winnipeg, aging (the environment or the climatic) causes most deterioration. For example, with annual average ESALs of approximately 125,000 (Load Group-2), the terminal PCI value 50 is reached in about 34 years. With approximately three times the loading (Load Group-3), the model predicts a PCI value 50 in about 31 years.

Other research

This study of Winnipeg truck route pavements is not unique in the sense that other Canadian studies have also found, or used models to suggest, that truck axle loads play a relatively small role in pavement deterioration. For example:

- An Alberta model from the early 1980s, based on data from a large number of pavement sections, suggests that truck loadings are only weakly related to pavement deterioration. For example, with annual loadings of 17,500 ESALs pavements on these low-volume roads need resurfacing in about 26 years. With loadings of four times this amount, resurfacing is required in about 22 years.[7]

- A model of pavement performance developed in Newfoundland, based on data from the TransCanada Highway, does not include the "number of heavy vehicles" as an explanatory variable as the research leading to the development of this model showed that this variable was not significant.[8]

- In costing work for the Royal Commission on National Passenger Transportation, the assumption was made that environmental factors accounted for between 40% (freeways) to 80% (low-volume rural roads) of pavement deterioration.[9] These assumptions are based on an analysis of flexible pavements in Ontario using the OPAC model of pavement performance.

So, while the finding that truck axle loads are not that important for pavement deterioration on low-volume roads is not new, what may be unique about the Winnipeg research is that these findings have been extended to: (a) urban roads, (b) where the roads are truck routes, and (c) where a rigid pavement is involved (most other studies, indeed most inter-city pavements in Canada are flexible pavements).

Policy implications

The pavement design, subgrade conditions, climate, traffic loads and maintenance practices in Winnipeg are, of course, somewhat unique. For this reason, it is not strictly possible to extend the findings of this study to other roads. That said, the findings along with the other research cited, raise interesting public policy questions. Within the context of this paper, these are best characterized as "what if" questions.

First, is it possible that public resources are being wasted because pavements are over-designed? Pavement designers use the expected number of ESALs as an important determinate of pavement design (thickness). This is based largely on assumed relationships between pavement deterioration and truck traffic. Yet, particularly for low-volume roads, the actual empirical evidence in Canada has trouble finding much of a relationship between truck traffic and pavement wear. With pavement construction costs in the range of $150,000 to $250,000 or more per two-lane kilometre[10], even a pavement that is "5 percent over-built" rapidly adds up to many millions of dollars (say $10,000 per two-lane kilometre). This is a serious ("what if") public policy question in an era where budgets for road departments are being cut and, yet, the public appears to demand higher levels of road service—capacity, routes, or serviceability (smoother pavements).

Second, is it possible that the use of the road is being over-regulated? That is, suppose that pavements are over-designed **or** that pavements have to be designed as strong (thick) as they are in Canada because of the climatic (freeze-thaw, etc). In either case, there is a possibility that the axle loads allowed on these pavements are not too important. That is, they are not "important" in terms of the pavement deterioration. Again, this "what if" question is really only appropriate for Canada's low-volume roads. It is interesting to pose this question within the context of roads on the Prairies where volumes tend to be low and where maximum tandem-axle weights have increased from 16,000 kg to 17,000 kg (1988) and where the provinces are now considering a further increase to 18,000 kg.[11] Is there an implicit assumption here that, in fact, axle loads are not as important to the performance of the pavements on the Prairies as sometimes assumed? Whatever the answer, the point is that there is a direct relationship between trucking costs and allowed maximum axle

loads for most bulk or dense freight: freight costs fall as allowed axle weights increase.

Third, is it possible that the common belief about truck taxes is in error? That is, while full highway cost allocation studies have never been done in Canada (there have been a few partial ones), it is commonly understood that truck taxes (fuel, registration) should be high because truck axle loads account for the greatest proportion of pavement costs. For example, in the United Kingdom, where cost allocation studies are taken seriously (there is a link between the findings and the tax regime) total pavement costs are simply allocated on the basis of total ESALs.[12] Cost allocation methods in the United States and Australia result is a somewhat lower allocation to trucks as factors other than ESALs are recognized as important in pavement performance. But Canada is different—harsher climates, lower traffic volumes, etc—and it may well be that truck axle loads are relatively unimportant on many low-volume roads in terms of cost responsibility. Even though there is not a direct link between cost causality and taxes in Canada, there are assumptions made about which road users should pay more or less than other road users. Perhaps these assumptions need examining.

Comment

Winnipeg pavements—and indeed much of Canada's highway infrastructure —experience low truck volumes. At the same time, they experience high environmental loadings in terms of changes in temperature, subgrade moisture, and so on. From this research, Winnipeg's PCC pavements do not show signs of truck load deterioration. This suggests that a number of transportation and tax policies in Canada may well be open to question.

END NOTES

11. All provinces met in Toronto on February 6 and 7th, 1997 and again on March 7, 1997 to consider changes to Canada's national agreement on truck weights and dimensions. One of proposals discussed was to standardize tandem axle loads at 18,000 kg and to increase tridem axle loads to as high as 25,000 kg (or higher in Eastern Canada). At the time this paper is being written, these proposals are being considered by the deputy ministers of transportation in each province.

12. U.K. Department of Transport, *The Allocation of Road Track Costs*, various years, London.

1. Mohammed Alam is employed by DS-Lea Consultants Ltd and is enrolled in a PhD program in civil engineering at the University of Manitoba. This paper is based on his M.Sc thesis. Mr. Alam was also involved with the TransPlan 2010 study from which traffic data has been used. Alan Clayton is a professor of civil engineering at the University of Manitoba. He was Mr. Alam's thesis advisor. Fred Nix is a transportation consultant based in Ontario.

2. Alam, M. M., 1996, *Influence of Truck Traffic in Pavement Deterioration in the City of Winnipeg* M.Sc Thesis, University of Manitoba.

3. TransPlan 2010, 1996, *Profile of Urban Goods Movements*, Winnipeg.

4. Lodge, C., 1995, *Variability in Pavement Distress Measurement*, Undergraduate thesis, Dept of Civil and Geological Engineering, University of Manitoba.

5. Fekpe, E.S.K, & A. Clayton, 1994, "Truck Weight Models and Knowledge Transfer", Dept of Civil and Geological Engineering, University of Manitoba

6. A "3-S2" is a 5-axle tractor-semitrailer, a "3-S3" has 6-axles, an A-train is a double trailer configuration using an A-dolly converter, a B-train is a double trailer configuration that uses no converter dolly, "LT" and "HT" are light and heavy straight trucks.

7. Cheetham, A. & J.T. Christison, 1981, "The Development of RCI Prediction Models for Primary Highways in the Province of Alberta," Alberta Transportation, Edmonton.

8. Papagiannakis, A.T., D. Shea, J. English & W.T. Beckett, 1990, "Development of Pavement Performance Models for Newfoundland," RTAC annual conference, St John's.

9. Nix, F.P., M. Boucher & B. Hutchinson, 1990, "Road Costs." Royal Commission on National Passenger Transportation, Vol 4 of Final Report, Ottawa.

10. These are within the range of the numbers used in the work for the Royal Commission (op cit Nix, Boucher & Hutchinson) which, admittedly are now somewhat dated, are for flexible pavements, and are typical Ontario values. The point here is not to provide any accurate estimate of average pavement costs in Canada. Rather, it is simply to give a flavour to the possible misallocation of resources if pavements are being over-designed.

The Economy-Transportation Linkages in Ontario: A Sectoral Perspective

N. Bedi, Ministry of Transportation, Ontario*
and
A.A.Kubursi, McMaster University

Introduction

Transportation is integral to Ontario's economic and social well-being. The production, distribution and consumption of almost all goods and most services directly embodies transportation. Yet, there is a huge gap in knowledge on the role and importance of transportation in the changing economy. Statistics such as tonne-kilometres of cargo moved and vehicle-kilometres travelled, while essential to the management of many aspects of the transportation business, are not sufficient for placing transportation within the wider economic context.

Provincial income and product accounts provide aggregate measures of the contributions of economic sectors and activities in terms of value added, employment, expenditures and other relevant indices. These aggregate indices do not provide much information on how any sector relates to other sectors and activities in the economy. This is especially problematic for transportation, given its extensive linkages to other sectors. An additional complication is that transportation is defined very narrowly in the provincial income product accounts. Excluded are transportation services not purchased in the marketplace and some key expenditures on transportation services by households, government and businesses.

This paper is based on a broader study undertaken to identify, develop and quantify a large set of indicators and measures to gauge the many complex linkages and contributions made by transportation to the economic wealth and wellbeing of the province of Ontario over the past two decades. Selected aspects of that study are reported here.

* The views expressed are those of the authors and not necessarily those of the ministry.

The Sectoral Perspective: Is Transportation Just Another Sector?

Each economic system has a unique internal structure. The study of this internal structure is a prerequisite to the proper understanding and management of the economy. Consequently, techniques such as "input-output" analysis and/or process analysis have been used to describe the internal economic framework of many countries.

Inter-industry tables are now widely used accounting frameworks for the analysis of sectoral linkages. In this framework, each sector is considered to buy its inputs from other sectors and sell its outputs to other sectors and, in the "open" version of the system, to an "autonomous sector" which has no output of its own. The "autonomous sector", if there is one, represents final demand and is explained within the model. Analysis of these tables through appropriate techniques (e.g. input-output analysis) provides valuable insight into the indirect relationships of an economic system and their consequences.

This paper uses input-output data for Ontario for 1979, 1984 and 1990, with detailed analysis of the economic and technical implications of sectoral interdependencies involving the transportation sector in the Ontario economy. Space permits the display of only the 1990 data. Implications of the following will be examined in some detail:

- the nature and extent of indirect and induced links among the various sectors of the economy with the transportation sector, with associated multipliers for the transportation subsectors;
- the position of the transportation sector within the different types of productive sectors classified according to their input uses and output distribution;
- the nature and extent of forward and backward linkages among sectors connected to the transportation sector;
- the determination of measures of dispersion of the various coefficients of linkages, and
- the identification of "key" sectors of Ontario's economy in relation to the transportation sector.

Transport Cost Intensities in Ontario's Structure of Production

An industry may directly sell or buy from only a few industries, but its customers and suppliers may be connected with many industries. This industry may thus have a profound influence on the economy through its indirect relationships with other industries. The technical **input-output matrix** is a valuable resource in understanding such relationships.

The impact of transportation cost on the competitiveness of industries in the domestic and international markets is a policy concern. The output tables for Ontario in 1979, 1984 and 1990 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 include commercial (for hire) air, water, rail, 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.

The **intensity** of transport cost for an industry is defined as the cost of transport as a percentage of the total cost of its output. The concept of **transportation margins** is important in estimating the transportation cost. In the Use 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, 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.

Services, the transportation sector itself and forestry had the highest transport cost intensities in Ontario in 1979. Some industries, such as wood, furniture and printing, do not appear to make large purchases from the transportation sector directly but they pay significant transportation margins. Services show a transportation cost intensity of $0.14 for 1979 without transportation margins and about $0.15 with the margins. The total transportation cost intensities for transportation and forestry were $0.07 and $0.063 respectively. Nonmetallic minerals and refined petroleum ranked

next. Sectors with high transport cost intensity are also sectors with high direct value added coefficients. This fact suggests that 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. Industry rankings by relative transportation cost intensities are very similar for 1979, 1984 and 1990. However, the magnitudes of the intensities have changed markedly. Transportation, services, forestry, mining and refined petroleum show relatively high transport cost indices, but their values are far below those in the 1979 with the exception of the transport sector itself (Table 1).

Table 1

Ontario
Transportation Intensities, Margins And Value Added, 1990

	Transportation Intensities	Transportation Margins	Transportation Cost Shares	Value Added Coefficients
Transportation	0.1610	0.0043	0.1653	0.5318
Fishing	0.0557	0.0008	0.0565	0.4311
Refined Petroleum	0.0225	0.0038	0.0263	0.1190
Other Services	0.0182	0.0021	0.0203	0.3899
Mining	0.0089	0.0035	0.0125	0.6705
Forestry	0.0077	0.0026	0.0103	0.6439
Metals	0.0072	0.0325	0.0396	0.2699
Chemicals	0.0057	0.0100	0.0157	0.3873
Paper	0.0051	0.0187	0.0238	0.3694
Food & Beverage	0.0046	0.0146	0.0192	0.3193
Non-Metallic Minerals	0.0044	0.0119	0.0163	0.4672
Trade & Finance	0.0040	0.0004	0.0044	0.7121
Agriculture	0.0033	0.0073	0.0106	0.3643
Textiles	0.0029	0.0077	0.0106	0.6135
Transportation Equipment	0.0028	0.0071	0.0099	0.2382
Clothing	0.0024	0.0070	0.0094	0.4261
Fabricated Metal	0.0023	0.0105	0.0128	0.4580
Construction	0.0023	0.0083	0.0106	0.5058
Wood	0.0022	0.0201	0.0223	0.3472
Electrical Products	0.0021	0.0050	0.0071	0.4614
Furniture	0.0015	0.0095	0.0110	0.4106
Machinery	0.0012	0.0063	0.0075	0.4075
Printing	0.0008	0.0109	0.0118	0.5562
Misc. Manufacturing	0.0006	0.0037	0.0043	0.4160
Foreign Export	0.0140	0.0220	0.0360	N/A
Interprovincial Exports	0.0390	0.0350	0.0740	N/A

Source: Econometric Research Limited

The results indicate that bulky natural resources and high valued commodities require directly or through purchased inputs, large transport inputs in their production processes, whereas textiles, agriculture, electrical products and clothing have low transportation costs per unit of output.

Transportation cost intensities are relatively high for exports generally, but more so for interprovincial exports than for foreign exports. The high ratios of transport to other costs in the export sector are typical of moving goods to distant markets. This is a potential drag on the competitiveness of Ontario exports, particularly to other provinces.

Another fact to be noted is the composition of Ontario exports. Transport equipment exports are significant. They accounted for about 40% of total Ontario foreign exports in 1979, and increased to about 44% in 1990. If we were to add the exports to other provinces, these shares would rise to 50% of total Ontario exports.

The Transportation Sector Industrial Linkages

To evaluate an industry's direct and indirect relationships with other industries, we have to evaluate the "matrix multiplier" $(I-A)^{-1}$. The gross output levels (x's) required to sustain a given vector of final demand (f) in the open model is determined by the following equation system:

$$\textcircled{1} \qquad x = (I-A)^{-1} f$$

If the inverse of (I-A) matrix exists, it may be expressed by means of binomial expansion:

$$(I-A) = I^{-1} + (-1)I^{-2}(-A) + \frac{(-1)(-2)}{2!}I^{-3}(-A)^2 + \frac{(-1)(-2)(-3)}{3!}I^{-4}(-A)^3 + \ldots$$

$$\textcircled{2} \qquad = I + IA + IA^2 + IA^3 + \ldots$$

$$= I + A + A^2 + A^3 + A^4 + \ldots$$

$$= \sum_{k=0}^{\infty} A^k$$

The inverse matrix, $(I-A)^{-1}$, indicates the total direct plus indirect output required per unit of final demand. The series in (2) simply explains the general composition of this total output requirement. The first term, I, accounts for one unit of output to be delivered to final demand. The second term, A, indicates the direct input required to produce this unit of final demand. The next term, A^2, indicates first-round indirect inputs required to produce the direct input A, and so on. Due to the fact that the elements of the A matrix satisfy the Hawkins-Simon condition $a_{ij}^{(k)}$ decreases as k increases and the $(I-A)^{-1}$ is approximated by the sum of the powers of A.

Income and Employment Multipliers in the Ontario Economy

The macro "Keynesian" multipliers are simply the overall total of direct and indirect effects of a dollar increase in final demand. This summing of direct and indirect income effects is quite similar to the summing of the direct and indirect output effects in the input-output context discussed in the preceding

section. In fact, it is also possible to use input-output techniques to evaluate the income effects due to a change in final demand. By its very nature macroeconomics is concerned with the economy at large but strictly at the most general level. This is also true of its income multipliers. The question of what industries will produce the extra output when final demand is increased is irrelevant to the macroeconomic perspective. Input-output analysis deals with smaller components of the economy than does aggregate macro-economic analysis. The emphasis is on individual sectors, not the national total. In this way the transport sector is seen as a subcomponent of a larger system of interacting sectors.

Table 2 presents multipliers and other sectoral indicators that position the transportation sector in relation to other sectors in the economy for 1990. The objective here is to examine the transportation sector (transportation services and transport equipment) within the overall structure of production in Ontario for three specific years for which an input output table exists. The sectoral perspective is examined by using traditional output and income multipliers, direct value added and labour income (wages and unincorporated income) coefficients, jobs per million dollar of output, the share of labour in total income of the sector and finally average labour productivity. Twenty four sectors are evaluated on the same platform using these indicators.

The income and output multipliers are suggestive of the efficiency with which these sectors generate their output and income per one dollar of final demand. Industries (sectors) with high income multipliers are those that generate large incomes when the demand for their output increases. They may do so through a high direct income or through the relationships they maintain indirectly with sectors that are efficient in generating income per unit of their output.

Table 2

Comparative Industrial Indicators by Sector, 1990

Sector	Output Multipliers	Income Multipliers	Direct Value Added	Wage Value Added	ncorpora Income	Labour Income	hare of Labou Value Added	Employment Per Million $	verage Labou Productivity
Agriculture	2.54208	1.06753	0.36426	0.11952	0.0610	0.18056	0.4957	18.8	$53,246
Forestry	2.34795	1.30277	0.64387	0.32160	0.1719	0.49353	0.7665	10.6	$94,604
Fishing	2.90976	1.33556	0.43114	0.31628	0.0431	0.35935	0.8335	30.0	$33,288
Mining	2.01940	1.17002	0.67047	0.22610	0.0036	0.22975	0.3427	5.4	$184,386
Food & Beverage	2.64934	1.04940	0.31927	0.17114	0.0007	0.17185	0.5382	4.3	$230,919
Textiles	2.30110	1.23575	0.61352	0.40633	0.0012	0.40757	0.6643	5.3	$190,074
Clothing	2.50316	1.17179	0.42610	0.32390	0.0026	0.32652	0.7663	23.1	$43,262
Wood	2.85277	1.21915	0.34717	0.26380	0.0047	0.26850	0.7734	9.4	$106,872
Furniture	2.60865	1.16804	0.41060	0.29546	0.0023	0.29775	0.7252	9.6	$103,828
Paper	2.67327	1.14436	0.36941	0.25123	0.0000	0.25123	0.6801	5.3	$189,405
Printing	2.54115	1.28661	0.55624	0.38496	0.0031	0.38805	0.6976	9.3	$107,358
Metals	2.69866	1.03231	0.26993	0.20393	0.0001	0.20401	0.7558	4.8	$208,568
Fabricated	2.61176	1.19190	0.45800	0.33065	0.0024	0.33305	0.7272	9.8	$101,779
Machinery	2.39163	1.05260	0.40748	0.26558	0.0006	0.26620	0.6533	6.9	$145,724
Transportation. Equ	2.58295	0.87671	0.23819	0.14951	0.0001	0.14962	0.6282	3.5	$285,828
Electrical	2.40155	1.12142	0.46138	0.30741	0.0002	0.30759	0.6667	7.3	$137,187
Non-Metallic Miner	2.40825	1.16962	0.46722	0.29486	0.0010	0.29590	0.6333	7.4	$135,783
Refined Petroleum	1.44443	0.33335	0.11898	0.04100	0.0002	0.04116	0.3459	1.6	$611,254
Chemicals	2.39254	1.03700	0.38731	0.22560	0.0002	0.22575	0.5829	4.2	$240,320
Misc. Manufacturin	2.50681	1.15723	0.41601	0.29232	0.0012	0.29352	0.7055	9.7	$102,901
Construction	2.52104	1.23094	0.50579	0.33210	0.0388	0.37088	0.7333	6.4	$156,135
Transportation	2.29782	1.16819	0.53177	0.26997	0.0070	0.27697	0.5208	8.3	$119,976
Trade	2.27214	1.33974	0.71209	0.34108	0.0566	0.39773	0.5585	11.3	$88,120
Other Services	2.91691	1.28611	0.38992	0.22820	0.0666	0.29478	0.7560	22.1	$45,346

Source: Econometric Research Limited

A high value added coefficient is indicative of a high degree of internal processing. A sector that ships raw materials without any further processing typically has low value added coefficients. Sectors with high value added per unit of output are those that purchased inputs that are processed within the respectice sectors, adding significant value through the use of labour, capital and entrepreneurship.

A high labour income share in value added reveals the labour intensity of the production process of the sector. Refined petroleum, for example, uses few workers and as such should show a low share of labour income per unit of output or in value added.

Obviously, a high employment content per million dollar of output designates the sector as a significant contributor to employment generation in the province. This may often be consistent with low average labour productivity in the sector.

The transportation and storage sector in 1979 had below average output multipliers, above average income multipliers, a very high labour income share in value added, a relatively high employment content and not surprisingly a relatively low average labour productivity. Transportation and storage were effective in generating income and employment for the province but did not contribute much to its overall labour productivity.

By way of contrast, the transportation equipment producing sector shows a relatively high output multiplier (among the top 5 sectors) and one of the lowest income multipliers. This combination is indicative of a highly specialized but integrated production structure of the auto industry in North America. Establishments fabricate only a small part of the total product they sell. At the same time, transport equipment production in 1979 had a relatively low direct value added coefficient but a surprisingly high labour share in value added. The latter follows from the type of labour used in production and the high wages they received, given their high productivity. Alternatively, it is not surprising to find that employment per million dollar of output in this sector is relatively low (9.1 jobs) while average labour productivity is relatively very high ($109.7 thousands). The sector, however, sustains very high skilled, highly productive and well paid labour.

The combination of the transport and storage sector with the production of transport equipment sector within a broader and more comprehensive transportation sector provides a more "balanced" sector with even output,

income, employment and productivity contributions.

The consideration of structural change over time by using the input output tables for 1984 and 1990 reveals some interesting developments. The output and income multipliers of transportation and storage services declined between 1979 and 1984 but increased substantially in 1990 (see Table 2). On the other hand, employment per million dollar of output declined continuously over the period with the concomitant increase in average labour productivity.

Similar trends are observed in transportation equipment production. The results in Table 2 show a small but steady rise in the output and income multipliers of the sector but a drastic decline in employment per million dollar of output, with a notable rise in labour productivity.

The increased capitalization of the transportation sector, including both transport equipment and transportation services and storage, the anticipation of NAFTA and the intensification of global competition have combined to increase the efficiency of this sector. This increased efficiency was not accompanied by commensurate gains in employment.

Transportation and storage and the production of transportation equipment compare very favourably with other sectors in terms of both of the indicators, but particularly in the 1990 case.

The sectoral perspective on transportation from the selective indicators above is one of an efficient sector with high labour income ratios and substantial contributions to output, income and productivity. Missing from this presentation is the enormous contribution the transport equipment sector makes to the balance of payments of Ontario with the rest of the country and the rest of the world. The transport equipment sector is, after all, the leading export sector in Ontario.

Is Transportation A Key Sector in the Ontario Economy?

Hirschman defines a key sector as a sector with high forward and backward links with other domestic sectors. Hirschman's definition, however, does not impose any restrictions on variability of the linkages. This is a crucial issue because a sector can have very large links to only one sector. Far more important here is the spread of these links over many sectors. That is why we define the key sector to be one with large backward and forward links to many sectors. We gauge the links of a specific sector by the magnitude and the

spread of its purchases from and sales to other sectors.

Variability is measured by the lack of concentration of these purchases or supplies in a small subset of sectors.

A high v_j may be interpreted as indicating that a particular industry draws heavily on one or a few sectors and a low v_j as a sector drawing evenly from other sectors. Similarly, one can interpret that v_i's in the same way.

Adopting Hazari's criterion, a key sector is one which has:

(a) both u_i and u_j greater than, and

(b) both v_i and v_j are low relative to their averages.

This definition of key sectors can again be identified with Hirschman's definition of a key sector as one with high forward and backward links. Hirchman's definition, however, does not impose any restrictions on variability.

In Table 3 we present a two-way classification of forward and backward linkages and their respective distributions in 1990. Obviously the sectors that fall under high (backward links) L_j and low (variability) V_j reveal a high absorption rate from a large number of sectors in the economy. Similarly, sectors with high (forward links) L_i and low (variability) V_i show high forward linkages with a large supply network.

The sectors that possess both high L_i and low V_i, and high L_j and low V_j are key sectors in the economy. Surprisingly, only the food and beverages sector qualify for the key sector designation in Ontario. In 1979, transportation and storage was one of the sectors that showed higher than average forward linkages (1.82) and a low coefficient of variation (1.25), signifying that it supplied a large number of sectors with its services. On the other hand, it had low backward linkages (0.93) and a high coefficient of variation (2.70) indicating a low absorption rate of the products of only a few sectors in Ontario. By way of contrast, the transportation equipment sector had low forward linkages and these were highly concentrated in few sectors, in 1979.

The backward linkages of both the transportation services and transportation

Table 3

Forward Linkages and Coefficients of Variation, 1990

Sector Name	*Variability (V_i)*	*Forward Linkage (L_i)*
Sectors with Low Forward Linkage and Low Coefficient of Variation		
Agriculture	3.11	0.76
Mining	3.04	0.66
Paper	3.07	0.75
Printing	2.97	0.69
Primary Metals	2.92	0.83
Fabricated Metal	3.12	0.67
Transportation Equipment	2.95	0.91
Chemicals	2.48	0.93
Misc. Manufacturing	2.86	0.76
Sectors with Low Forward Linkage and High Coefficient of Variation		
Forestry	4.74	0.44
Fishing	3.69	0.61
Textiles	4.10	0.48
Clothing	4.20	0.47
Wood	4.02	0.53
Furniture	4.42	0.46
Machinery & Equipment	4.02	0.51
Electrical	3.40	0.63
Non-Metallic Minerals	3.94	0.52
Refined Petroleum	4.49	0.43
Construction	3.21	0.60
Sectors with High Forward Linkage and Low Coefficient of Variation		
Food & Beverage	2.09	1.12
Transportation	1.14	2.02
Trade & Finance	0.67	3.16
Other Services	0.59	5.04

equipment sectors did not change much in relative magnitudes between 1979

and 1990. Almost the same backward linkage pattern was observed between 1979 and 1990. On the other hand, the forward linkages of the transportation equipment sector underwent some significant changes between 1979 and 1990. In fact, the coefficient of variation of the forward linkages increased to above average values between 1979 and 1984 and then decreased to below average values by 1990.

For both transportation subsectors, the observed forward and backward linkages have, however, changed in absolute terms between 1979 and 1990. These developments are indicative of a process of greater integration of the Ontario sectors with foreign and other Canadian sectors outside Ontario. The 1990 coefficients amplify these trend and capture the maturation process that the transportation sectors have undergone between 1979 and 1990 (see tables 3 and 4). The transportation equipment sector remained exactly within the same blocks of designation in terms of backward linkages throughout the period. What changed were the absolute values of the coefficients. This is more pronounced in the case of forward linkages than in the case of backward linkages.

While only one sector qualified for the key sector status in Ontario between 1979 and 1990, other sectors may qualify, were we to relax the strict conditions (a) and (b). Tables 3 and 4 show some of these possible candidates that may have qualified as key sectors but did not because of slight violation of one of the conditions in (a) and (b) while ranking better than average on the rest. Transportation equipment and transportation and storage both are good potential candidates. Transportation and storage would benefit from larger backward linkages while transportation equipment would benefit from larger forward linkages.

It is to be noted that the key sector status should not be judged outside the full appreciation of the many indicators that were used in this study. The focus of this section is on the way the transportation activity, be it in terms of services or equipment, is a critical link in the production chains of the province. The objective here is to quantify the extent to which the health of the Ontario economy is linked to the health and proper functioning of the transportation sector. This perspective is all the more important when issues of efficiency and competitiveness are taken into account. The transportation sector, narrowly or broadly defined, is shown to be a critical sector in Ontario not only in terms of the size of its direct contribution to the economy, but also in terms of the linkages to other critical sectors in the province.

Table 4

Backward Linkages and Coefficients of Variation, 1990

Sector Name	*Variability (V_j)*	*Forward Linkage (L_j)*
Sectors with Low Backward Linkage and Low Coefficient of Variation		
Forestry	2.35	0.95
Textiles	2.28	0.93
Machinery & Equipment	2.25	0.97
Electrical	2.35	0.97
Non-Metallic Minerals	2.23	0.97
Sectors with Low Backward Linkage and High Coefficient of Variation		
Mining	2.56	0.82
Refined Petroleum	3.33	0.58
Chemicals	2.55	0.97
Transportation	2.86	0.93
Trade & Finance	2.93	0.92
Sectors with High Backward Linkage and Low Coefficient of Variation		
Agriculture	2.42	1.03
Fishing	2.25	1.18
Food & Beverage	2.38	1.07
Clothing	2.18	1.01
Wood	2.06	1.15
Furniture	2.08	1.05
Paper	2.28	1.08
Printing	2.22	1.03
Primary Metals	2.38	1.09
Metal Fabricated	2.17	1.06
Misc. Manufacturing	2.31	1.01
Construction	2.10	1.02
Sectors with High Backward Linkage and High Coefficient of Variation		
Transportation Equipment	2.68	1.04
Other Services	3.11	1.18

There exists a high correlation between sectoral contributions to GPP and the intensity of transportation costs in total production cost of these sectors. The high forward and backward linkages and relatively low coefficients of variation of transportation services and transportation equipment suggest that the economy of Ontario is strongly dependent on the proper functioning and the efficient use of the transportation infrastructure and services.

Economic Structural Change and Transportation

Structural change is a significant feature of the transportation sector in both of its components. The output and income multipliers of transportation services declined between 1979 and 1984 but increased substantially in 1990. Given the large and rapid increases in oil prices during the period, these changes are easy to understand. Similar trends were observed in the transport equipment sector. The increased capitalization of the transportation sector (engendered by declining real rates of interest), the anticipation of FTA and NAFTA and the intensification of global competition have combined to increase the efficiency with which both transportation services are delivered and transport equipment are produced.

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. It is the nature of the products exported and the location of the export markets that account for this fact. Nonetheless, this reduces the competitiveness of Ontario Exports in distant provinces where American firms in the proximity may be more competitive. This must be balanced against the enormous contribution the transport equipment sector makes to the Ontario and Canadian balance of payments.

There is an evident maturation process through which the two transport sectors have passed between 1979 and 1990. This process has evidently accelerated after FTA and NAFTA were concluded. Both trade agreements have increased the pace and level of integration of the North American markets. Globalization of production has also increased the diversification of the sector and its material sourcing. There is a definite trend towards lower domestic contents and increased capitalization, efficiency and exports. These trends are evident in the indicators displayed in this study.

Summary and Conclusions

The importance of the transportation sector to the economy is generally not questioned in the literature. There is, however, little consensus on the exact

boundaries (what activities it should include) of this sector and the magnitudes of its importance (contributions to GDP and employment). The issues are both conceptual and empirical.

The study reported in this paper attempts to draw a more comprehensive and consistent picture of the transportation sector's contribution to Ontario's economy than has been available in the past. Showing how this contribution has evolved over time makes it clear that transportation is still a key sector in Ontario. It is highly developed and efficient in delivering transportation services and products whose health is of great importance to the economy. This work needs to be extended (e.g. subprovincial analysis) and accompanied by additional analyses, such as the description and quantification of specific mechanisms through which transportation infrastructure contributes to growth in output and productivity. Together, such analyses will provide the insight necessary in addressing the contemporary economic policy questions aimed at ensuring a responsive yet economically efficient transportation system.

References

Dougberty, C. 1991. "A Comparison of the Performance of the U.S. and Canadian Economies: 1960-1987." Draft Report.

Eberts, R.W. 1990. "Public Infrastructure and Regional Economic Development." Pp 15-27 in Economic Review. Volume 26, Number 1, Federal Reserve Bank. Cleveland, OH.

Goode, F.M. 1989. "Rural Industrial Location versus Industrial Growth", Annals of Regional Science, Vol. 23:59-68.

Statistics Canada 1996. "The Input-Output Structure of the Canadian Economy 1992", Catalogue no. 15-201-XPB, Ottawa, Ontario.

Thompson, G.L., B. Weller and E.W. Terrie. 1993. "New Perspectives on Highway Investment and Economic Growth", Paper No. 93-0597 presented to the 72nd Annual Meeting of the Transportation Research Board, Washington, D.C.

Acknowledgements
The authors wish to acknowledge the editorial contribution of Michele Babin.

THE SASKATCHEWAN TRANSPORTATION MODEL AND ITS USE IN REGIONAL TRANSPORTATION PLANNING

by Doug Neis, Paul Rachar and Daryl Nixon
Saskatchewan Highways and Transportation

BACKGROUND

Introduction

Perhaps the most important issue facing road authorities in Saskatchewan is the rapidly changing grain handling and transportation network. Changes are placing an increased burden on the rural road network during an era of government restraint and debt management. Changing grain transportation is the driving force behind the recent formation of Regional Transportation Planning Councils (RTPC s). The Saskatchewan Transportation Model (STM) was developed in 1992 as a planning tool to help the Saskatchewan Department of Highways & Transportation (DHT) and RTPC s prepare for grain transportation rationalization.

Regional Transportation Planning Councils (RTPC's)

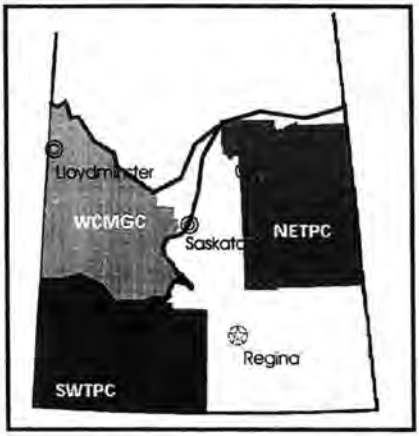
In their 1994 report, the Transportation Policy Council, appointed by the provincial Minister of Highways & Transportation and comprised largely of municipal government leaders, identified an opportunity to move decision making to the local level. Stakeholders who have to live with the decisions must be given the opportunity to play a major role in determining their transportation system. They recommended to develop a process to permit elected representatives of municipal government to participate more effectively in transportation related decisions. This recommendation led directly to the formation of RTPC s.

The emerging regional (as opposed to local) nature of grain movement has created an even greater role for RTPC s. Saskatchewan rural municipalities (RM s) are relatively small geographical areas. Changes

which occur outside an RM s boundaries can have significant social and economic consequences within that municipality.

The first RTPC formed on March 10, 1995 is the *Southwest* Transportation Planning Council, Inc. (SWTPC). On February 24, 1997, potential members of the *Northeast* Transportation Planning Committee (NETPC) agreed in principle to form an RTPC. The *West Central* Municipal Government Committee (WCMGC), dealing with all types of government issues, formed a Transportation Subcommittee in February, 1995. This subcommittee is considered to be the equivalent of an RTPC. The map in Figure 1 shows the boundaries of these three organizations within Saskatchewan s grain producing area. Interest has been expressed in forming councils in other areas of the province.

Figure 1: RTPC Boundaries



Urban and rural municipal representatives of the RTPC s work cooperatively together, and with DHT, to decide the future needs and priorities for the Saskatchewan road network. However, the ultimate responsibility still lies with each jurisdiction for its portion of the overall network. This partnership in planning approach keeps with the long-standing spirit of Saskatchewan cooperatives.

Changes in Grain Transportation

Saskatchewan, often referred to as the Canadian grain basket , accounts for 55% of the nation s export grain industry. Agriculture in Saskatchewan, of which grain is the primary component, accounts for by

far the largest percentage of the total economy of any province at about 12%. The great distance to ports means that grain transportation efficiencies are vital to Saskatchewan s economy in the global marketplace.

The grain industry has seen many changes over the last quarter century, and the industry is presently undergoing a massive rationalization. Abandonment of grain dependent branch lines, and consolidation of the grain elevator network are the key components of the rationalization.

Changes are occurring at a highly accelerated rate as a result of the August 1, 1995 elimination of the Western Grain Transportation Act (WGTA) and Crow Benefit rail freight rate subsidy, plus the new Canada Transportation Act (CTA), which provides a streamlined process for CN and CP to divest themselves of unprofitable branch lines. The following is a summary of changes affecting grain transportation.

Rail Branch Line Abandonment - The original 16,000 km rail system has been reduced to the present 5,700 km branch line system and likely will be shorter by the Year 2000. There is potential for a select few of the more viable branch lines to be saved by short line operators.

Grain Elevator Consolidation - The bottom line cause of increased grain traffic and resulting road damage is the increasing spatial distribution of grain delivery points. Consolidation would still be occurring to a large extent even without rail line abandonment, but the two issues do go hand in hand.

Older, smaller low throughput (LTP) wooden crib elevators, once commonplace to the prairie landscape, are being replaced by fewer, larger concrete high throughput (HTP) elevators[1]. The number of Saskatchewan elevators and delivery points have been steadily decreasing from 1,940 and 954 respectively in 1971, to 663 and 433 in 1996. With the accelerated rate of change, it s anticipated that there could be as few as 250 delivery points by the Year 2000, and less than 100 after the rationalization process settles down early in the 21st century.
HTP elevators provide added services, such as; grain cleaning, drying, blending, and agricultural information/resource centres. They are able to offer producers financial incentives to overcome higher trucking costs to attract greater volumes, due to multi-car loading incentives from the railways, inland processing advantages, and large economies of scale. It

[1] LTP s typically have a storage capacity of 2,500 - 6,000 tonnes and handle about 20K - 25K tonnes annually. HTP s have about 10K - 30K tonnes storage and handle 75K to 250K or more annually.

is questionable whether these incentives will still be offered once HTP elevators are the rule, and network rationalization is complete.

Larger Trucks - Longer hauls are more efficiently handled by larger (5+ axle) commercial trucks as opposed to 2 and 3-axle farm trucks. Once the grain is on a large commercial unit, the added cost to haul an incremental greater distance is small.

Inter-Elevator Transfers - describes the practice of moving grain from LTP to HTP elevators. In this case, LTP elevators act as satellite collection points for HTP elevators. Inter-elevator transfers are generally performed by truck, and can be even more damaging than if the satellite point didn t exist. Transfers often involve backhaul over producer haul routes, and are concentrated over short periods of time. This practice may be diminishing, as some major grain companies move to a best place first policy and procure grain directly off the farm.

Shift in Agricultural Activities - The loss of the $721 million/yr. Crow Rate Benefit, or two-thirds of the cost of moving export grain from elevator to port, is making other activities more attractive. Activities such as; domestic food processing of grains and oilseeds, livestock (primarily hogs), and specialty crops (potatoes, pulses, forage, etc.) are on the rise. These shifts will be most notable in central and eastern Saskatchewan, which are further from the port of Vancouver[2]. Some of these alternative activities are more damaging to roads than export grain is. Processing plants truck in raw materials from great distances, and truck out much of their product and bi-product. Specialty crops such as potatoes can be much higher yielding than grain.

Road Impacts

Saskatchewan has the most extensive rural road network in the country, with about 26,000 km of provincial highway and 160,000 km of rural road serving just over one million people. The system was developed with the intention of providing access to a close knit network of elevators and farm sites. Producers typically had no more than 10 to 15 km to haul to their nearest local elevator. This has resulted in a network of mostly low volume/low standard roads.

Thin membrane surfaces (TMS s), which are unique to Saskatchewan, are a thin (20 to 50 mm) asphalt mat placed directly on top of the subgrade. These surfaces were intended to provide a dust, mud, rock free

[2] The majority of Saskatchewan s export grains are shipped through the west coast with the growing dependence on Asian markets. Only the southeast corner of the province has Thunder Bay as their primary port.

alternative to gravel surfaces, but are not designed to withstand large volumes of heavy truck traffic. There are 9,360 km of provincial and some municipal TMS roads in Saskatchewan. Some of these serve as major feeders to new HTP elevator sites. The result of this is inordinately high maintenance costs and/or costly upgrading.

Incremental municipal road costs are highly uneven. RM s adjacent to new HTP elevators are bearing the brunt of the cost with none of the economic benefits from these facilities. Saskatchewan s RM s are too small to absorb large budgetary increases.

There are traffic impacts to consider other than road surface structural impacts caused by increased truck traffic. More trucks on narrow (mostly gravel and TMS) rural roads is a safety concern. There is also an environmental concern with greater fuel emissions.

System Optimization Approach to Grain Transportation Planning

Ideally, stakeholders should be working toward improving overall system efficiency. There are four primary stakeholder groups (producers, railways, grain companies and government) and four cost elements (trucking, elevation, rail and road) to consider.

Trucking Costs - are of primary concern to producers. HTP elevators offset trucking costs with incentives to encourage longer hauls.

Elevation Efficiencies - are of primary concern to grain companies. Handing costs alone can be deceptive in that HTP elevators provide more services, and can attract larger volumes to help increase market share.

Rail Costs - are of primary concern to the rail companies. In Saskatchewan that generally refers to the Class 1 duopoly of CN and CP. In some regions of the province either company holds a complete monopoly on rail service. Short lines can potentially operate a viable branch line service at a lower cost than can CN or CP.

Road Costs - are the primary concern of municipal and provincial governments. Producers are also indirectly affected through higher taxes. Unlike the other cost elements, road costs are not strictly user pay, especially when it comes to regional heavy truck traffic. Costs incurred by another jurisdiction have no impact on an individual producer s decision regarding where to haul. Grain and rail companies feel that government money will always be available for changing infrastructure needs. If it was practical to implement a true user pay system, perhaps road impacts would be given their due consideration.

Stakeholders have separate and competing interests. Naturally, each will make decisions in their own best interest. In the end, it is the producers who pay the price, but they also have the ultimate power if they act collectively. It is questionable whether efficiencies realized by the rail and grain companies through system rationalization will be passed on in part or in their entirety to producers.

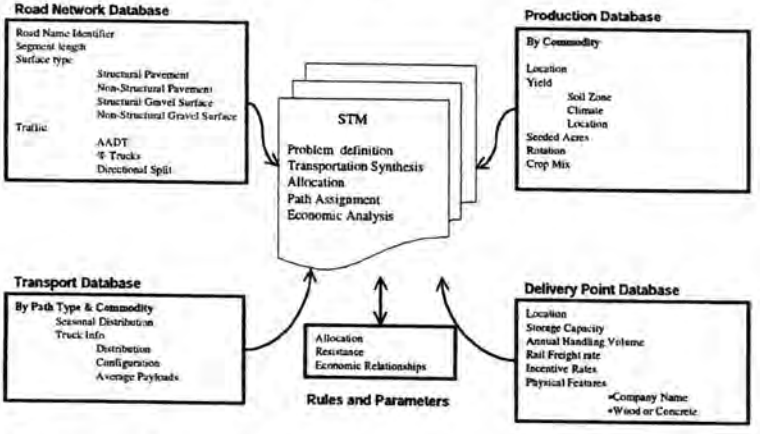
TRANSPORTATION MODELING

General Overview

In the early 90 s, DHT put together an initiative to develop a comprehensive transportation planning model called the Saskatchewan Transportation Model (STM). STM is essentially a set of GIS based analytical tools used primarily to quantify the costs associated with the structural changes occurring in the grain handling and transportation industry. This paper specifically makes reference to grain transportation modeling; however STM has been used to model other commodity traffic as well.

The model simulates origin to destination commodity traffic flow based on a given set of parameters or rules that determine routing. The commodity paths are stored in a database by path type, each path type is superimposed and accumulated for each road and highway link. The traffic modeling outputs can be passed on to economic cost models to estimate incremental system costs (road, elevation, rail, and trucking).

Figure 2: System Approach Used in Saskatchewan Transportation Model



Information Sources

The Saskatchewan Transportation Model incorporates several databases designed to represent origin and destination nodes, links for the rural municipal road and highway networks, characteristics of the transport system, and the rules and parameters that the model runs under.

Road Network Database- Provincial road network databases contain tabular information that includes road location identifiers, direction of travel, segment lengths, classification, surface type, and traffic volumes. The road network databases include both the provincial highway network and the rural municipal road networks.

Grain Production Database- The grain production database contains information specific in the calculation of crop volume. Grain production is calculated for each township location by crop type based on soil zone type, crop yields, crop rotations and mixes. Information is provided by Saskatchewan Crop Insurance, Prairie Farm Rehabilitation Administration, and Saskatchewan Agriculture and Food. Grain commodities include the seven major grains and oilseeds grown in Saskatchewan[3].

Delivery Point Database- Consists of elevator and agricultural processing information collected from the Canada Grains Commission and the Canadian Wheat Board. This information is supplemented by information provided by individual grain companies. The information includes delivery point locations, storage capacity, annual handling volume, rail freight rates, typical incentive rates, owner, and elevator type.

Transport Database- Consists of information that describes how and when the commodities are being hauled. The seasonality and mixture of truck types being used varies for each commodity and path type.

Rules and Parameters

The rules and parameters determine O-D commodity allocation and the paths that are taken. Rules and parameters are tied to economic relationships such as path resistances, rail freight rates, trucking rates, and elevator incentives.

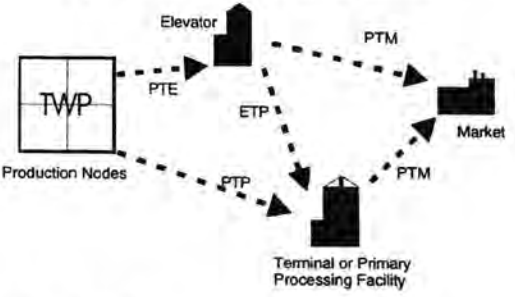
STM Modeling Components

Because of the scale and complexity of grain logistics, it is very difficult to simulate all the possible decisions that a producer will make in the delivery of grain. The approach used in the STM is to: 1) understand the major commodity movements, 2) research the physical constraints that limit commodity flows, and 3) develop the economic relationships that drive producer decision making.

[3] Wheat, Oats, Durum, Barley, Rye, Canola, & Flax

To simplify our understanding of grain commodity movements it is necessary to break it into components that are easily modeled. It s assumed that each commodity is hauled from the farm to the elevator collection points (PTE). A portion of the deliveries may a transferred from the country elevator system to the larger terminals located on the mainline or even specialty processing facilities (ETP), with the majority remaining on site for rail delivery to port for export (PTM). In addition, some of each commodity is also hauled directly from the production nodes to the larger terminals or specialty processing facilities (PTP). The concentration of haul and the distribution of trucks varies seasonally for each haul type. Truck size tends to vary with length of haul and haul type.

Figure 3: Showing Grain Logistics.



Problem Definition

The systems approach to grain transportation requires a clear understanding of the level of rationalization in both rail and elevation networks, as well as policy changes that may be occurring in agriculture that affect production (quantity and location) and processing diversification.

Problem definition provides the framework for synthesizing grain logistics using the Saskatchewan Transportation Model. GIS software is used to setup the origin and destination locations and transportation network options that will be identified as possible scenarios. The expected outputs are scenario arrays defining the changes projected to occur regarding network consolidation, transportation economics, and commodity diversification.

Transportation Synthesis

The model combines the spatial analysis capabilities of GIS with the data manipulation sophistication of commercially available software packages such as spreadsheets and databases. A shortest path routing algorithm is used within GIS to determine a matrix of origin-destination path string combinations for four different path categories. The path strings are determined for producer to elevator (PTE), elevator to processor transfers (ETP), producer to processor (PTP), and processor to market hauls (PTM).

Allocation

Once a path string matrix is created, each commodity is allocated to the elevator using a linear programming algorithm. When the commodity is allocated to the elevator system based on production and handling constraints, an allocation matrix is used to assign commodity flows to the path string matrix. The allocation of grain from origin to destination is defined in the allocation matrix. Since the allocation table controls the volume of commodity on each path, the model has the flexibility to consider different alternatives by changing the allocation

The linear programming model used to determine the allocation matrix has an objective function that minimizes differential cost to the producer. The producer s cost is the difference between trucking cost and site cost.

H = T + (R I)
T = Trucking costs (distance based)
R = Rail Freight rates
I = Elevator incentives such as blending, elevation, storage, screening, unit car benefits.

The objective function is minimized with regards to production and elevator capacity constraints:

Minimize $C = H_{ij} X_{ij} + H_{i+1\,j+1} X_{i+1\,j+1} + .$

$X_{ij} + X_{i+1\,j+1} + . \leq T_i$

$X_{ij} + X_{i+1\,j+1} + . \leq E_i$

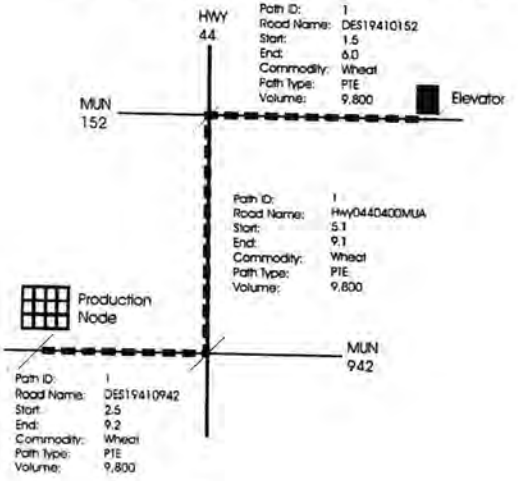
Where:

H The Producers cost of transporting grain from the township (i) to elevator (j).
X The tonnage hauled from township (i) to elevator (j)
T The production constraint at township (i)
E The capacity constraint at elevator (j)

Path Assignment

The *path string matrix* contains links to all road segments traveled for each path using a process called dynamic segmentation. The dynamically segmented road sections link to the road network databases by the road name identifier and by start and end kilometers. Through the allocation table, each path string is linked by path and commodity types to the production and transport databases

Figure 4: Origin-Destination Path string



Economic Analysis

Commodity traffic volumes generated by STM can be downloaded to a number of economic analysis models to determine road, rail, elevator, and trucking costs. Each of these models determine the total grain transportation systems cost.

<u>Trucking Cost Model</u>- The trucking cost model links vehicle-operating, labor, and administration costs to the truck distribution and segment haul lengths to estimate the cost of hauling grain for any given scenario. Vehicle operating costs per tonne-km for each truck configuration and surface type includes: capital depreciation, utilization, fuel, oil, maintenance, tire, insurance and registration.

<u>Elevation Cost Model</u>- The elevation cost model considers the average elevator operating costs and fixed costs for concrete and wooden elevators of different capacities. Operating costs include labor,

maintenance, utilities, and marketing. Fixed costs include capital depreciation, taxes, and insurance.

Rail Cost Model- The rail cost model is used to estimate the annual cost of operating either a class1 or short-line operation. The model considers the average operating and fixed costs per tonne kilometer needed to operate a branch line. The operating costs considered are: labor, fuel and oil, track maintenance, engine maintenance, administration, and marketing. The fixed costs considered are: annualized cost of rail at salvage value, depreciated cost of locomotive, taxes, and insurance.

Road Cost Model- The road cost model is an experience-based model that estimates incremental costs of maintenance and capital upgrading due to modeled grain traffic on each road segment. Capital improvement decisions are based on the probability that a given existing road classification is suitable to carry the increased truck traffic. Maintenance costs are estimated using experience based cost functions. All costs are annualized over a 20 year period.

REGIONAL TRANSPORTATION STUDY RESULTS

DHT has recently performed STM traffic modeling and study analysis for the SWTPC and WCMGC.

Southwest Transportation Planning Council Study

Only producer to elevator (PTE) grain traffic was modeled for the existing and two future rail/elevator network scenarios. A systems optimization approach was taken; considering road, rail, elevation and trucking costs for the handling of the region s 2.44 million annual tonnes of export grain.

Scenarios

Delivery points for each of the following scenarios are shown in the traffic flow maps in Figure 5.

Scenario 1 - represents the highest predicted level of elevator consolidation. Seven delivery points would serve the region, including two located outside the region. These points are located on the CP main and non-grain dependent branch lines.

Scenario 2 - represents a modified version of the existing scenario, with 45 of the existing 65 delivery points plus the branch lines serving these points assumed open. Many of the 20 eliminated points are quite small

and will be phased out shortly. This scenario provides a good benchmark with which to compare the two future scenarios.

Scenario 3 - represents a future scenario with a much lower level of consolidation than Scenario 1. There are 24 delivery points including the two outside the region.

Traffic Flow Maps

STM model accumulated grain traffic for the three southwest scenarios is shown in Figure x below. Note the much greater traffic volumes in Scenario 1, where grain is forced to move up to 150 km along N-S corridors from the Montana border to Swift Current and Gull Lake.

Figure 5: High Impact South West Road Sections Scenario 1

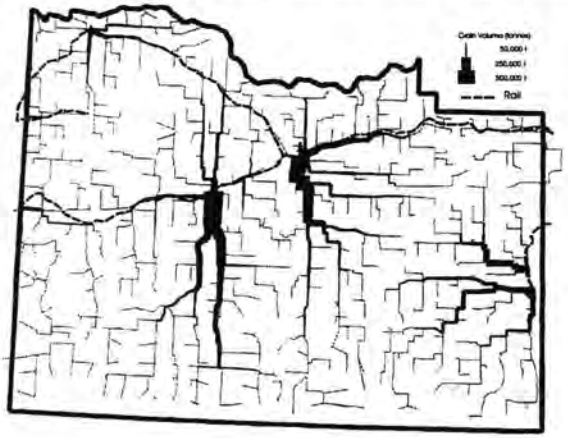


Figure 6: High Impact South West Road Sections Scenario 2

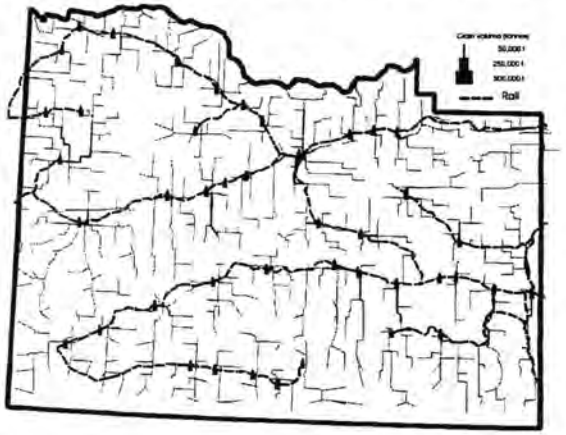
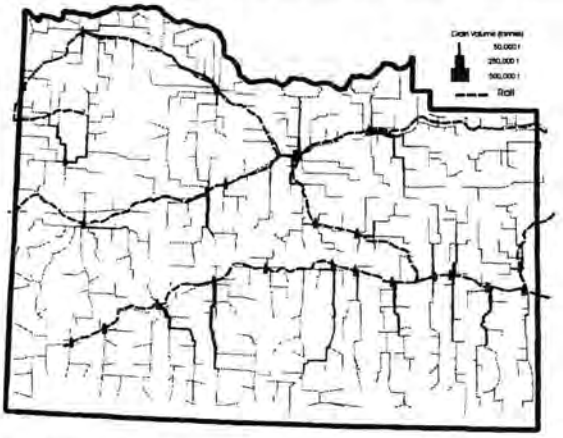


Figure 7: High Impact South West Road Sections Scenario 3



Systems Cost Analysis

Figure 8: Southwest Transportation Study Scenario Costs ($ million)

Scenario	1	2a	2b	3a	3b
Avg. Truck Haul (km)	66.8	24.6	24.6	32.6	32.6
Elevation	26.3	19.2	19.2	19.7	19.7
HTP Value Added	12.2	1.7	1.7	3.8	3.8
Net Elevation	14.1	17.5	17.5	15.9	15.9
Truck Cost	30.7	19.9	19.9	24.3	24.3
Rail Cost[4]	5.1	20.1	15.2	15.7	12.2
Subtotal	**49.9**	**57.5**	**52.6**	**55.9**	**52.4**
Road Impact Cost	14.3	8.6	8.6	11.6	11.6
Overall Net Cost	**64.2**	**66.1**	**61.2**	**67.5**	**64.0**

When road costs are not considered, Scenario 1 with the greatest level of consolidation appears most attractive. However, when road costs are added to the equation, a more moderate level of consolidation would appear prudent from an systems optimization perspective. Viable short lines have the potential to reduce operating costs in Scenarios 2b and 3b.

[4] Rail costs for Scenarios 2a and 3a represent estimated Class 1 carrier costs. Rail costs for Scenarios 2b and 3b represent estimated short line operating costs.

Further Work

The results of this preliminary planning study have led to initiatives in; short line rail and alternative loading methods feasibility, traffic analysis for other commodities (oil), and developing a road management plan for the region.

West Central Municipal Government Comm. Transportation Plan

Unlike the earlier southwest study, the west central study was only concerned with road impacts related to three issues:

Grain Industry Changes - elevator consolidation, rail line abandonment and transfers to inland grain terminals.

Food Processing - This considers growing value added domestic processing activity in areas such as; canola crushing, malt barley processing and flour mills, and is tied to the abolishment of the Crow Rate subsidy.

Increased Irrigation and Expanding Lucky Lake Area Potato Industry - Potatoes are generally far more damaging to roads than is grain (much higher yields, 100% trucked, concentrated hauls).

Although the scope of the project was to model as much agricultural traffic as practical, many activities of much lower order of magnitude could not be practically modeled.

Scenarios related specifically to grain elevator and food processing delivery points were developed for the Years 2000 and 2005. There was no existing scenario run to compare with the future scenarios.

The study s main objective was to determine which road sections are likely to be heavily affected by future developments related to the three stated issues. Figure 9 below provides an indication of which road sections are highly likely to require upgrading by the Year 2005, as predicted by STM traffic volumes and the road impact cost model[5].

[5] The present road impact cost model is used for macro cost analysis. It is not intended to be used to predict individual road performance.

Figure 9: High Impact West Central Road Sections (Year 2005)



This study was the testing ground for many of the enhanced STM processes described earlier, such as accumulating traffic for multiple commodities and path types, and applying seasonal distributions to traffic.

BIBLIOGRAPHY

1. A Methodology for Measuring the Impact of Rail Branch Line Abandonment on Provincial and Municipal Roads in Saskatchewan , ADI Limited, February, 1989.

2. A Description of the Grain Handling and Transportation System , DHT and Saskatchewan Agriculture & Food, for the Grain Task Force, December, 1992.

3. Formation, Organization & Operation of the Southwest Transportation Planning Council , DHT for SWTPC, May, 1996.

4. Southwest Transportation - Pulling Together, Moving Ahead , SWTPC Inc., August 15, 1996.

5. West Central Saskatchewan - Moving With the Times , WCMGC, January 16, 1997.

PERFORMANCE EVALUATION BY NORTH AMERICAN CARRIERS: PRELIMINARY THOUGHTS

by Mary R. Brooks
Dalhousie University

Abstract

Performance evaluation in the U.S. transport system is well underway according to the trade press; its use by Canadian firms is less well known. This paper will report some of the preliminary findings of the **second** phase of a three phase research study now under way to examine the assessment of carrier performance by users and suppliers of North American transportation services. Phase I examined the approaches of cargo suppliers (manufacturers) to the evaluation of their transport suppliers' performance, in order to understand the business practices taking place. Detailed case studies of user companies is the part of Phase I which remains to be completed. Phase II will examine similar elements from the viewpoints of carriers and logistics service firms. This paper focuses on the preliminary findings of the **carrier** component of Phase II and is intended to stimulate debate on the entire topic of buyer-seller relationships and the current realities of selling transport services.

Introduction

In Phase I, it was surprising to find very few differences between Canadian and American companies in terms of their evaluation of transport companies. It was also noteworthy that few differences existed between firms shipping high value goods and those shipping low value goods. However, there were significant differences between companies operating according to Just-in-Time (JIT) requirements and non-JIT companies in their use of quality programs and of formal carrier performance monitoring. From Phase I findings, it was concluded that firms involved in JIT processes are more likely to engage in carrier monitoring than those not involved. Phase I also found that, although QS9000 has been touted as a vehicle for streamlining vendor selection processes, even manufacturers with JIT systems have not wholeheartedly adopted it nor believe it to be a good proxy for carrier performance evaluation. Many transport buyers have developed their own in-house quality programs. Formal carrier

performance monitoring generally does not take place in smaller manufacturing concerns. Although audit was the most preferred method of supplier evaluation, and third party audits are generally available to smaller firms, third party audits were not widely used. This paper will examine the view of carriers with respect to the monitoring of their own performance and how they review that performance with the buyers of their services.

Literature Review

In the last half decade or so, the market for transport services has changed dramatically. Deregulation, time-based competition and the resultant focus by many buyers on JIT strategies have resulted in the buyer-seller (shipper-carrier) relationship for many companies being a partnership. This trend is reinforced by the implementation of global production and distribution systems by many transport buyers. Time-defined or other forms of contractual relationships (including strategic alliances) govern an increasing number of buyer-seller relationships. There is a move away from price-oriented transaction-specific carrier selection and towards alliances and negotiated transport service contracts (Brooks, 1995a; Gibson *et al*, 1993; Crum and Allen, 1991; Kleinsorge *et al*, 1991). Lieb and Millen (1988, 1990) noted that JIT implementation leads to changes in criteria by which carriers are selected and in the number of carriers used. Phillips (1991) concluded that more innovative contractual arrangements will continue to devel op. Bowersox and Daugherty (1995, p. 77) noted that "firms can be expected to place greater emphasis in the future on monitoring and measuring performance." Therefore, this research investigates the processes of carrier performance monitoring and accompanying performance measures (or metrics) used.

Transport performance literature generally takes either a macro- or microeconomic viewpoint. The work of the Transport Studies Group of the University of Westminster (two undated studies) focuses at the industry level. Likewise, Data Envelopment Analysis (DEA) has been used to track the efficiency of carriers' performance over time (Kleinsorge *et al*, (1991, 1992). These are not, however, appropriate for firms seeking to conduct a self-assessment of their performance in order to implement a strategy of continuous improvement. There have also been a number of articles on performance measurement from the perspective of the purchasing manager, often drawing conclusions about the types and effectiveness of those systems in place (e.g., Dumond, 1991). This work has generally not focused specifically on the purchase of transport services.

However, both the trade and academic press remain unclear as to how widely adopted either logistics management or supply chain management has been in Canada. There is a well-established and growing body of literature in the logistics performance field; metrics such as order fill ratios, order cycle times and the like are commonly used to measure performance. It is the author's view that carrier performance evaluation is one tool used by supply chain or logistics managers to acquire, secure or maintain competitive advantage and that these management fields are much larger in scope that of transport performance evaluation. This research is narrowly focused on transport performance measurement at the firm level.

A review of the literature points to two areas of research relevant to this project—carrier selection criteria and carrier performance measures. Generally, most research has been focused on the former, and predominantly based on shipper surveys. The criteria by which shippers choose carriers have been studied for over two decades and these studies were broadly reviewed, from the shipper point of view, in the paper prepared on Phase I (Brooks, 1995b). The reason for research emphasis and interest in carrier selection criteria usually stems from the relative importance individual carriers place on the **findings** of the research; understanding of shippers' selection criteria can be used by a carrier to adjust or redirect business strategies and tactics. Knowledge of the standard of service desired better enables the firm to meet the customer's expectations and secure the continuance of the business. However, if selection criteria are a moving target over time and vary significantly between segments of the market for a single mode, as found by Brooks (1995a), this type of research only provides a snapshot, albeit one that can be used to develop the variables examined in bench-marking a firm's performance in a planned program of continuous improvement.

The NEVEM Working Group (1989) present a process for evaluating logistics performance, noting that logistics objectives have to be set (delivery or throughput time, delivery reliability, flexibility and inventory level) and that, as these are inherently contradictory, tradeoffs will need to be determined as part of the optimization process. Furthermore, they note that a system of performance indicators must allow for problem analysis and assignment of responsibility for the problem. What is not clear from their work is the extent to which it has been put into practice or whether it is at all suitable for smaller firms.

As for specific literature on carrier performance metrics, however, there has been limited work done, predominantly with a modal focus. As an example, Tyworth and Wu (1993) examined the value of standards conformance in the area of on-time shipments by less-than-truckload carriers.

Some in the industry argue that the relationship marketing literature should be explored. They note that the marketing of transport services has always been a traditional "relationship"-marketed service because of customer reliance on carrier sales representatives and the intangibility of the service. O'Neal (1989) concludes that relationship marketing is more likely to be effective under the following conditions: a long time-frame, high supplier switching costs, high perceived risk, substantial investment and intangibility of the service. Altho ugh the lastobviously applies to all transportation purchases, for a significant volume of these purchases the time-frame is short, switching costs are low, risk is low for low value goods, and there is minimal investment. These factors move closer to the effective relationship marketing end of the spectrum when JIT operations are introduced, goods have a higher value, and delay has significant consequences. There is no doubt that some carriers practice relationship marketing but it serves as only one use of the information flow resulting from a performance assessment effort on the part of a carrier, and not necessarily the driver of performance measurement and monitoring.

In sum, future research into carrier performance must focus on: (1) the performance measures used to evaluate carrier performance, and (2) the processes used for collecting and evaluating such data so that cost-effective means may be found for smaller firms to take advantage of such practices. The research reported in this paper begins to scrape the surface of what needs to be undertaken with respect to both. The paper identifies the incidence of performance evaluation on the part of carriers as well as reports on the broad performance measures in use.

Methodology

In the overall research program, three groups—buyers, carriers and logistical service firms—are being contacted in two countries, Canada and the U.S.A. In Phase II a random quota sampling approach targets 240 firms in each of four groups: Canadian carriers, U.S. carriers, Canadian logistical service firms and U.S. logistical service firms, all to be contacted by mail with a questionnaire.

The target of 240 Canadian carriers was not met because there were not sufficient unrelated firms found. (Continued random sampling only yielded duplicate records, e.g., another branch office of the same company.) This paper reviews the approach taken with carriers (all modes), as logistical service firms have not yet been contacted. However, it was interesting to discover that 53% of the carrier respondents indicated that they act as logistical service firms! The response results are shown in Exhibit 1, while Exhibit 2 illustrates the modal spread of the respondents.

The response rate is within acceptable limits. Altho ugh it is intended that a number of research hypotheses be tested, the carrier subset is too small to draw definitive conclusions about the differences between Canadian and U.S. carriers and between those with monitoring programs or not at this point in the Phase II research. The findings reported here are largely descriptive—defining the incidence of and practices in carrier performance monitoring.

Exhibit 1: Response Experience

	Canada (1)	**U.S.A. (1)**
Questionnaires Mailed	215	240
Usable Questionnaires Returned	55	38
Questionnaires Returned but Deleted From Analysis (2)	31	
Other Returns (3)	24	42
Replied, Refusing to Participate (4)	31	18
Non-respondents	102	141

Notes: (1) One questionnaire mailed to a Canadian firm was returned by its U.S .parent. One questionnaire mailed to a U.S. firm was returned by its Canadian parent.

(2) Three of these were deleted as a review of the responses proved they were agents for ports or carriers without operating influence. One was deleted because two questionnaires were returned by one company.

(3) Address Unknown, Moved with Expired Forwarding Order or No Longer Active as a Business.

(4) A number of companies refused because of company policy, insufficient volume to be interested, or they did not believe the study was relevant to their business.

Exhibit 2: Respondents By Mode

Mode (1)	Canada	U.S.A.
Truck	39	33
Ocean	22	17
Intermodal	13	17
Rail	14	12
Air	14	9
Other	3	1
Total (2)	55	38

Notes: (1) Those companies also supporting third party logistical services totalled 53% of respondents.
(2) The totals do not add as more than one mode could be checked. Respondents checked an average of 2.07 modes.

Findings—Performance Evaluation Practices

Transport supplier evaluation

A total of 93 questionnaires were analyzed—55 from Canadian companies and 38 from American companies. Both geographic groups overwhelmingly (in excess of 85% of firms) supported the notion that transport carriers should formally monitor their own performance independent of whether their customers request such monitoring. A significant percentage of the carriers believed monitoring should be done by the purchaser of the transport services; this point of view was expressed by 36% and 42% of Canadian and American firms respectively. Less than 10% of respondents were of the opinion that performance evaluation should be conducted by a third party audit, either at the customer's or carrier's discretion. The derogation of third audit by carriers is not surprising given its lack of favour with users as well (Brooks, 1995b). Some carriers expressed the view that carriers benefit from performance monitoring through continuous improvement, and one observed that "all service suppliers should be confident enough to publish their standards, allowing competitors and customers to monitor service success." It was added that "If it cannot be measured, it cannot be managed."

As seen in Phase I, more than one method of evaluation is perceived as useful for evaluating carrier performance. When questioned about the use and type of formal program to evaluate its own performance, 80 of the 93 firms indicated they have a formal evaluation program in place (89% and

82% for Canadian and American firms respectively). Three-quarters of the firms with formal monitoring programs have developed their own in-house program. Contracted third party audits were seldom used, while spot monitoring and customer programs were each used by one-third of respondents. Other suggestions included service failure reports (completed by the sales department), satellite tracking and customer service surveys.

Of those users monitoring carrier performance, Phase I found that two-thirds used audits and a similar percentage used process reviews. Therefore, it was decided to ask carriers who monitored their own performance about their customers' practices. The number of users who continuously monitor carrier performance was estimated by the carriers to be 37%. Carriers estimated the percentage of customers using spot monitoring at 44% (for both Canadian and American customers). Carriers did not believe that many of their customers used third party audits to monitor performance, with only 9% of firms estimated to be using this technique. Only one Canadian firm believed that none of its customers monitored its performance. It is noteworthy that some firms did not feel comfortable providing, or could not provide, a percentage estimate of their customers' behaviour. (Sixteen companies did not answer the question or provided only partial answers.)

The question posed to the survey participants with formal monitoring programs was "how often should your performance be reviewed with your customers?" The overwhelming responses were customer-focused—as often as the customer warrants because it depends on the customer. (See Exhibit 3.) Of those recommending periodic reviews, quarterly reviews are the most favoured with monthly reviews preferred by 22% of respondents. Annual reviewsreceived little support, although some respondents thought that all customers should receive feedback at least annually.

In contrast, all carriers were asked how frequently users seek to review their performance. Most notably, carriers believed that reviews are primarily sought by transport buyers when there is a problem with performance (37%) or when sales staff are meeting the customer (20%). Monthly reviews were also indicated (22%). (See Exhibit 3.) While there was little consensus on what is an appropriate time-frame for conducting periodic regular reviews, and it appeared that many carriers believed shippers seek reviews only when performance is not acceptable, monthly reviews were mentioned by one out of five respondents.

Exhibit 3: Performance Review Timing

Frequency with which (1)	Carriers should review performance with customers (Q7)	Customers seek a review (Q9)
As often as the customer warrants	59%	NA
Review performance only when there is a problem	NA	37%
Monthly	22%	22%
Only when they are meeting with the carrier's sales personnel	NA	20%
Quarterly	33%	16%
Other (2)	8%	18%
Annually	4%	13%

Note: (1) Totals do not equal 100% as many respondents checked more than one.
(2) Other included different timing alternatives.
NA=Not asked. The wording of the two questions was not entirely similar given the different focus of the questions.

The responding firms that used a formal program to evaluate transport performance were then asked to report the method(s) employed to evaluate such performance. The majority of respondents, 91% (consistent for both geographic groups), indicated the use of a service performance measurement for performance evaluation. Customer complaint levels were used for evaluating performance by 57% (49% and 71% for Canadian and American firms respectively) while process review was the third most popular method of evaluation, with 55% of respondents using this method (49% and 65%, Canadian and American respectively). Alternative methods of performance evaluation, such as analyzing claims ratios or sales growth, satellite tracking, customer surveys, and so on, were s uggested by 20% of firms, while performance measurement by third party audit was used by a mere 14% of responding carriers.

When responding carriers were asked if they employed a formal system for reporting to their customers on service performance, 39% of Canadian and 52% of American carriers responded affirmatively about using a report card system, while 43% of Canadian and 35% of American firms did not employ any type of formal reporting system. Of those carriers using a report card system, only a few did so for all customers; report cards tended to be used for key customers (56%) or for those requesting them (35%). The form of "report cards" may be as varied as EDI or other computer-generated

transit time or delivery time reports to peer audits to daily conference calls with key customers. The open-ended part of the question generated as wide a variety of responses as there were respondents, but indicated significant ability on the part of many carriers to develop a custom solution for the customer demanding specific information. Less than 20% of either Canadian or American carriers indicated a formal reporting system other than the report card approach.

For those firms using a report card system for performance assessment, 82% of Canadian and 100% of American carriers used the system as a method of evaluating and/or measuring process improvement. Only slightly fewer used the system as part of a sales strategy to clients or to provide a formal feedback mechanism. A report card as a tool for general discussion ranked fourth while one used to document bonuses, penalties, service-related guarantees or items (including rates) for negotiation ranked fifth and sixth.

Perhaps most interesting is the reporting of formal carrier performance monitoring by company size and how that differs between users and carriers. In the study of **users** (Brooks, 1995b), it was clear that formal monitoring was more likely to be undertaken by large rather than small companies. In that study, none of the companies reporting sales of under $5 million conducted formal monitoring of carrier performance; only seven users with sales under $25 million undertook any type of formal performance monitoring with inspection as the prime method. As can be seen in Exhibit 4, even small **carrier** companies participated in formal monitoring, although their use of report cards (29% use them) as a method is less than that of the group as a whole. A review of these small carrier companies showed 81% had in-house performance evaluation programs with 43% conducting spot monitoring. The widespread use of performance evaluation in support of improvement in operations was noteworthy.

Exhibit 4: Monitoring Activity By Size of Firm

Number of employees	Total number of companies	With formal monitoring	With report cards
Less than 100 employees	39	29	10
100 to 499 employees	16	14	10
500 to 999 employees	13	13	11
1000 to 3000 employees	7	7	6

More than 3000	18	17	11
No Answer	0	0	0
Total	93	80	48

Annual sales	**Total number of companies**	**With formal monitoring**	**With report mards**
Less than $25 million (1)	28	21	7
$25-99 million	33	28	19
$100-249 million	7	7	6
$250-999 million	9	9	7
$1 billion or more	13	12	8
No answer	3	3	1
Total	93	80	48

Quality program activities and perceptions

Traffic Management's survey of its readers in 1992 (reported in the May 1993 issue) indicated that 7 out of 10 manufacturing companies had a quality program. This high percentage was not supported in Phase I where only 44% of respondents indicated participation in this type of program (Brooks, 1995b). Of the 93 responding carriers in Phase II, 51% indicated their company currently participates in a formal "quality" program or process. This was true of both Canadian and American subgroups. Only 6% indicated that they participated in such programs in the past but had discontinued the practice, while 43% responded that they never participated in a quality program at any time. Reasons for discontinuance were varied but similar to those noted by users: lack of internal support, too busy with other activities, or dissatisfaction with results.

Carriers from all modes were represented in the users of quality programs. Of those companies that participated in a formal quality program or process, by far the two most popular types of programs indicated are in-house systems (representing 58% of the group), and ISO 9000 quality assurance (indicated by 40% of the group). More American firms currently use in-house programs than Canadian carriers. Other quality systems are used but not in substantial numbers: Malcolm Baldridge by 12% of this group, contracted third party audits by 6%, customer or industry association designed programs by another 8% ,and other programs, such as TQM,

Philip S. Crosby Quality College, statewide quality awards, and so on, by 12%. Some carriers had more than one quality program in place.

Respondents were next asked if their customers required them to participate in such programs, and/or what percentage of customers did so. No carrier responded that all its customers required participation in such programs, while only one American firm indicated that more than 75% of customers require it to do so. Of those currently using a quality program, 25% indicated they do not know what percentage of customers demand participation in a quality program while 35% estimated that it is less than 25% of their customers. The remainder were split evenly between the two categories: 50-74% or 25-49% of their customers require quality programs.

Carriers' attitudes are not as derogatory of ISO 9000 as were those expressed by users. Generally, responses were positive but indicated that many users inquired about ISO 9000 status but did not require certification to conduct business with the carrier. One carrier noted that if the service performance is solid, customers will use them without formal certification. Another carrier observed that European trades seek ISO 9000 but it is not important for Asian business. A number of carriers reported that they focus on those quality processes important to their key customers while others recorded find that customers did not demand a particular quality process.

Other Findings

Companies were also asked about their use of EDI systems with major customers. Half of the respondents stated that they used EDI systems for all customers who wished to participate, while another 21% indicated they only used EDI systems with major customers. The remainder replied that they did not offer EDI capability.

Exhibit 5: Most Commonly Identified Service Elements for Monitoring

Service Element	Canada	U.S.A.	All
On-time delivery	57%	65%	60%
Accurate billing	49%	61%	54%
On-time pick-up	47%	52%	49%
Loss/damage experience	49%	45%	48%
Accurate documentation	41%	42%	41%
Equipment cleanliness	33%	39%	35%
Right equipment supplied	31%	32%	31%
Cycle time	16%	26%	20%

Note: By those companies participating in carrier performance monitoring programs. In no particular order, other service points included telephone accessibility/ease of contact, quality of telephone response/telephone etiquette, response time for booking or confirmation, call-back times, accounts receivable payment history, timely quotations, actual versus advertised transit times, wait time, driver appearance/attitude, and equipment delays. Those reported only once have not been listed.

Findings—Performance Measures

There was remarkable consistency between Canadian and U.S. carriers on the key service points that should be included in an effective monitoring system. On-time delivery is clearly critical, as is accurate billing. However, as can been seen from Exhibit 5, a wide range of metrics are used to measure performance. In addition, 16 firms (8 Canadian and 8 American carriers) indentified more than 10 **additional** service points that should be included in an effective monitoring system.

But what specifically is measured by carriers? Only 6 carriers provided "report cards" for detailed review. One of those was not an analytic report card as it did not measure discrepancies in any manner but reported to the buying company the usual shipment information, including pick-up date, due date and undelivered shipments. A second report card rated document accuracy of the carrier's subcontracting or interlining partners as well as the general service provided. The former was rated subjectively on a weekly basis and the results were plotted (in a graph) over a period of time; the latter was a descriptive report to be signed by both carriers. The third, fourth and fifth report cards were customer satisfaction surveys, relying on the customer to evaluate the carrier's performance on subjective factors. Only the sixth was a detailed report of its own performance across a very large number of measures in four categories—customer satisfaction, safety and compliance, operational efficiency and employee satisfaction. Each of these categories was weighted and, within each, specific factors were assigned weights. The specific factors were measurable and included, for example, percent of documents handled on time, percent of documents with billing errors, booking desk availability, cargo shipped as booked, percent of claims processed within 30 days, number of hazmat violations, and so on. It appeared from the structure of the report, however, that it was

intended only for internal use and not for distribution to customers as a report card.

Conclusions

On the basis of the survey of users of transport services, Phase I concluded that many of the carrier performance measurement systems used by manufacturers do not appear to be that well established scientifically, and remain the purview of large companies The results of the carrier survey are not similar. Carrier participation in systems monitoring own performance is high and the size of firms undertaking performance evaluation is quite diverse (Exhibit 4).

In Phase I, manufacturers often relied on subjective evaluation. The number of manufacturers employing quantifiable objective metrics was small and the metrics used appeared to be very company-specific. This was also true of the carrier respondents supplying report cards for review. The reliance of carriers on customer surveys, based on subjective evaluation was noticeable.

The carrier performance evaluation systems used by manufacturers appear to have caught the imagination of the trade press while it appears that carriers, on the other hand, recognize the importance of such evaluation for long-term survival and quietly incorporate it into their business processes.

Acknowledgments

The author would like to thank student research assistants Patty Hope (MBA '96), Robert Hollis (MBA '97) and Stephen McSweeney (MBA '97) for their assistance in managing the logistics of the research data collection and analysis.

References

Bowersox, Donald J. and Patricia J. Daugherty (1995), "Logistics Paradigms: The Impact of Imformation Technology," *Journal of Business Logistics*, 16, 1, 65-80.

Brooks, Mary R. (1995a), "Understanding the Ocean Container Carrier Market—A Seven Country Study," *Maritime Policy and Management*, 22, 1, 39-50.

Brooks, Mary R. (1995b), "Performance Evaluation in the North American Transport Industry," Transportation Research Forum Annual Conference *Proceedings*, 1, 226-43.

Crum, Michael J. and Benjamin J. Allen (1991), "The Changing Nature of the Motor Carrier-Shipper Relationship: Implications for the Trucking Industry," *Transportation Journal*, 31, 2, 41-54.

Dumond, Ellen J. (1991), "Performance Measurement and Decision-Making in a Purchasing Environment," *International Journal of Purchasing and Materials Management*, 27, 2, 21-31.

Gibson, Brian J., Harry L. Sink, and Ray A. Mundy (1993), "Shipper-Carrier Relationships and Carrier Selection Criteria," *The Logistics and Transportation Review*, 29, 4, 371-82.

Kleinsorge, Ilene K., Phillip B. Schary and Ray D. Tanner (1991), "The Shipper-Carrier Partnership: A New Tool for Performance Evaluation," *Journal of Business Logistics*, 12, 2, 35-57.

Kleinsorge, Ilene K., Phillip B. Schary and Ray D. Tanner (1992), "Data Envelopment Analysis for Monitoring Customer-Supplier Relationships," *Journal of Accounting and Public Policy*, 11, 4, 357-372.

Lieb, Robert C. and Robert A. Millen (1988), "JIT and Corporate Transportation Requirements," *Transportation Journal*, 27, 3, 5-10.

Lieb, Robert C. and Robert A. Millen (1990), "The Responses of General Commodity Motor Carriers to Just-in-Time Manufacturing Programs," *Transportation Journal*, 30, 1, 5-11.

NEVEM Working Group (1989), *Performance Indicators in Logistics: Approach and Coherence*, Bedford, UK: IFS Publications.

O'Neal, Charles R. (1989), "JIT Procurement and Relationship Marketing," *Industrial Marketing Management*, 18, 1, 55.

Phillips, Laurence T. (1991), "Contractual Relationships in the Deregulated Transportation Marketplace," *Journal of Law and Economics*, 34, October, 535-64.

Traffic Management (1993), May.

Transport Studies Group University of Westminster (undated), *Road Transport Industry—Financial Trends*, London: RHA Financial Services.

Transport Studies Group University of Westminster (undated), *Performance Trends in Road Transport*, London: RHA Financial Services.

Tyworth, John E. and Haw-Jan Wu (1993), "Conformance-To-Standard for On-Time LTL Transportation: How Much Is It Worth?" *Journal of the Transportation Research Forum*, 33, 1, 109-28.

Re-engineering the Logistics of Canadian Grain Handling, or "Back to the Sacks?"

"May you live in interesting times". Ancient Chinese curse

Dr. Barry E. Prentice, Director, Transport Institute
Jake Kosior, Research Associate, Transport Institute
Jim Thomson, Mercatus Ventures

Introduction

Complacency in business is always risky, but it is particularly dangerous during periods of economic transition. Rapid changes in global trade are being spurred and re-oriented by growing world populations and rising incomes. Competitive advantage is being sought through technological advances in computers, genetics, robotics, telecommunications, and transportation. Governments are intensifying competition with policies of deregulation, "free" trade, and privatization. The drivers of structural change have seldom been more diverse, or more profound. To survive, businesses are being forced to re-engineer their manufacturing and logistical processes to lower costs and increase quality.

A logistics process is a sequence of value-adding activities that puts the right product in the right place at the right time. The best logistics process is the one that achieves this mission at the least total cost, and maximum customer value. Any logistical process that has been in place for some time is a likely candidate for re-engineering.

Re-engineering a logistical process involves radical redesign and the use of technology to achieve a quantum improvement in supply chain performance. This paper explores the opportunity to re-engineer the logistics of the Canadian grain handling system. The first section sets out the historic development of the current bulk handling system. This is followed by a discussion of the economic advantages and disadvantages of bulk handling, and the role of the grading system. Subsequently, the analysis examines five principals of logistics that make the containerization of grain a competitive alternative. The paper concludes with a discussion

of a new institutional framework (electronic markets) that would facilitate the re-engineering of the grain trade.

Re-engineering Grain Handling in the 19th Century

Before 1850, most grain was marketed in sacks and depended almost entirely on water transport. Sacks had certain advantages given the logistics of the period. Sacks of grain could fit into the awkward spaces of the river boats and could be carried on a man's shoulder across gang planks, down a set of stairs and along narrow corridors. The disadvantages of sacks were equally compelling. Handling was very labour-intensive and consequently expensive. There were few opportunities for economies of size and the water routes were circuitous and slow.

Marketing grain in sacks incurred high transactions costs. No buyer would purchase grain sight unseen. This meant that each lot of sacks had to be kept separately, with a corresponding paper trail. All risk of physical loss, as well as the risk of a price change, was borne by the shipper. Consequently, freight insurance was a major cost associated with handling grain in sacks (Cronon, 1991).

Re-engineering of grain handling in the 1850s was caused by technological and institutional changes that accompanied the invention of the telegraph (1844) and the expansion of the railway. The speed of oncoming trains had to be very slow with a manual system to signal an approaching train. The speed of the telegraph signals enabled the trains to run faster and safer. This encouraged the building of rail lines over longer distances and increased their competitiveness for freight. Only 2,800 miles of railway track had been built by 1840. The railway network expanded to 9,000 miles within six years of Samuel Morse's invention.

A method of handling grain in bulk was demonstrated by Joseph Dart at Buffalo, NY, in 1842. The railways were early promoters of bulk handling[1] because it reduced their labour costs and sped up the loading of railcars. Lack of an accepted grading system impeded bulk handling of grain. As long as ownership and value were distinguished by individual lots of grain, commingling was impossible. The first grading system for grain was introduced in 1856 by the Chicago Board of Trade. The ability to mix lots into a fungible commodity eliminated the need for buyers and sellers to be

in physical proximity to the product, or to each other, when conducting their transactions.

The superior speed and lower cost of the railway would have increased bulk transport in any case, but the impact of the telegraph on transactions costs, sealed the fate of traditional grain handling. Information about prices, quality and quantity moved at the same speed as the transport of goods prior to the telegraph. Prices became closely linked between surplus and deficit regions with the advent of "electronic" communications. This greatly reduced the price risk of shippers and enabled buyers to purchase when price was most advantageous.

Bulk handling, grading standards and the telegraph were the prerequisites for a commodity futures exchange. The futures market came into being over the period 1853-65. Once established, traders could make transaction decisions in remote markets based on telegraph quotes of price and grade. Telegraph-enabled buyers and sellers could hedge their transactions risk by committing to firm prices for future delivery.

The grain handling system in North America was well on its way to being completely re-engineered by the end of the American Civil War. Sacks were still used (especially for the movement from farm to country elevator) but the superiority of bulk movements was evident. The costs of trading and transporting grain from the interior of the continent fell dramatically and the volume of trade[2] increased exponentially. This opened the settlement of the great plains, and ultimately, the prairies of western Canada.

Economics of Bulk Handling Systems for Grain

The cost advantages of bulk handling stem from automation and economies of size. Mechanizing grain loading and unloading has continued to increase labour productivity with every new generation of elevator. The limit to country elevator size is determined by the costs of collecting grain from more distant farms, the density of production and the number of competing facilities.

The economies of size in movement are associated with equipment utilization and labour savings. The efficiency of bulk handling increases with the use of specialized equipment like grain hopper cars, which are faster to load and unload. Rapid car turn around increases the utilization

of the railway. This can be further improved by increasing the number of cars assembled at bulk loading sites. The ultimate efficiency in bulk handling is a unit train (104 cars) that has one pickup and one delivery.

The savings in packaging costs is a further important advantage of bulk handling. No packages are needed in bulk shipments because the product takes the shape of the shipping vehicle. Hence, a double benefit: avoidance of the costs of packaging and the cost of shipping the package.

The disadvantages of bulk handling are related to quality and inventory costs. Handling grain repeatedly damages its inherent quality. Physical handling splits some kernels and abrades the seed coat. It also creates the potential for dangerous dust explosions. The necessary dust control measures at warehousing and loading facilities add to the cost of bulk handling.

Commingling grain creates an average quality product. The average quality is based on the lowest common denominator specified in the grading system. This leaves no incentive for any improvement in the system. If the grading standard calls for one percent foreign matter, no less will be provided. For example, grain companies will add back dockage to meet the maximum grade allowance.

Bulk handling systems tend to have empty backhauls. The specialized design of bulk handling equipment makes it difficult to move products in both directions. Terminal elevators are designed to load grain to export, not to receive imports. Most grain hopper cars return empty to the country elevators. The bulk ships often arrive at the port in ballast. All the transportation and handling costs of the bulk system must be paid by the fronthaul shipper.

The lower unit costs of bulk transportation are traded off against the higher costs of financing and storing large pipeline inventories. A 45,000 tonne grain boat needs at least this quantity of the correct grade(s) assembled at the port. Any delay in assembling sufficient inventories can lead to large demurrage charges.

The cost of financing pipeline inventories depends on the velocity of marketing. The slower the inventory turnover, the higher the holding costs.

Risk of physical loss due to spoilage and pests increases the longer grain sits in the pipeline.

The advantages of bulk movements stem from the economies of size in handling and transportation. These savings are traded-off against the losses in product quality and higher costs for inventory and storage. The direct, or "visible" costs of bulk handling have determined its dominance. The indirect, or intangible cost trade-offs, however, are becoming increasingly important. The negative impacts on marketing are subtle. The key to these impacts is the grading system.

Economics of Grading Systems for Grain

The grading system was developed when information was expensive to collect and costly to transmit. The information pertaining to the quality attributes of grain could be reduced to a single code (e.g.1CRSW). Grading lowered transactions costs when telegraph messages were charged by the word. This benefit of grading has evaporated now that voluminous information can be stored electronically and transmitted at negligible cost (e.g., Internet).

The grading system created a fungible commodity that reduced clerical work. There was no need to keep paper records of each lot, with multiple carbon copies filed and stored. A simple warehouse receipt with a grade specification was all the paper documentation needed to claim and exchange ownership. Grading reduced the effort and increased the speed of order processing.

The clerical benefits of grading have also diminished as computer technology has advanced. Electronic Data Interchange (EDI) and bar codes have revolutionized information processing. Couriers sort and track millions of individual envelopes on a global scale. Customs have electronic clearance procedures in which no paper documents are required. Just-in-time manufacturers use EDI to issue orders, receive and process purchases and make financial settlement without human intervention. The difference in processing cost between one large order and a thousand small orders has become inconsequential.

Grading systems have opportunity costs. Market prices of grain adjust to the lowest common quality dominator. Blending better quality grain with

lower quality product is a commonly accepted practice. This may raise the value of the lower quality grain, but it denies the benefit to the producer of higher quality grain. The benefits of superior quality are lost through commingling, or captured by grain handlers.

The grading system in Canada requires new varieties of wheat to be visually distinguishable from the existing CRSW varieties. This condition is based on the need to maintain the quality reputation of these milling wheats. High yielding strains of lower protein wheat varieties have been prohibited because they could be mistaken in the bulk handling system. The opportunity cost of limiting the genetic potential of Canada's wheat breeding programs is significant.

Consistent quality has become more important as food processors have become more sophisticated. The quality attributes identified in grading systems are becoming less relevant. Millers observe that individual grain varieties of the same grade react differently. Processors are required to adjust their formulas continuously to compensate to differences in the "average" quality. The benefits of "consistent" quality are lost through commingling varieties. Some buyers are beginning to request specific grain varieties to minimize the variances in the grade classification.

Strategic Concepts of Logistics for Re-engineering

Innovations in transportation and communications create opportunities for developing improved logistical systems. The alternative to handling grain in bulk is to ship grain in ISO containers. Containers would be loaded at the farm, or at a consolidation facility, e.g., a country elevator. Containers would be trucked to double-stack train terminals, and forwarded to marine container ports. Grain containers would move with other containerized cargo to foreign buyers. At no point, except loading at a country elevator, would the containerized grain share the bulk handling system.

A key problem in re-engineering logistical systems is to escape entrenched ideas. Logistical theory has developed strategic concepts that are useful in assessing the prospects for re-engineering marketing channels (Ballou, 1992). Five strategic concepts of logistics theory are used to argue why the grain handling system should be re-engineered to move grain in containers.

Mixed systems are superior to pure systems.

Congestion cost theory supports the principal of mixed systems. As increasing volumes are forced through any fixed capacity, congestion will lead to diminishing returns. Congestion can be reduced by expanding capacity, but marginal costs tend to increase as new bottlenecks are encountered. Shifting some activity to an alternative system can lower total costs.

A mixed system is always superior if the process is subject to fluctuating volumes. The low utilization of the fixed capacity during the off-season can make it less expensive to use an alternative that has variable capacity. This is the argument for using public warehousing to supplement a private warehouse during a peak sales period. It is less expensive to use a small private warehouse at full capacity all year, and hire public warehousing as needed. Rather than incurring the cost of a large private warehouse that is only used at capacity for a very short time, total cost of a mixed system is less.

Bulk handling accounts for 99 percent of Canadian grain exports. The products moved in containers are usually bagged (e.g., lentils) and destined for consumption without further processing. In 1997, the port delays at Vancouver are expected to add $50 million to the cost of the bulk system. The use of the container alternative would have reduced the impact of these delays and lowered the total system cost.

Variety exacts its price.

The greater the variety of products in a logistical system, the higher the inventories necessary to maintain customer service. Product variety increases pipeline inventories. This raises the logistical costs for storage and financing inventories. Greater variety also lowers the average shipment size. Assuming the same volume of demand, dividing the product into more classifications produces smaller shipments.

Variety is the soft underbelly of the bulk handling system for grain (Prentice and Craven, 1980). The more products that the bulk system has to maintain separately, the less efficient that it becomes. The demand for "Identify Preserved Grains" (IPG) is increasing: organic wheat, variety preferences, specific quality attributes, e.g., protein. Advances in genetics and the

demands for product differentiation threaten to congest the bulk system with further variety. The bulk system could operate more efficiently if the lower volume, small shipments are moved in containers.

One size does not fit all.

A differentiated distribution strategy may be applied to products, sales volume or customers. Some consumers are willing to pay more for premium service, while others are just interested in the lowest price. A standard level of service dissatisfies the "picky" customers, and has attributes for which the "cheap" customer will not pay. Similarly, some products have high margins and are very popular, others may be offered more as a service than a profit centre. Providing the same level of logistics service in both cases is not warranted.

Bulk handling favours a push, rather than a pull system of marketing. Large stockpiles are created and deliveries are made in boatload quantities. If the customer wants lowest cost, average quality and will accept these volumes, the bulk handling system is ideal. Not every customer needs or wants the volumes that are most "economic" for bulk delivery. Some buyers process volumes that could be handled on a Just-in-time (JIT) basis using containers.

Delay commitment to the final product until the last possible moment.

This logistical strategy is used to lower finished inventory costs and increase customer service. The classic example is the distribution of paint. Only untinted paint is shipped to retail outlets. Tints are added after the customer chooses the desired colour. This reduces inventories held at retail and eliminates obsolescence while improving the selection of colours available. The strategy of delaying commitment is now being used in a range of products. For example, assembly plants have been established in the Netherlands to finish the manufacture of Japanese computers and other electronic equipment for the European market.

Bulk handling reduces the foreign miller's opportunity to tailor processed grains to the exact specifications of the buyer. Commitment is made to the quality of the final product, as soon as the grain is commingled at the country elevator. Containerized grain would delay commitment and give the foreign processor a multitude of options. Each container would be bar-

coded with information on grain variety and exact quality attributes. The foreign processor could blend grain with different quality attributes from a small inventory of bar-coded containers to fit a wide spectrum of finished product specifications.

It's Total Costs that matter.

There is a tendency in a logistical system to view the next agent in the marketing channel as the "customer". Great efforts may be made to reduce costs and improve the service to this customer, without considering the impact on the entire marketing chain. Only system changes that lower total costs to the end consumer make everyone better off. The principal of the total cost concept is that producers need to be as concerned about their customer's customer as they are with their customer. Some firms have found that getting their immediate customer to bear more costs is possible, if they can lower the total costs to the ultimate customer.

The Canadian grain handling system is designed as if all the important transportation and handling costs ended at the port of export. Most Canadian grain exporters have no idea of the total costs to the processor in the foreign country. Reducing the costs of the Canadian portion of the system shifts costs to the buyers. Large bulk shipments may be the lowest cost method of moving grain from farm to port, but imposes high inventory holding and storage costs on importers. A JIT container system could have higher transportation costs in Canada, but would virtually eliminate the storage, inventory holding and shrinkage costs for the foreign processors.

Intangible costs, such as the reliability, are also important. The greater uncertainty of the bulk handling system adds to costs. Foreign buyers must have contingency plans and additional inventories to guarantee service to their customers. A rough comparison of the shipping time for the bulk handling system and the proposed container system is presented in Figure 1. The comparison is only approximate because no informed opinion could be obtained in Canada of the unloading time and storage of grain in foreign import terminals. These data show that the bulk handling pipeline is four times longer than a container system. A 25 percent variation in delivery time of the bulk system would equal the time required for a container movement.

The costs of the relatively slow bulk system are reminiscent of the sacks grain it replaced. Assuming FOB sales contracts, the Canadian producers are financing inventories for 10 times longer in the bulk system. These costs are generally absorbed in prices, or final payments to producers for the Canadian Wheat Board.

Figure 1 Approximate Shipping Time Comparison for Bulk Handling and Containerization of Grain

Bulk Handling System	Days	Container System	Days
Farm Storage		Farm Storage	
Local Delivery	1	Local Delivery	1
Primary Elevator	40	Intermodal Terminal	2
Rail Hopper Cars	11	Double-stack Train	2
Export Terminal	19	Intermodal Port	2
Bulk Shipment	15	Container Ship	11
Import Terminal	10	Intermodal Port	2
Local Delivery	1	Local Delivery	1
Final Customer		Final Customer	
Total	97		21

Cost Trade-offs: Bulk versus Containers

Cost trade-offs are central to the re-engineering of logistics. The scope and design of logistics involve a balance between conflicting activities. Logistical functions can never be eliminated, only the cost of one activity can be traded-off against the cost of another activity. For example, transportation costs can be reduced by opening additional warehouse distribution points, but to maintain the same level of customer service, the inventory holding costs must increase.

Cost of physical handling versus container transfer. Grain is loaded and unloaded at least four times before reaching the final customer. Containerized grain would be handled once and the container transferred three times. The economics of size in bulk handling may not be totally exhausted, but further gains are likely to be marginal. Container systems are experiencing rapid productivity improvements. Computer controlled cranes and robotic trolleys are reducing labour costs and increasing handling speed.

Pipeline storage costs versus use of containers. Farmers already possess most of the storage they require to protect the crop after harvest. The storage provided in the bulk handling pipeline duplicates on-farm storage. The container serves as transportation unit and storage. Farmers could use containers to supplement on-farm grain storage. When the grain was sold, the storage would move into the handling system.

Economics of shipping versus inventory holding costs. The speed of the container system, and the opportunity to use Just-in-time scheduling reduces the costs of financing pipeline inventories. Lower inventories also reduce the risk associated with spoilage and shrinkage due to pests. On the other hand, approximately two tonnes of metal have to be physically moved with each 20 tonne container shipment. Adding 10 percent to the weight of the shipment increases the cost of transport.

Empty backhauls versus tare weight of containers. Containers are the new "boxcars" of the railways. Whereas bulk systems experience empty backhauls, the availability of empty containers would attract a variety of freight. If no backhaul can be obtained however, the container is likely more expensive to return empty than the bulk handling equipment.

Low freight costs/average quality versus freight premiums/exact quality. The bulk system offers low-cost service but delivers only average quality grain. Containers cost more to transport, but can deliver exact specifications. Some buyers may find the quality benefits more than offset the extra transportation charges. Its worthwhile bearing in mind that the value of grain contained in a loaf of bread, or a bottle of beer, is measured in pennies. A penny more for transportation might be easily extracted as a quality premium for the final product.

Factors Favouring Containerization of Grain

A range of technological and economic factors favour the containerization of grain.

Freight costs: Container rates for grain to Asia Pacific are about 30 to 50 percent higher than the direct costs of bulk handling. Several improvements are likely to narrow this difference. The container fleet has nearly doubled in size since 1985 (Worldwide Container Growth, 1997). In addition, the capacity of the containerships has nearly doubled. In 1980, containerships carried 2,000 to 3,000 TEUs (twenty foot equivalent unit). The new containerships are 6,000 TEUs, while 8,000 TEU ships are on the drawing boards. Freight rates for grain will be bid down as steamship lines compete for cargo to fill these new ships. A 6,000 TEU ship will load the equivalent of 15 double stack container trains. Container terminals are becoming highly automated to maintain a 24-hour turn around in port.

The railways are improving their container service. The first double-stacked container train service was introduced by APL in 1984. The Association of American Railways estimates that double stacking lowers line-haul costs by up to 40 percent (Martin, 1996). Container volumes in North America have enjoyed a compound growth rate of 5.7 percent since 1988. Better service and lower costs are anticipated as the railways upgrade their container terminals and add equipment.

Further cost improvements are likely as the container system becomes more mature. Steamship lines jealously guard their container fleets. The railways have invested in "domestic" containers that are larger, and do not move overseas. The pooling of container fleets is beginning to gather support. Increasingly, the industry is pointing to "Gray Boxes," which are owned by third parties, as a method of increasing vehicle utilization.

Communication Costs: The revolution in communications can be summed up with one word: Internet. The information highway may be the most important innovation of this age, but it is still in a primitive state. Proposals to launch a network of satellites that would create real time access to the Internet are being considered seriously. Visions of global commerce via the Internet are already taking shape. The impact of the Internet on grain marketing in the 21st Century could be akin to the changes that occurred in the 19th Century when the telegraph was introduced. Just

as the telegraph decentralized the physical transaction of grain, the Internet could decentralize the electronic transaction of grain. Information is replacing the need to store large inventories and reducing the economies of size in order processing.

Price/quality Considerations: Bar codes and computer data bases reduce the effort of tracking container shipments. Demand for "designer grains," or IPG, will lead the move to larger container volumes. Quality premiums to farmers are encouraging the production of organically grown crops, specific varieties of mainstream crops, and a variety of "special crops" that range from herbs to pulses. These products cannot move economically through the bulk handling system. As their volume grows more products will be added to the list of containerized grain.

JIT service to foreign grain processors is an advantage of containers that bulk shipment can not provide. Given the success of JIT in manufacturing, it is only a matter of time before processors begin to demand JIT grain shipments. If the benefits from delayed commitment can be added to the reduction of inventory costs, the volume of container movements will accelerate.

Electronic Markets for Prairie Grain

The impact of communications technology and containerization could be far reaching for the grain industry. The potential exists for farmers to develop electronic markets that replace the middleman function performed by grain handlers. Operationally, farmers could post container loads of products on an Internet site, and field purchase offers via e-mail. Potential buyers could be sent grain samples by FedEx. Transactions could be consummated with container shipments that never touch the current bulk handling system, or use the services of grain handlers. Naturally, some institutional arrangements would have to be put in place to assure the farmer received payment, and the buyer received the actual quality according to the contract. These institutional arrangements constitute the requirements of an electronic market.

Electronic markets enable traders to buy and sell complicated products without being in physical contact. There are certain institutional prerequisites that must be developed for an electronic market (Prentice and Mulligan, 1996). A computer-supported system must be used to search and

negotiate the transaction. There must be membership rules, a method of quality checking and a system for settling transactions. Some elements of an electronic market for containerized grain are in place; the missing pieces present no significant barrier.

The Canadian Wheat Board (CWB), which has a monopoly on the export of wheat and barley, could administer an electronic market for these crops. Producers could test their harvest to determine high quality (e.g., protein) wheats that they wish to market in containers. The wheat would be stored in containers and bar codes would identify the container and samples sent to the CWB. The CWB could confirm the quality and list the containers of wheat on an electronic data base. As sales of specific qualities of wheat were made, the CWB could pull the containers forward to fill customer orders. The CWB would guarantee the quality and ensure that the producers were paid.

The need for containers could create a second electronic market for farmer-owned "Gray boxes". Essentially, producers would purchase containers and form a leasing pool. While the containers were at the farm, the producer would receive no payment. When the grain was sold, the container would enter a leasing pool that would return a payment to the farmers. In the next harvest season, the producer could request the delivery of a container, and the cycle would begin again. Containers last about seven years. Farmers would not necessarily receive the container they purchased, but would receive "rights" to a container for seven years. Container leases could be traded electronically, with bids and offers entered by producers, grain handlers, carriers and buyers. This would encourage utilization of equipment and maximize producer returns.

Conclusions

After 150 years of growth and development, most grain exports are marketed through grain elevators, rail hopper cars and bulk carriers. The bulk handling system is so ubiquitous and entrenched as to be uncontested. No doubt it appeared the same to the handlers of bagged grain when the bulk method was first introduced. There have been many refinements and productivity improvements in the bulk handling system, but the basic concept remains unchanged. The economic rational for the creation of bulk handling of grain have changed. Technological advances in information

collection and transmission are now make container shipment a rival for the bulk handling of grain.

The motto of the Industrial Age (1850-1975) can be summed up with the expression: *"If it ain't broke, don't fix it!"*. In the modern era, which some have termed the Information Age, this expression could be changed to *"If it still works, it's probably obsolete!"* Rapid changes because of new information technology are forcing many systems to be overhauled. These changes spare no sector of the economy, least of all the Canadian grain handling industry.

References

Ballou, Ronald H. *Business Logistics Management.* Third Edition, Englewood Cliffs, NJ: Prentice-Hall, Inc., 1992.

Cronon, William. *Nature's Metropolis.* New York, NY: W.W. Norton & Company, 1991.

Kindleberger, Charles P. *Foreign Trade and the National Economy.* New Haven, CT: Yale University Press, 1962.

Martin, James D. "Special Report: Intermodal" *Transportation and Distribution.* February 1996: A-X.

Prentice, Barry E. and Jack W. Craven. "The Economic Potential for Improving the Canadian Grain Handling System through Containerization", *Canadian Journal of Agricultural Economics.* Proceedings Issue, August 1980, 119-128.

Prentice, Barry E., and John O. Mulligan. *An Electronic Market for Truckload Freight: Economic Definition, Existing Structure and Missing Prices.* University of Manitoba, Transport Institute, Research Bulletin #15, Oct. 1996.

Wilson, Charles F. *A Century of Canadian Grain.* Saskatoon, SK: Western Producer Prairie Books, 1978

"Worldwide Container Growth" (source: Clarkson Containership Register), *Inbound Logistics.* Vol. 17 No. 1, January 1997.

1. The Canadian Pacific Railway offered free lease sites on sidings to companies willing to build "standard" elevators of not less than 25,000 bushels. "This offer was accompanied by an undertaking by the railway not to allow competitive loading from flat warehouses or from farmers' wagons at points where elevators were located." (Wilson, 1978).

2. Re-engineering a logistical system is trade creating. Kindleberger (1962) has described the effect of innovations in transportation and communications as having a double impetus for trade. Innovations that lower the cost of logistics increase profits for the producer while it lowers prices for the consumer. The incentive to buyers combined with the incentive to sellers creates the double impetus to expand trade.

NON-STOP APPEAL: VIABILITY OF LONG-RANGE LOW-DENSITY AIR ROUTES

Abstract

This paper explores consumers' preference for direct, intercontinental service and the technical feasibility of developing the niche of international long-thin routes. By modifying a route choice and market share model derived from US domestic data, insights can be gained as to how the preference for price, service frequency and convenience of travel interact to influence the market appeal of a direct intercontinental service. Meanwhile, improvements in technical capabilities and operational economics of the latest generation of long-range jets render this niche an attractive strategic option in an increasingly globalized airline industry. The success of Air New Zealand in so doing is evidence of the viability and strategic importance of this niche.

Acknowledgements

The author gratefully acknowledges the advice and support provided by Dr. Tae H. Oum, W.J. van Dusen Professor of Management at the University of British Columbia, and information provided by Airbus, Air New Zealand, Boeing Commercial Airplanes and *Frequent Flyer*.

Overview

Much research has been conducted on the supply and demand of air services in the United States, especially on the hub-and-spoke route networks (Oum and Tretheway, 1990) developed since the deregulation of the industry. The primary appeal of hub-and-spoke systems is their efficiency in geographical coverage through traffic consolidation. The consolidation of traffic allows higher frequency of service between the hub and non-hub cities, which in turn can stimulate demand primarily by reducing the amount of schedule delay. The inconvenience caused, however, is usually considered to be more than compensated for by the increase in flight frequency in short-haul domestic travels. However, as the travel distance increases, the amount of trip circuitry also tends to increase. At some point, there exists a trade-off between the preference

toward flight frequency and the preference toward trip convenience.

Meanwhile, improved technical capabilities and operational economics of the latest generation of long-range jets allow airlines to offer more long-haul, intercontinental flights with lower break-even demand. This favours the development of direct, point-to-point intercontinental services rather than routing through hubs and represents a viable strategy for smaller carriers to enhance their international competitiveness in niche markets despite the monolithic strengths of global mega-carriers.

In this paper, estimates from a route choice and market share model will first be examined. Then, with reference to the latest generation of long-range jets, the technical viability of direct, non-stop service will be explored. Finally the success by Air New Zealand in developing the niche of long-thin intercontinental routes will be used as a case study.

Market Appeal of Direct Intercontinental Services

Ideally, industry data can be used to quantitatively estimate consumer's preference on travel choice selection. However, by the very nature of long, thin intercontinental routes, difficulties arise in collecting samples of statistically significant sizes on those who have opted to travel on direct flights versus others on connecting services. This is exacerbated by the lack of published international origin-destination data.[1]

Instead of resorting to primary data, insights on the market appeal of direct intercontinental services can be gained through analyzing key research findings on travel preference in the US domestic market (where the data is most complete), and then by extrapolating consumers' preferences on longer-haul travel. Industry surveys are a good starting place for this discussion.

The result of two surveys on important factors affecting an airline passenger's travel decision is shown in Table 1. Airline's reputation was shown to be a significant factor. Schwieterman (1995) estimated the effect of an airline's reputation on ticket price, but the study was limited to US carriers only, therefore precluding further inclusion of this variable in a route choice decision model. The affiliation with a frequent flyer program is also an important factor in a travel decision, but there is simply not sufficient comprehensive, disaggregate data for analysis here.

Table 1. Factors Affecting Travel Decisions

Factor in Order	Frequent Travellers*	OAG Users^
1st (most important)	Schedule	Departure/Arrival Times
2nd	Frequent-flyer program	Air Fare
3rd	Price	Frequent-flyer incentives
4th	Non-stop flight	Route

Sources:

* Annual Reader Poll, *Frequent Flyer*, December, 1995

^ *Pocket Flight Guide Usership Survey – The Harris Poll*, Official Airline Guides, 1995

As evident from the survey, the notion of schedule convenience has two components: frequency of flight (schedule) and ease of travel (the convenience of an itinerary). The former is obviously an important decision. From a firm's perspective, Kahn (1993) showed in detail how United Airlines' dominance in frequency (over American Airlines') from Chicago O'Hare contributed to its disproportionate number of passengers carried and revenues earned. As well, Morrison and Winston (1986) reported that a doubling of frequency of air service would lead to a 21% increase in the demand for air service by business travellers.

The second aspect of schedule convenience, the ease of travel, is also an important attribute in an individual's travel decision. This refers to the number of en-route stops, or whether or not a connection is necessary at a hub airport. This will be explored in more details later.

Certainly, price is an important factor, although its relative importance varies from frequent flyers to leisure travellers. Oum, Gillen and Noble (1986) found that in the United States, the price elasticity for business travellers was -1.15 and that for leisure travellers was -1.5. Agarwal and Talley (1985) reported an elasticity of -0.8425 for international air travel to/from the United States. De Vany (1974) explicitly derived price elasticity values for US domestic travel as an increasing function (absolute value) of trip distance. At the same time, a similar relationship for time elasticity (as a function of trip distance) was derived.

Insights from a Market Share Model

Significant research has been conducted in the travel behaviour in the US. A mathematical model derived from these findings will be adapted to approximate the appeal of direct intercontinental services. Certainly, the results of this technique have to be viewed with caution, since there is an implied assumption that the preferences for some intrinsic trip qualities for both long-haul and short-to-medium-haul travels are similar and do not vary significantly with time. But in the absence of primary data, the insights gained can act as reasonably useful guideposts.

Kanafani, Ghobrial and Hansen have modelled the decision-making process in flight route selection in the US. All of them have used a logit model for the route choice probability (P, manifested in market shares) of traffic on a given flight route (i) between a given city pair (r):

$$P_r = \frac{e^{V_r}}{\sum_i e^{V_r}}$$

where V is a deterministic utility of a given itinerary. Since it is the difference in the deterministic utilities of competing flights that define their respective market share, the absolute value of utilities is not important.

Kanafani and Ghobrial (1985) modelled the utilities based on short-haul domestic travel. Their attributes included travel time, daily frequency, fares, aircraft size and connectivity. Summarized in the connectivity term were the differences among non-stop service, on-line service with stops and indirect connecting service. However, the deterministic effects of these attributes were all simply additive, thereby limiting the model's use to only short-haul domestic traffic.

Ghobrial (1989) examined the competition between US and foreign air carriers on 27 international routes. The use of the logarithm of weekly frequency and the definition of convenience of service as the ratio of non-stop flights to total flights in a specific market were instrumental in modelling travel behaviour, but the omission of price effect precludes its full adoption for intercontinental modeling.

Hansen (1990) estimated the deterministic utility from passenger itinerary and airline flight schedule data (in 1985), based on the US

Department of Transportation 10% flight coupon data. He estimated the functional form of the deterministic utility of a direct service[2] as:

$$Vd = aF + q_o \ln(f) + Vdct$$

where F is the fare in dollars, f is the daily frequency, Vdct is a constant reflecting consumer preference for non-stop service, and a and q_o are also constants as shown in Table 2. Similarly, the utility for on-line connecting service was estimated as:

$$Vc = aF + s\,CIRC + q_1 \ln(fmax) + q_2 \ln(fmin)$$

where CIRC is the trip circuitry in miles, fmax and fmin are the maximum and minimum frequencies of the connecting "spoke segments", and s, q_1, and q_2, are constants as shown in Table 2.

Table 2 Market Share Model Estimates from Hansen's Study (1990)

Coefficients of Attributes	Estimated Value	Standard Error
ln (direct freq), q_o	1.29	0.17
ln (max direct freq), q_1	0.33	0.14
ln (min direct freq), q_2	0.78	0.10
Direct service utility, Vdct	2.72	0.81
Fare in US $, a	-0.0045	0.0010
Circuitry in miles, s	-0.0029	0.00026

271 observations were made, with an R-squared value of 0.74

In intercontinental travels, the effect of higher trip circuitry can be approximated by the CIRC term. The use of the natural logarithm of frequency and the differentiation of direct versus connecting frequencies are generally applicable to long-haul flights. The effect of price has also been incorporated. These make Hansen's model generally applicable to approximate intercontinental travel behaviour.[3]

To do so, these estimates can be applied to a sample routing, say, from Vancouver to Osaka, assuming a symmetry of demand and the offering of only one fare (full economy) per airline. At this writing, according to the *Official Airline Guide*, three airlines: Air Canada, Japan Airlines and United Airlines (AC, JL and UA respectively), offer convenient on-line service on this route (without much backtracking),

and are likely to attract the majority of traffic in this market. Their itineraries are shown in Table 3, and for the purpose of illustration, only these three will be considered.

Table 3. Sample Itineraries from Vancouver to Osaka

Carrier	Stop	Flight time, hr	Transfer time, hr	Daily freq.	Circuitry miles
AC	No	10:50	0	5/7	0
JL	NRT[4]	12:10	1:45	1	671.7
UA	SFO	13:25	3:30	1	1307.4

Airports: NRT-Tokyo Narita, SFO-San Francisco International

If all three airlines offer the same fare of US$1,200 (with other attributes being equal), and if the travel choice selection pattern on the intercontinental market of Vancouver-Osaka is the same as would be expected in US domestic flights, then according to Hansen's model, the market share that would be gained by these three airlines can be expressed as a function of the frequency of non-stop service offered by Air Canada. However, using his original model, even as both UA and JL drop their fares to 75% of that of AC's, the non-stop flight would still be very popular, capturing more than 60% market share with the offering of only 1 weekly flight. This certainly appears to be too favourable for direct intercontinental service, and more stringent conditions need to be applied.

As alluded to previously, De Vany (1974) estimated that as the trip distance increases, the absolute values of price elasticities increase while those of time elasticities decrease. Halving the circuitry coefficient would likely give a more stringent estimate in Hansen's model, with the lower bound of it being zero. A simultaneous doubling of the fare coefficient would reflect a significantly higher sensitivity to fare. The exact magnitude of these adjustments required to model intercontinental air transport demand is not known. However, based on the standard errors associated with these attributes, as reported by Hansen, these two modifications do represent rather stringent conditions to model intercontinental traffic.

Figure 1 shows the effect of lowering fares by UA and JL if AC continues to charge $1,200 and flies 3 non-stop flights a week, based on

these modified conditions. As evident from the figure, even if Air Canada operates only 3 non-stop weekly flights, it would still attract a higher proportion of origin-destination traffic than either JL or UA, given a fare difference of $200. This is consistent with Hansen's original prediction in the domestic setting that even when no extra distance is involved, 15 flights per day on each of two flight segments would be required for a hubbed service to gain a 50% market share if the only other competitor is a direct service of one flight per day.

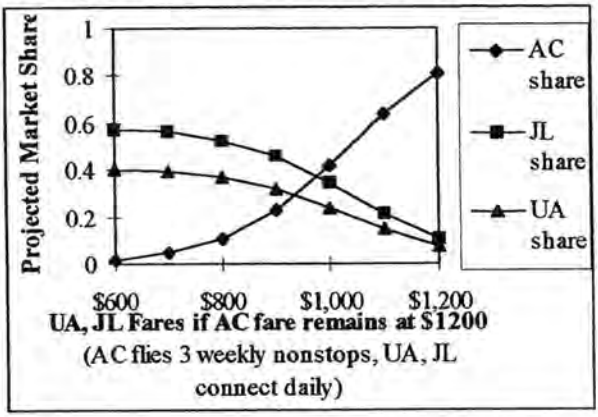


Figure 1. Effect of Price on Market Share, Modified Hansen's Method

The effect of connecting time has not been segregated from the direct service utility in Hansen's estimate. Alternatively, Kanafani and Ghobrial (1985) estimated that the amount of travel time has a linear effect on deterministic utility, with an associated coefficient of -0.897, which, when converted to trip circuitry at 800 km/h, presents a more favourable condition to non-stop service than Hansen's estimate. The inclusion of additional time delay, with Hansen's estimate, does seem to have a significant impact on the predicted market shares (see Figure 2).

Figure 2 shows the effect of non-stop service frequency on the market shares of the Vancouver-Osaka market, assuming that AC charges $1,200 and both UA and JL $200 less, based on modified Hansen's estimates. Even under various stringent conditions as discussed above, the market appeal of a direct intercontinental flight[5] is certainly respectable compared to that of connecting services.

Given the attractiveness of direct intercontinental service, what then,

if any, has prevented airlines from fully exploiting this niche? The availability of aircraft technology has been a major factor, and will be discussed next.

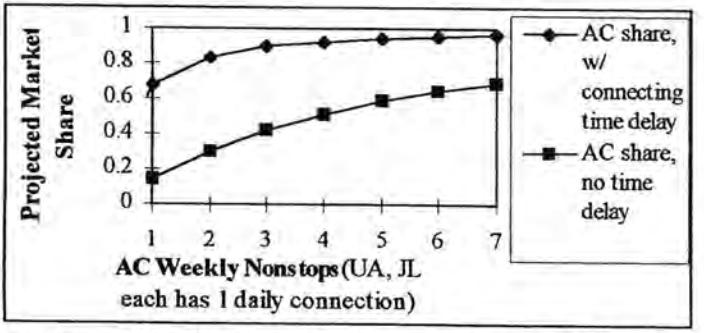


Figure 2. Effect of Non-stop Frequency on Projected Market Shares

Latest Long-Thin Intercontinental Jets

Ideally, the costs of providing these non-stop services can be compared with those of connecting service, thus arriving analytically at a precise measure of benefit in developing one type of service over the other. In reality, however, the effect of hub transfers and scale of economies accomplished through consolidated maintenance and operation bases, the distribution of stochastic delays, and the difference in management structures can add significant uncertainty to such analysis.

Meanwhile, from a global perspective, if the "right" aircraft technology is available, and if an airline can demonstrate success in using these aircraft types to develop the intercontinental long-thin niche to gain international competitiveness, then it is reasonable to conclude that such markets are both viable and of strategic importance.

Technological capabilities had, until the past decade, been a major impediment in the full development of thin, direct intercontinental markets. The 747 had the most impressive payload-range capability, but its sheer size also translated to a larger number of break-even load factor. The smaller 747SP was relatively costly to operate. Some versions of the DC-10 and L1011, notably the DC-10-30 and L1011-500, have been designed to carry fewer passengers over shorter distances than the 747, and were moderately instrumental in developing intercontinental

linkages, albeit often with intermediate stops.

The latest generation of "long-thin" intercontinental jets has certainly changed the scene. The range and typical mixed-class seating capabilities of these and other older aircraft types are shown in Figure 3.

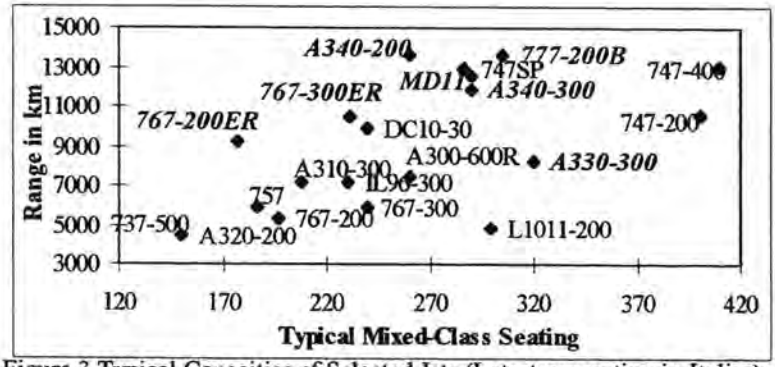


Figure 3.Typical Capacities of Selected Jets (Latest generation in Italics)
Sources: Airbus Industrie, Boeing Commercial Airplanes, Green & Swanborough (1988), *Flight International* (1993) and McDonnell Douglas Aircraft.

The latest generation of intercontinental long-thin jets includes the Boeing 767's and 777's, Airbus A330's and A340's, as well as McDonnell Douglas MD-11. In some occasions, the 757-200ER, A310-300 and A300-600R are also used to perform intercontinental flights, particularly across the Atlantic and between Europe and the Middle East. As evident from Figure 3, these aircraft types provide impressive range capability but relatively low seating capacity. Figure 4 reveals that these aircraft types have significantly lower long-range fuel consumption than their predecessors.[6]

Apart from fuel economy alone, the latest generation of long-range, low-density jets generally has relatively low operating costs.[7] Despite the economics of these latest intercontinental long-thin jets, since many of these have just entered scheduled service in the past 2 years or so, the benefits from their deployment and the market niches developed may have yet to mature. Nevertheless, the 767's have been used in the transoceanic flights since the late 1980's, and the success in their intensive use in long-thin routes will be particularly insightful.

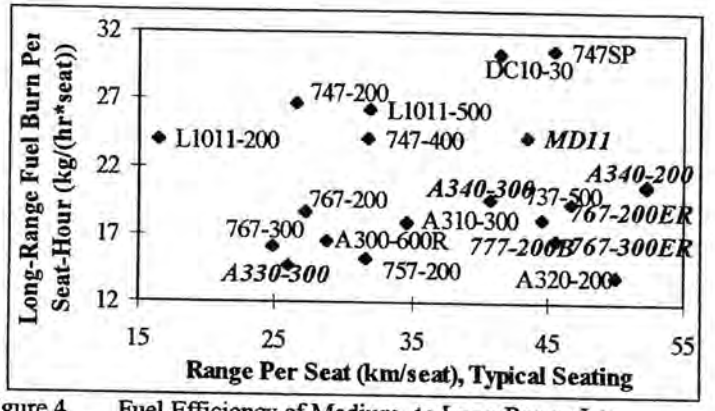


Figure 4. Fuel Efficiency of Medium- to Long-Range Jets
Sources: Green and Swanborough (1988), *Flight International* (1993),

Air New Zealand's Case

Since the late 1980's, the use of 767's over transoceanic services became more popular. However, for the purpose of this paper, the use of such aircraft in developing the intercontinental long-thin niche has to be shown in attracting and possibly stimulating origin-destination traffic between the cities served, rather than merely collecting traffic from a particular region. In this case, Air New Zealand's success serves as an excellent example.

The geographical location of New Zealand offers little opportunity for an airline to develop a comprehensive network for connecting traffic. In 1986, Air New Zealand only flew to 2 destinations in Asia, using the 747-200: Tokyo and Singapore (twice weekly to each from Auckland, airline timetable, 1986). Reciprocal services from the region originated from these two countries only. Across the Pacific, it connected Fiji and Tahiti directly with Los Angeles, again with the 747-200.

In contrast, Qantas of Australia, backed by the enormous size of its home market, flew daily to Hong Kong and Singapore, near-daily to Bangkok and Tokyo, and regular service to Kuala Lumpur, Beijing, Manila and Indonesia. Further, passengers travelling from these places to New Zealand, can often find convenient on-line connections via Qantas. For Air New Zealand to continue operating successfully as an independent, international carrier, it had to find a way to bypass Qantas'

hubs and fly direct to other growing regions of the world.

Since 1987, Air New Zealand has embarked on an expansion plan in the Pacific region, making use of both the underlying consumer preference for direct service and the economics of the 767's. Today,[8] the airline flies to 10 destinations in the region (36 flights a week in total), but the 747-200 is used only to fly to three high-density markets: Tokyo, Taipei and Hong Kong (8 weekly flights total). Meanwhile, the 767 has also been used to connect the South Pacific islands of Western Samoa, Tonga and the Cook Islands directly to the US. Figure 5 compares the growth in Air New Zealand's long-range fleet from the end of 1986 to 1995, providing proof to the airline's increased emphasis in the intercontinental long-thin niche.[9]

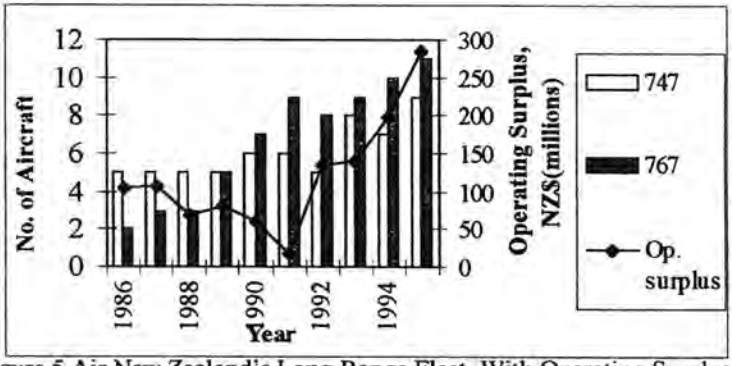


Figure 5.Air New Zealand's Long-Range Fleet, With Operating Surplus
Source: Air New Zealand Annual Report (1996)

Some operational statistics of Air New Zealand from 1986 to 1994 is shown alongside that the world's total in Table 4. It is evident from the table that during this period, the growth in revenue-passenger-kilometres (RPK) for the airline has almost doubled that of the world total, while its growth in the number of passengers carried has been in line with the world's average. The relatively low growth in freight-tonne-kilometres (FTK) carried compared with the world's figure is indicative of the slower-than-average growth of New Zealand's home market.[10] These, combined with the much faster growth of its 767 fleet compared with that of the 747, are clear evidence of the impact of the airline's newly developed intercontinental long-thin markets.

Table 4. Operational Statistics of Air New Zealand in 1986 and 1994

Category	Air New Zealand 1986*	1994*	% change
Pax No.	4.2 mil	5.8 mil	+38
RPKs	8.0 bil	15.6 bil	+94
FTK	329 mil	493 mil	+50

Category	World Statistics 1986*	1994*	% change
Pax No.	960 mil	1.3 bil	+37
RPKs	1,452 bil	2.3 tril	+56
FTK	43.2 bil	97.2 bil	+125

Sources: (*-includes domestic operations) Air New Zealand (1996); *Air Transport World* (1986, 1994)

Strategic Importance and Implications

The successful development of the niche of intercontinental long-thin service through the use of advanced aircraft technology by Air New Zealand can be considered as a practical confirmation of the market share modelling and the operational economics of the latest generation of intercontinental aircraft. Certainly, sound management strategies are important in the success of implementing such strategies. Nevertheless, there is clear evidence of the economic viability, technical feasibility and strategic importance of the direct, intercontinental long-thin markets. This, in particular, has some serious implications for airline strategy planning elsewhere in the world.

Similar to New Zealand, the geographical location of Canada limits the opportunities for Canadian carriers to build effective global hub-and-spoke networks. As well, carriers in the neighbouring United States have comparative advantages on traffic feed owing to their much larger home market and coverage. Passengers to and from Canada, for instance, often observe that a higher frequency of service to an international destination can be provided by an American carrier, albeit with the penalty of added trip circuitry, travel time and inconvenience. Reflecting on Air New Zealand's experience, the recent introductions of non-stop services like Vancouver-Beijing by Canadian Airlines International and Toronto-Tel Aviv by Air Canada appear to be in the right direction.

On the transborder front, the recent near-"Open-Skies" air transport agreement between Canada and the US opened up new opportunities for carriers on both sides of the border. Together with the introduction of the Canadair Regional Jet (CL-65), the Airbus 319 and the latest generation of the Boeing 737's, the transcontinental version of the "long-thin" niche may be of strategic importance in increasing transborder market shares.

Conclusions

To an individual traveller, the higher frequency offered by a hub-and-spoke network between non-hub cities often demands additional time and inconvenience. On intercontinental flights, this inconvenience could be more significant, and beyond a certain point, the comfort of a non-stop flight may be more important to a passenger than frequency and price.

By modifying results from Hansen's study on US domestic travel behaviour, some insights into how the frequency of a non-stop intercontinental service can affect a carrier's market share in face of competition from hub-based connecting service. These findings, however, should be interpreted with caution, as a number of important factors, including frequent flyer programs and airline reputation, still have not been incorporated.

In any case, given the underlying demand for direct, non-stop service, the availability of a host of new aircraft types appear to have rendered the niche of intercontinental long-thin routes a viable strategic option for airlines around the world.

References

Abrahams, M. (1983): "A Service Quality Model of Air Travel Demand: An Empirical Study", *Transportation Research*, Vol.17A, No.5, p.385-393.

Agarwal, V. and W. Talley (1985): "The Demand for International Air Passenger Service Provided by U.S. Air Carriers", *International Journal of Transport Economics*, Vol.12 No.1, p.63-70.

Air New Zealand (1996): *Annual Report 1995*, p.30-31.

Air New Zealand: Various issues of international timetable.

Airbus Industrie (1987): "A330/A340", *Air Transport World*, Vol.24 No.2, p.28.

Airbus Industrie: Company aircraft fleet brochures.

Boeing Commercial Airplanes: Company brochures.
"Commercial Airliners of the World" (1993), *Flight International*, Vol.144 No.4393, 27 October-2 November, p.66-84.
De Vany, A. (1974): "The Revealed Value of Time in Air Travel", *The Review of Economics and Statistics*, Vol.56 No.1, February, p.77-82.
Force, B. (1990): "Cleaning the Air", *Business Traveller*, July, p.20-24.
Ghobrial, A. (1989): "Competition Between U.S. and Foreign Air Carriers: An Econometric Approach", *International Journal of Transport Economics*, Vol.16 No.1, February, p.19-33.
Green, W. and Swanborough, G. (1988), *"Airliners"*, p.5-130.
Hansen, M. (1990): "Airline Competition in a Hub-Dominated Environment: An Application of the Nonco-operative Game Theory", *Transport Research B*, Vol.24B No.1, p.27-43.
Kahn, A.E. (1993): "The Competitive Consequences of Hub Dominance: A Case Study", *Review of Industrial Organization*, Vol.8, p.381-405.
Kanafani, A. and A. Ghobrial (1985): "Airline Hubbing - Some Implications for Airport Economics", *Transportation Research*, Vol.11, p.1-8.
McDonnell Douglas/Boeing: Company brochures for the MD-11.
Morrison, S. and C. Winston (1986): *The Economic Effects of Airline Deregulation*, Brookings Institute, Washington D.C., p.15-21.
Official Airline Guide, Reed Travel Group, various issues.
Oum, T.H., D.W. Gillen and S.E. Noble (1986): "Demand for Fareclasses and Pricing in Airline Markets", *Logistics and Transportation Review*, Vol.22 No.3, p.195-222.
Oum, T.H. and M.W. Tretheway (1990): "Airline Hub and Spoke Systems", *Journal of Transportation Research Forum*, Vol.30 No.2, p.380-393.
Oum, T.H. and A. Taylor (1995): "Emerging Patterns in Intercontinental Air Linkages and Implications for International Route Allocation Authority", *Transportation Journal*, Vol.34 No.4, Summer, p.5-27.
Schwieterman, J. (1995): "Hedonic Price Assessment of Airline Service Quality in the U.S.", *Transport Reviews*, Vol.15 No.3, p.291-302.
TravelScan (1995): "Aircraft Operating Data", *Air Transport World*, Vol.32 No.3, March, p.98; Vol.32 No.4, April, p.87.
"World Airline Statistics-1986", *Air Transport World*, Vol.24 No.6, p.68.
"World Airline Statistics-1994", *Air Transport World*, Vol.32 No.6, p.56.

Endnotes

[1]The International Civil Aviation Organization (ICAO) collects data on

each flight segment as well as same-flight traffic, but not connecting traffic. The U.S. Department of Transportation collects 10% of industry ticket samples, but this is limited to passengers on U.S. airlines. Both the U.S. Immigration and Naturalization Services (INS) and Transport Canada have limited origin-destination data, but with no account on itineraries.

[2]In Hansen's model, direct, on-line, same-plane service is not differentiated from non-stop service. It is reasonable to assume that for an intercontinental flight, as long as the longer sector is sufficiently close to the total stage length, the difference between a true non-stop and a direct flight is negligible. Otherwise, the utility of the flight would be somewhere between a true non-stop flight and an on-line connection.

[3]In fact, the changes in the total route-specific market size as a result of carriers' competition may also be significant. In any case, a carrier's market share should be a reasonably good indicator of its competitiveness.

[4]Passengers travelling on Japan Airlines actually have a choice of airports in Osaka. But only the connection to Kansai Airport will be considered here.

[5]In fact, the model does not measure the appeal of direct flights in general, but only non-stop ones. However, if the two flight segments of an intercontinental flight are far from equal, the difference in utility between the two longer-haul markets can be assumed to be more or less equal.

[6]Also, fuel consumption per passenger-kilometre decreases as the trip distance increases. This decrease can be as much as 20% for a London-Istanbul flight compared with a London-Paris flight by a 767. See Force (1990)

[7]See TravelScan (1995).

[8]Air New Zealand also uses fifth freedom rights between Australia and other Asian countries, but to a limited extent.

[9] In a similar vein, with the 747-400, Air New Zealand launched Auckland-Los Angeles non-stop and same-plane one-stop service to London, forcing its direct competitors to either drop the route (Continental) or adopt the same itinerary (United, British Airways).

[10]This observation is also consistent with the fact that Air New Zealand's domestic revenue-passenger-kilometres figure fluctuated between 1.365 million and 1.637 million only from 1986 to 1995 (Annual Report 1995).

VIRTUAL TRANSPORTATION AS A SOLUTION TO URBAN TRANSPORTATION PROBLEMS: A CRITICAL ANALYSIS OF THE POTENTIAL OF TELECOMMUTING AS A WORK TRIP REDUCTION STRATEGY

Introduction

Transportation in urban space has become increasingly difficult. The automobile, the dominant transportation mode choice, is restricting the movement of people during peak demand periods. The rising levels of traffic congestion caused by overloaded road transportation infrastructure, is evidence of this crisis. The mass utilization of the automobile has had a lasting effect on society, the economy and the environment[1]. Policy makers have planned and implemented numerous strategies to attempt to solve the problem of congestion and environmental degradation. The least supported solution, virtual transportation shows the most promise in eliminating or reducing urban transportation problems. The development of a global information economy and recent advancements in technology make virtual transportation a viable solution to the peak demand problem.

Virtual transportation has the potential to reduce traffic congestion and the associated environmental impacts by reducing the generation of work trips or by altering their spatial and temporal components. A clear understanding of the root of the transportation problem will explain why many of the contemporary proposed solutions are ineffective. The success of the transfer from physical to virtual transportation largely relies on how well implementation strategies are designed. Virtual transportation will have massive repercussions on society. Public policy issues and societal impacts will need to be mediated. Virtual transportation, although in its infancy, has shown results that support theoretical expectations.

The Urban Transportation Problem

Mode Choice

The philosophy of travelling to work was born in the industrial revolution when masses of people left the home for the factory. The

spheres of work and home were separated. Social practices prescribe the precise timing and locations of events that confine human movement to set rhythms between activity sites[2]. The invention and introduction of the Trolley car increased the spatial separation between home and work by allowing families to settle further from their place of employment. Trolley lines shaped the development of the city.

The automobile was a catalyst for the rise of a suburban culture in North America. Cities decentralized after World War II as widespread automobile ownership brought great increases in personal mobility, further increasing the distance between home and work. Land use policies and the baby boom fuelled the massive suburbanization that took place. Land use planning departments sought to safeguard property values by separating land uses. Suburban housing patterns resulted in large distances between places, making automobile mobility a necessity. Public investment in automobile infrastructure, namely expressways (1945-1970), opened up new land for development, decreased commute times to work and stimulated the economy. Refitting the urban environment to accommodate the car required a vast capital expenditure which translated into jobs and business activity[3]. Human settlements became auto-centric with urban design, urban infrastructure and policy making all catering to the car[4]. Auto-centric land use patterns make public transit logistically and economically infeasable.

The growth in automobiles has surpassed the growth of the population. Today, 70% of freeways have become congested to some degree as freeway capacity is surpassed, causing the length of time to commute to work to drastically increase[5]. The ever expanding urban periphery has increased the average number of vehicle miles travelled (VMT). The automobile meets up to 90% of all travel requirements[6]. Freeways were designed to move people between the periphery and the central business district or manufacturing regions. The rise of the information economy and municipal policy decisions have deconcentrated jobs across the urban space. The radial orientation of freeway networks is out of synchronization with the demand of intra-urban flows. In summary, the Post-Industrial metropolis is faced with inherent structural transportation problems that can only be solved by reducing the amount of travel demand serviced by the automobile.

Travel Demand

Transportation is a by-product of trying to accomplish some non-transport activity. Work trip travel demand is a derived demand. Home is

at one end of 80% of all trips made each day in the city. Work trips account for up to 50% of all intraurban trips[7]. Traffic congestion results from the peaking of travel demands during rush hours. Rush hour periods are between 6:00am to 9:00am and 3:00pm to 6:00pm. Standardized work hours are the culprit for this double peaking phenomenon. The absolute magnitude of the peak travel demand is increasing rapidly and is a function of three variables: activity accessibility, automobile ownership and labour-force participation.

Auto-centric urban design and spatially constricted automobile travel create concentrated travel flows. Transportation planners did not accurately forecast the future demand that would be placed on the freeway networks. People are willing to travel long distances to access economic activities. The utility of affordable housing surpasses the disutility of long commute times and distances. The length of trips measured in distance, time and cost, has been found to increase as city size increases and density decreases. Hence, the length of peak demand periods increase.

The growth in automobile ownership (population of cars) has outpaced the growth in human population in cities. Motor vehicle registrations in Canada increased by over three million from 1980 to 1992 or an average annual increment of 2.6 percent[8]. Expressways became congested as design capacity was surpassed.

Transportation planners designing the urban freeway network did not forecast the impact on demand caused by the growth in female labour-force participation. This growth was fed by the woman's movement, the development of a consumer society and the changing composition of the household. Peak travel demands soared.

Environmental Impacts

The automobile is responsible for a multitude of negative environmental impacts. The automobile pollutes the environment, murders and injures million of people and causes damage to property, consumes vast reserves of energy and occupies inordinate amounts of space. Advances in technology which mediate the impacts of the automobile are negated by the increases in its population.

The internal combustion engine consumes energy stored in petroleum and emits pollutants. Fossil fuel reserves are being depleted. Internal combustion engines that are fuelled by diesel fuel release particulates into the atmosphere. These particulates are a known carcinogen. The main pollutants emitted from automobiles propelled by the internal combustion engine are: Carbon Monoxide, Carbon Dioxide,

Volatile organic Compounds and Nitrous Oxides. Congestion and increases in vehicle miles travelled only magnifies the magnitude of the emissions outputted by an automobile during a work trip.

Automobile infrastructure consumes vast quantities of space and phenomenal quantities of natural resources. Roadways partition or destroy neighbourhoods, farmland and wildlife habitat. Automobiles generate large amounts of noise pollution. The production of cars and auto infrastructure consume large amounts of finite resources and their disposal generates vast quantities of waste. Clearly, a new transportation system is required that is environmentally sustainable and does not restrict freedom of movement.

Potential Solutions to the Urban Transportation Problem

The government and the market possess the power to develop an environmentally friendly urban transportation system that can efficiently handle travel flows. The government can mandate change or provide incentives for change. The market reacts to both government regulation and consumer demand with the aim of continually maximizing profits. The market drives our consumer society. The solution to the urban transportation problem will require a balance between government intervention and market forces. Proposed solutions to the causes and effects of single modal choice, unevenly distributed travel demand and finite automobile transportation infrastructure are:

1> Technologically fixing the automobile
2>Increasing the capacity of auto infrastructure
3> Changing the transportation modal split

Technological Fix

The most strongly advocated solution to the environmental problems of the automobile is technologically taming the emissions produced by the automobile. The removal of lead from fuel has solved the lead pollution problems of the automobile. The amounts of carbon monoxide emitted by the automobile has been largely tamed by the introduction of the two and three way catalytic converter. The catalytic converter does not perform its role until it reaches a high temperature, making it relatively useless on short errand trips. In addition, carbon monoxide emissions are replaced by higher carbon dioxide emissions,

which is a greenhouse gas. The ongoing evolution of more and more fuel efficient internal combustion engines has led to a reduction in energy depletion per vehicle and an increase in nitrous oxides emitted. However, most pollution reduction brought about by technological improvements has been offset by the explosive growth in the size of the automobile population. The development of an electric propulsion motor will solve the pollution shortcomings of the internal combustion engine. The electric vehicle solves only one part of the urban transportation problem, pollution. The large problems of congestion, land consumption and people and property damage are not resolved.

Increasing the Capacity of Automobile Infrastructure

The urban transportation problem of congestion can be solved by increasing automobile infrastructure capacity. The flows on a network can be increased in two ways: increase the physical dimensions of the network links or improve network link efficiency. The unusual level of affluence that facilitated the mass highway construction era does not exist today. Governments have already accumulated massive deficits and tax-payers are not willing to incur the costs associated with building more highways. The civil liberties movement and the current environmental conservation movement provide formidable obstacles to solving the problem of congestion with this solution. Increasing capacity only improves the utility of the car and with time, congestion will reappear as more demand is placed on the transportation network link.

Network link efficiency can be improved by the use of traffic management technology. Intelligent Vehicle Highway systems (IVHS) is the solution currently supported by the highway engineering community as the Interstate Highway Construction Program has been largely completed[9]. The end of the cold war has left defence contractors without a future so they are supporting IVHS as an avenue for future profits. IVHS is a short term solution to congestion that will not be ready for introduction for decades. Development has just begun. The pursuit of this solution is purely influenced by politics. Augmenting capacity will only solve the problem in the short term. The problem to solve is how to manage continually increasing unevenly distributed flows of people across increasing distances. The solution lies in changing the modal split and reducing the number of work trips generated.

Changing the Modal Split

In theory the urban transportation problem may be solved by

changing the modal split of work trip demand. Government mandates and incentives could be used to change people's behaviour by making modes other than the automobile more attractive and decreasing the utility of single occupancy automobile travel. The alternate travel modes are: high occupancy vehicles (car pooling), public transit and virtual transportation (telecommuting).

Government mandates and incentives to change the travel behaviour of people have been ineffective, expensive and frustrating. The dispersion of economic activities across urban space in combination with residential land use patterns make public transit an unviable option. Park and ride schemes make trunk travel by public transit possible but congest local areas and fail to solve the problem of reaching the work destination from the nearest transit stop. Public transit is also very expensive to build and maintain. Peaking demand requires a large fleet which is under used most of the day and night. Virtual transportation is a simple solution that is inexpensive, requires little government intervention and drastically reduces emissions.

The Virtual Transportation Solution

Virtual transportation has the potential to solve the urban transportation crisis within a short period of time with little cost to employees and employers and can be implemented with little government incentives and mandates and at a reasonable cost. The solution attacks the root of the problem: the temporally concentrated generation of work trips across ever increasing distances and their modal split. The emergence of a global information economy and recent advancements in technology make virtual transportation an inevitable solution to the urban transportation crisis.

Virtual Transportation Defined

The concept of virtual transportation assumes many names and associated definitions. Telecommuting and teleworking are the two most common labels of the phenomenon. Virtual transportation is still in its infancy, which results in definitions changing as information technology evolves and new applications for the technology are discovered. For the purpose of this paper, virtual transportation (telecommuting) will be

defined as:

> "a philosophy of allowing an employee to perform required tasks full-time or part-time from his\her choice of location by using information technology".[10]

Telecommuting removes or reduces the geographic constraints of time and space. Virtual transportation will reverse the work-location trends of the 19th century industrial revolution a systematic reappraisal of the spatial organization of economic institutions.[11] Information technology is ready today to make virtual transportation possible.

Information Technology

A full review of all of the present and prospective developments in Information Technology is beyond the scope of this paper. A brief review of some of the key technologies is included to demonstrate the feasibility of telecommuting today. The adoption of evolving information technology has accelerated due to increased power, user friendliness, dependability, exposure and the decreased cost of the technology. [12]

The backbone of information technology infrastructure are the communication networks. They are comparable to freeways and road systems. There is a hierarchial ordering of communication flows. Telephone, Cable TV (coaxial cable) and wireless are three possible interactive telecommunication mediums. The key network for telecommuting currently under development is the ISDN - Integrated Services Digital Network. ISDN will enable data in any format (voice,data, images, music and video) to travel on the same cables. The carrying capacity of a network is measured by the bandwidth of its cables and the type of switches. The telephone line with twisted copper cable has a very low carrying capacity and is a constraint to connecting the home to ISDN. Cable TV coaxial cable has 100 times the carrying capacity of the phone line. However, current infrastructure is not point to point and two way. Expensive upgrading will be required if the cable TV network is to be used as a communication corridor. Fiber optics are the superhighways of the telecommunications network. They carry 10000 times the capacity of the copper telephone line due to their broad bandwidth.[13] Satellite corridors are faster than coaxial cable and less than fiber optics. The recent substantial enhancements in transfer capacity of communication networks have been the result of the introduction of digital switching (versus

analogue) data compression technologies and fibre optic trunk lines. Digital switching has increased the ability of service providers to manage data streams. Service providers hope to have a fully digital trunk network shortly.[14]

The weakest link in telecommunication infrastructure is the absence of a complete fibre optic network. The integrated services digital network is the only medium that can handle the anticipated flows of data and transmit them at a reasonable speed. The cost of replacing copper telephone cables is prohibitively expensive due to the decentralized low density residential pattern of contemporary cities. Virtual commuters have different demands for services. Many economic activities could be transferred to the home without the need for ISDN services. The construction of neighbourhood telecommunication centres (teleports) will make telecommuting possible in the short term for those economic activities that require full time or periodic ISDN services, without the massive expenditure of upgrading the whole network immediately. Cellular technology may be another solution. Currently it is not sufficiently developed to access its potential as a communications network.

Telecommunication Centres

Telecommunication centres are a solution to many of the problems limiting the introduction of telecommuting. In theory, they would be structured in the form of a community office. They would be evenly distributed across urban residential space and the size of each 'teleport' would differ according to local demand. They would be equipped with all the office equipment, amenities and information technology required to perform any information economy activity. Hardware would include: personal computers with fax-modems and video-phones and a variety of software, audio telephones with numerous services (call forwarding, waiting, voice mail), photocopiers, scanners and video-conference equipment. A single fibre optic highway could support 20 000 telecommuters. Gann[15] has hypothesized a 'teleport' for 1000 telecommuters would require: 1000 computers, 200 fax terminals, 200 ISDN circuits, 200 data only circuits, and 10 video conference suites.

The strength in the theory of telecommunication centres lies in their ability to supplement the amenities available in the home at an easily accessible location and a short distance away. The problematic part of the idea lies in financing. Who should pay to build and operate these facilities:

municipal government, federal government, Industry or the users? A combination of sources is required. Ideally they should be controlled by a government body to ensure equal access. Financing could include usage tolls paid by employers, a reasonable membership fee for employees and heavy subsidization by fuel taxes or even road pricing. Road pricing would be an excellent way to generate the capital required for the initial construction of facilities.

Telecommuting Explained

Coordination between business and government is necessary to make the transition to virtual commuting smooth. The success of the transfer from physical to virtual transportation largely relies on correctly selecting economic activities that are telecommunicable given the technology available at the specific period in time. Telecommuting will require time and creativity to implement. The information workforce will need time to adjust and social impacts will need to be mediated.

Business and government will receive benefits from the adoption of telecommuting. The reduction in work trip generation and the reduction in its temporal and spatial components will solve the urban transportation problem. The reduction of temporally concentrated demand on freeway networks will eliminate congestion and reduce emissions. The reorganization of society by telecommuting will make any locality, region or country that much more competitive in the global information economy. The province of New Brunswick was first in Canada to fully digitize its communication networks and as a result has been successful in attracting high technology and information based employment. Economic growth can only benefit the people and business. Telecommuting has been demonstrated to decrease the costs of production by lowering capital expenditure (land) and increasing worker productivity. Virtual transportation also has the ability to increase worker satisfaction. Table 1 demonstrates the potential of telecommuting to both satisfy employers and employees. The idea of the neighbourhood telecommunications centre would solve the majority of problems for both employers and employees. The centre is located close to home and open 24 hours a day. Video phone technology will remove the feeling of disconnectedness with colleagues but will bring personal attire to the forefront again. The teleport in theory will create a closer sense of community
and maintain a sense of corporate culture. Unfortunately, not all workers are eligible telecommuters.

TABLE 1 ADVANTAGES AND DISADVANTAGES OF TELECOMMUTING[16]

The Individual Employee

ADVANTAGES

-Less commuting time and costs (monetary and physical)
-Flexible work schedule and more time for home and family
-Ability to plant roots in a community and work in a wider catchment area
-Employment opportunities for the disabled
-More effective use of face to face contact
-Increased job satisfaction

DISADVANTAGES

-Isolation - lack of interaction with co-workers
-Potential of gaining experience from peers is reduced
-Chances for personal promotion are reduced with horizontal organizational structure
-Less political power within the organization
-Danger of exploitation -work during leisure hours to make deadline
-Loss of contract status, benefits?
-Pressures of family and work in the same sphere

The Employer

ADVANTAGES

-Time and space advantages of telecommuting in a global market
-Early birds or night owls can work when they are most productive
-Higher productivity (25% average)
-Lower facility/amenity costs
-Easier to recruit staff - contract staff for rushes.
- wider catchment area to recruit staff
-Compliance with local commuting reduction regulations
-Reduced labour militancy

DISADVANTAGES

-incompatibility with current work ethos, corporate identity
-Greater demand to supervise and monitor progress
-Technical problems of transporting or transferring work.
-Cost of setting up telecommuting eg. training
-Security of information
-Loss of productivity in transition period

The majority of current information production and processing positions are telecommunicatable. The actual potential population differs substantially within the literature. Estimates range from ten percent to seventy five percent. Quantifying the number of present telecommuters has also been problematic given the variety of defining parameters. The population has been measured as between 10 000 and 10 000 000 in the United States. The key determinants for evaluating the potential

transferability of an economic activity to a remote site are:

1> The degree that the activity is information based
2> The amount of physical interaction required by the activity
3> The measurability of output

Table 2 illustrates some of the many positions that can be performed from a remote site using information technology. Clerical and professional occupations are the simplest to transfer to remote sites.

TABLE 2 EXAMPLES OF JOBS WITH TELECOMMUTING POTENTIAL[17]

Accountant	Journalist
Applications programmer	Lawyer
Architect	Market analyst
Auditor	Purchasing manager
CAD/CAM engineer	Realtor
Civil engineer	Statistician
Data entry clerk	Stock broker

Virtual transportation has the potential to solve the urban transportation problem by reducing the number of work trips generated and the magnitude of their spatial and temporal components. The development of a network of telecommunication centres, one within each neighbourhood would solve many of the problems associated with using one's home as an economic activity location. The full adoption of telecommuting may take up to ten years. Gradually, air quality would improve and congestion would become a memory of the past.

Virtual Transportation Performance Appraisal

Virtual transportation presently exists more powerfully as an ideological construct than as a reality. The phenomenon is in its infancy. There have been few published empirical studies verifying the theoretical trip reduction impacts of telecommuting. Moreover, it is not captured in any comprehensive statistical databases. An empirical study by Lund and Mokhtarian in 1994 and one by Mokhtarian, Handy and Solomon in 1995 show results that correlate with trip reduction potentials associated with virtual transportation.[18] The total amount of vehicle miles travelled, peak period trips, energy consumption and emissions decrease significantly. In addition, there appears to be no new generation of non-work trips now

that the automobile is liberated from its work trip duties. Very small reductions in single occupancy vehicles have been found to result in exponential decreases in congestion.[19] There appears to be no empirical data published on the impact on the temporal and spatial components of work trips by telecommunications centres, in an urban region. The true long term impacts of telecommuting remain unknown and many hypotheses are simply extrapolations from limited short term empirical studies.

Substantial case histories and anecdotal evidence suggests that telecommuting leads to gains in work productivity and job satisfaction.[20] A limited empirical study by DuBrin reveals the same results. The reoccurring themes that emerge from case studies of pilot projects are a strong resistance from management to changing work structures and a fear by workers of forced adoption of telecommuting which eventually leads to exploitation. Voluntary programs show disappointing participation rates. If telecommuting is to solve the urban transportation problem, government intervention is required to stimulate employers and employees to switch over from physical commuting to virtual commuting and legislation needs to be amended to protect workers from exploitation.

Public Policy Issues

Government legislation is required to maintain affordable and universal access to the networks of the informational economy. Information technology is very difficult to control. The cost of upgrading the information economy should be borne by the service carriers since they stand to profit. The only way to ensure equity in a telecommunicating workforce is for a government body to control price stability and service quality.

Zoning regulations and employment legislation need to be amended to address the potential problems of implementing telecommuting on a large scale. Current zoning practices and lease agreements forbid many economic activities from locating within the home. Many white collar occupations do not present the kinds of problems the codes were designed to control. Employment legislation currently contains few provisions to protect the telecommuter. Telecommuting provides the opportunity to employers to exploit workers by shifting the costs of information technology to employees, forcing employees to work for less pay, demanding unreasonable production schedules and cutting employee benefits. Regulatory protection is

necessary. Income tax regulations discourage the use of the home as a location of work. Business deductions are difficult to claim.

Societal Impacts

Virtual commuting has the potential to improve the quality of life of society. However, it may also generate negative impacts on the city. Improvements to the quality of life were summarized in Table 1. Potential negative impacts to the city include local traffic congestion, further decentralization of the city and creating a central business district real estate crisis. The introduction of telecommunication centres into the residential landscape may congest the roadways of neighbourhoods. Telecommuting enables employees to locate anywhere telecommunication networks service. Wilderness and rural regions would be theoretically threatened should service be provided to these regions. Urban planning decisions motivated by maximizing development could be a serious threat to the remaining fragile ecosystems untouched by development.[21] Telecommuting will decrease the need for commercial space by a magnitude never witnessed before. A substantial portion of the wealth of our capitalist society lies in real estate, especially commercial real estate. A vacancy crisis would fuel a large resistance by property owners against the mass introduction of telecommuting. Creative uses must be found for vacant commercial space to avoid the development of urban wastelands within the city.

Conclusion

The form of a city dictates what kind of transportation system is feasible. However, transport itself is a determinant of the shape that a city assumes. The relation between the two is a complex blend of cause and effect.[22] The most advocated solutions to the current urban transportation problem fail to address this relationship. The cause of the problem is temporally concentrated transportation demand. Physical transportation is not feasible given the urban landscape of a contemporary city. Virtual transportation is the mode of travel of the future. Telecommuting ignores time and space. The infrastructure of this transportation consumes negligible amounts of space and is environmentally sustainable. Telecommuting may restore the sense of place and community that was taken away by the automobile.

The introduction of road pricing could be used as a revenue generating mechanism to build a network of telecommunication centres. Road pricing would provide a strong incentive to commuters to substitute physical travel with virtual travel. Telecommunication centres could be accessed with small, environmentally friendly neighbourhood electric vehicles.[23] Society would not have to relinquish the tradition of auto mobility.

The real power of the telecommunicating phenomenon lies in its creative applications. Virtual shopping and virtual training are examples. The acceptance of telecommuting by society is evident in the strong demand for remote access to graduate university programs. People want the convience of remote access across any distance. Time is a finite resource in people's lives, it should not be wasted travelling. The automobile changed the world, virtual transportation will do the same.

Notes

[1] P. Keller and M.C. Rotach, "Opportunities and Risks Involved in Telematics for Traffic and Settlement - A Technology Assessment", in *Telematics - Transportation and Spatial Development*, ed. H.M. Soekkha (The Netherlands: Utrecht, 1990), 21-36.

[2] S. Hanson, *The Geography of Urban Transportation* (New York: The Guilford Press, 1986).

[3] Ibid., p. 106.

[4] D. Sperling, *Future Drive* (Washington, D.C.: Island Press, 1995).

[5] G. Ritter, and S. Thompson, "The Rise of Telecommuting and Virtual Transportation", *Transportation Quarterly* 48, No.2(1994): 235-248.

[6] Sperling, op. cit.

[7] Hanson, op. cit.

[8] Statistics Canada, *Human Activity and the Environment* (Ottawa: Ministry of Industry, Science and Technology, 1994).

[9] Ibid., p.21.

[10] S. Bramson, "Attitudes Towards Telecommuting: A Study of the Expectations and Feelings Associated with the Introduction of Telecommuting", Unpublished MBA research project, Montreal, Concordia University, 1991.

[11] M. Hepworth and K. Ducatel, *Transportation in the Information Age* (London: Transet,

1992).

[12] Bramson, op. cit.

[13] J.L. Swerdlow, "Information Revolution", *National Geographic* 188, No.4 (1995): 5-36.

[14] T. Melnik, "Marketing Telecommuting as a Product", Unpublished MBA business research project, Montreal, Concordia University, 1993.

[15] D.M. Gann, "Buildings for the Japanese Information Economy", *Futures* 23, No.5 (1991): 469-481.

[16] A.J. Durbin, "Comparison of the Job Satisfaction and Productivity of Telecommuters Versus In-House Employees" , *Psychological Reports* June, Part 2 (1991): 1223-1234; M. Gray, N. Hodson, and G. Gordon, *Telecommuting Explained* (New York: John Wiley & Sons Inc., 1993); Hepworth, and Ducatel, op. cit.; M. Mehlmann, "Social Aspects of Telework", in *Telework: Present Situation and Future Development of a New Form of Work Organization*", eds. K.B. Korte, S. Robinson, and W.J. Steinle (Amsterdam: Elsevier Science Publishing Company Inc., 1988); Melnik, op. cit.; M.H. Olson, "Organizational Barriers to Telework", in *Telework: Present Situation and Future Development of a New Form of Work Organization*", eds. K.B. Korte, S. Robinson, and W.J. Steinle (Amsterdam: Elsevier Science Publishing Company Inc., 1988); Staddon, op. cit..

[17] J. B. Sneider, and A. Francis, *An Assessment of the Potential of Telecommuting as a Work-Trip Reduction Strategy: An Annotated Bibliography* (Chicago: Council of Planning Librarians, 1989).

[18] J.R. Lund, and P. Mokhtarian, "Telecommuting and Residential Location: Theory and Implications for Commute Travel in a Monocentric Metropolis", *Transportation Research Record* No.1463 (1994): 10-13; P. Monkhtarian, S.L. Handy, and I. Salomon, "Methodological Issues in the Estimation of the Travel, Energy, and Air Quality Impacts of Telecommuting", *Transportation Research Part A. General* 29, No.4 (1995): 283-302.

[19] G. Ritter, and S. Thompson, "The Rise of Telecommuting and Virtual Transportation", *Transportation Quarterly* 48, No.3 (1994): 235-248; T. Weijers, R. Meijer, and E. Spoelman, "Telework remains Made to Measure", *Futures* 24, No.10 (1992): 1048-1055.

[20] DuBrin, op. cit.

[21] J.H. Snider, "Information Highway as Environmental Menace", *Futurist* 29, No.2 (1995):16-20.

[22] D. Sudjic, *The 100 Mile City* (Great Britain: Harper Collins Manufacturing Glasgow, 1993).

[23] Sperling, op. cit.

POWER FAILURE

An Examination into the Political Economy of the Electric Vehicle

By: Daniel Schwartz

Introduction

In the late 1890s and early 1900s, the electric vehicle (EV) stood poised, ready to become the world's vehicle of choice to succeed the horse-and-carriage. By the late 1920s, however, it was clear that the EV had yielded to its arch rival - the internal combustion engine vehicle (ICEV). Since this time, numerous opportunities for the resurgence of the EV have presented themselves, including the oil crises of the 1970s and the environmental movement of the 1970s and 1980s. These opportunities have also failed to give way to the EV. Most recently, California announced its zero emissions vehicle (ZEV) mandate, requiring that a growing percentage of each major automaker's sales in California be ZEVs. The required percentage is set at 2 percent for 1998, increasing to 5 percent in 2001 and 10 percent in 2003[1]. Again, however, the fate of the EV remains questionable, as automakers have persistently resisted the ZEV mandate, and consumers remain reluctant to adopt the new technology[2].

This paper will explore the political economy of the EV, and attempt to determine the reasons behind its chronic failure to become a mainstream technology. Although the paper will focus on the last decade, the earlier historical periods of importance to the EV (late 1800s, early 1900s, and the 1970s), will serve as valuable guide in interpreting present trends.

A common view represented in the popular discourse on this topic, is that EVs have failed simply due to a lack of technology: the technological constraint

---

[1] EVs are the only technological alternative for automakers to meet the ZEV mandate, since no other zero emission vehicle technologies are workable at the current time.

[2] Noel Perrin, Life with and Electric Car Sierra Club Books, San Francisco, 1994, p.198.

theory.[3] Some maintain that vested interests, such as the major automakers and oil companies, have collaborated to ensure the survival of the auto-industrial complex and the longevity of the ICEV: the conspiracy theory.[4] Still others claim that the EV has never become popular because it cannot match the performance of the ICEV: the consumerist theory.[5] The findings of this paper suggest that all of these responses are incomplete. There is no one overwhelming cause for the failure of the EV. To rely on one response is to divorce the failure of the EV from the larger socio-political framework in which it has evolved.

The literature on the political economy of technology distinguishes between the phases of technological change. For purposes of this paper, it is assumed that the EV represents an innovation that has not yet been widely diffused. Furthermore, the EV is a radical innovation, because it represents a significant break with the past. The EV is not merely an extension of the ICEV; it challenges fundamentally the hegemony of current automotive technology. Essentially, then, the EV embodies a new technological trajectory. As with any new environmental technological trajectory, there are many impediments operating in a multiplicative and highly interactive nature. This paper examines these impediments to a new technological trajectory, and applies them to the case of the EV. The preponderance of these factors stem from Kemp and Soete's `selection environment'. `Selection environment' emphasizes "the institutions involved and the mechanisms behind the selection of an innovation," and recognizes that "innovations [are] embedded in a technoeconomic system (which involves socioinstitutional elements).[6]"

---

[3] See for example, Deborah Gordon, Steering A New Course: Transportation, Energy and the Environment Union of Concerned Scientists, Washington D.C., 1991, p.96.

[4] See for example, Freund and Martin, The Ecology of the Automobile Black Rose Books, Montreal, 1993, p.136.

[5] See Michael Brian Schiffer, Taking Charge: The Electric Automobile in America Smithsonian Institution Press, Washington and London, pp.1-4.

[6] Kemp and Soete, 1992, p.445. The term `selection environment' is borrowed from Nelson and Winter (1977).

The paper concludes that failure for the dispersion of EV technology can be traced to a number of factors, including costs to industry and concern over profitability, apprehension over appropriability conditions, lack of technological opportunities and technological feasibility, a shortage of research and development, and consumer resistance. Factors in which there is mixed evidence (i.e. in which the impediment is not clear-cut) include, the mode of regulation, the effect of `clusters of technological trajectories', and market structure.

The first section of this paper will offer a brief history of the EV. The second section will take a detailed look at the factors which have impeded the diffusion of EV technology. The factors in which there is mixed evidence will be presented in the third section. This will be followed by a summary and some policy implications.

A Brief History of the Electric Vehicle

The development of the EV has been a history of peaks and troughs. From the first EV ever built in 1873 to General Motor's sporty 1991 Impact, this technology has continuously come in and out of the market, never quite able to solidify its position amongst automakers and drivers.

By the end of the nineteenth century, EVs actually constituted the mainstream in automotive development. Of the automobiles manufactured in the US in 1900, 1,575 were EVs, and only 936 units were gasoline powered.[7] With continuing improvements in the performance of the ICEVs and persistent limitations in battery technology, however, the gap between the performance of EVs and ICEVs widened. The appearance of the Ford Model T in 1909 heralded the popularisation of ICEVs and their long domination of the market. By the 1930s, despite a combined effort of both Thomas Edison and Henry Ford himself, the EV had disappeared from the mainstream market.[8]

The EV regained some popularity in Japan during the mid-1930s when wartime restrictions on the use of gasoline became severe, and again in the US and

---

[7] International Energy Agency, Electric Vehicles: Technology Performance and Potential, OECD, 1993, p.19.

[8] Ernest H. Wakefield, History of the Electric Automobile, Society of Automotive Engineers Ltd., 1994, pp.211-224; see also Michael Brian Schiffer, 1994, pp.153-174; and Sheldon Shacket, The Complete Book of Electric Vehicles, Domus Books, 1981.

Europe during the oil crises of 1974 and 1979.[9] During the late 1980s, a resurgence of the environmental movement, sparked especially by concerns over global warming, reignited interest in the EV. Although the environmental movement has wanned somewhat in the 1990s, California's ZEV mandate has once again peaked interest in EV technology.

Impediments to a New Technological Trajectory

What has prevented the diffusion on EV technology, especially in light of the profound solutions it offers to the environmental quandary? The following is an examination of the factors that have inhibited this `green' technological trajectory.

Costs to Industry and Profitability

Monetary costs and concerns over profitability are a central impediment to any `green' technology. Although macroeconomic trends suggest that pollution control need not be economically problematic, initial capital outlay for environmental technologies remains an obstacle for some industrial sectors.[10] Kemp and Soete (1992) claim that economic uncertainties extend beyond initial capital outlay to include broader economic concerns:

> "Many firms...will be reluctant to adopt cleaner techniques because of the economic risks involved. Production routines and procedures have to be changed and employees have to learn and become familiar with the new technology."[11]

Ultimately, industry concerns boil down to a question of profitability.

Such consternation over economic risk, has been a major impediment to the adoption of EV technology by automakers. Automobile manufacturers argue that each EV will cost an extra $10,000 to $20,000 US above what a comparable ICEV would cost.[12] They contend that such a cost would greatly reduce demand for their product. Moreover, automakers might be forced to take a loss on EVs and

---

[9] International Energy Agency, 1993, pp.20-21.

[10] OECD, "The Technology/Economy Programme, Technology and the Economy: The Key Relationshipch.9, 1992, pp.203-205.

[11] Kemp and Soete, 1992, p.451.

[12] Daniel Sperling, Future Drive: Electric Vehicles and Sustainable Transportation Island Press, Washington D.C., 1995, p.138.

compensate for this loss by raising the price on ICEVs[13]. Sperling (1989) argues that initial costs for EVs conceal the fact that EVs will actually save the consumer money if full "life-cycle costs" are considered. Moreover, mass production of EVs is likely to reduce the costs significantly due to economies of scale. Nonetheless, the cost impediment for automakers is not easy to circumvent.

Appropriability Conditions

Related to the issue of costs and profitability, is the issue of appropriability conditions. Kemp and Soete (1992) site appropriability conditions as a significant determinant in the decision to generate clean technologies. Appropriability conditions refer to "the means through which a firm can reap returns from its innovations, and hold off other firms from eating too much into these returns."[14] Appropriability conditions are influential because a firm will not undergo the financial risks of developing a `green' technology if they cannot reap the benefits.

There is some evidence to suggest that the major automakers' apprehension over appropriability conditions has delayed the diffusion of EV technology. Schiffer (1994), claims that a patent-sharing agreement instituted in 1914 by the National Automobile Chamber of Commerce, has stunted technological innovations for the past 100 years, because automakers "knew that what they invented in one year could be industrywide practice by the next."[15] Investing in new technologies did not pay because a company could gain no more than a temporary advantage over their competitors.

More recently, the 1991 creation of the United States Advance Battery Consortium (USABC), a joint business-government R&D effort, has come with certain stipulations regarding patent rights:

> "Battery manufacturers who want [the USABC's] money must agree to ...relinquish patent rights to any innovation they make under contract. The [Department of the Environment] has largely ignored cries that this stipulation...flouts a federal law requiring the government to favour small companies over giants in such matters...."[16]

---

[13] BusinessWeek, "Electric Cars: Will they Work? And Who will Buy Them?, May 30, 1994, p.111.

[14] Kemp and Soete, 1992, p.449.

[15] Michael Brian Schiffer, 1994, p.171.

[16] The Economist, "The Electric Car's Achilles'

Kemp and Soete (1992) note that government action of this sort in not uncommon with respect to environmental technologies: "In view of the public interest in rapid diffusion of clean technology, there is probably more government willingness to limit appropriation than in the case of `normal' technology."[17] Although far from conclusive, the government endorsed stipulations, regarding the USABC, are an indication of the automaker's anxiety over appropriability.

Technological Opportunities and Feasibility

Intimately tied to the issue of costs, profitability, and appropriability conditions, is the question of technological opportunities and feasibility. Following Dosi (1977), Kemp and Soete (1992) specify technological opportunities as a determinant of the decision to generate a clean technology: "As in the case of `normal' technology, these opportunities depend on accumulated scientific knowledge, available equipment and capabilities in organizations."[18] While technological opportunity is a measure of `technical achievability', technological feasibility also incorporates the issue of costs. A technology that is technically achievable, but is overly costly, is not technologically feasible. Both opportunity and feasibility are significant when examining the case of EVs.

As the historical account outlined above suggests, EVs have been the focus of much research for the last century. Although incremental improvements have been made over this time span, no breakthroughs have occurred, and the EV remains an inferior product to the ICEV in terms of driving range, speed, acceleration, and cost.[19] Sperling (1995), maintains that EVs have undergone major improvements in the last five years, in terms of the motor, power electronics, transmission, aerodynamics, tire friction, and other auxiliaries. For example:

> "...advances in power electronics have made possible alternating-current (ac) drives that are cheaper, easier to maintain, and more compact, reliable, efficient, and adaptable to regenerative braking than the direct-current (dc) systems used in

---

Axle," September 19, 1992, p.101.

[17] Kemp and Soete, 1992, p.449.

[18] Kemp and Soete, 1992, p.448.

[19] See Sperling, 1995, pp.46-55; see also Schiffer, pp.175-190.

virtually all [EVs] in the early 1990s.[20]

Charging devices, a major concern for EV proponents, have also seen major improvements in the last decade.[21] Schiffer (1994) concurs on the progress of the EV: "The upshot...is that electric vehicle technology has improved enough so that a viable car can be made today."[22]

By all accounts, however, batteries remain the 'weak link' of the EV. Lead-acid batteries are likely to dominate the EV market well into the next decade, because of their relatively low cost and high reliability. Improvements to lead-acid batteries are also continuing, and are likely to be significantly advanced by 1998.[23] Nevertheless, lead-acid batteries remain a questionable technological alternative in comparison to the ICE. Using lead-acid batteries, EVs will have a maximum driving range of 70 miles, and far less acceleration and power than the ICEV. Even the more promising battery technologies of the near future, such as nickel metal hydrade, zinc bromide, and zinc air, are inferior in performance to the ICE: "All chemical batteries that might be mass produced in the next decade seem likely to suffer from limited range, feeble acceleration, expensive replacement every three to five years and difficulties over recycling and disposal."[24] The technological opportunity clearly exists, but relative to the ICEV, it not difficult to see why automakers are hesitant to produce EVs, and consumers are reluctant to endorse them.

When costs are taken into consideration, the technological feasibility of EVs is also called into question. While current lead-acid batteries are relatively inexpensive, they require replacement every 20,000 miles.[25] Future technologies do not fare much better. Nickel-Metal Hydride, one of the more promising alternatives in terms of performance, is currently priced beyond market reach.

---

[20] Sperling, 1995, p.47.

[21] BusinessWeek, May 30, 1994, p.105.

[22] Schiffer, 1994, p.174; see also Freund and Martin, 1993, p.148.

[23] BusinessWeek, May 30, p.107; see also CQ Researcher, "Electric Cars," July 1993, p.580.

[24] The Economist, September 19, 1992, p.101.

[25] Sperling, 1995, p.48; see also Gordon, 1991, pp.95-98.

Research and Development

While technological opportunities and feasibility have been substantial impediments to the adoption of EVs, there is evidence to suggest that the deficiency in technological advancement is due to a lack of R&D.

The notion that research and development can be a significant factor in determining whether or not `green' technologies are adopted is widely recognized.[26] It is not only essential that R&D for `green' technologies take place, but that the funds be properly earmarked as well.

Sperling (1995) argues that EV technology has been largely ignored ever since the ICEV became the auto of choice in the US:

> "It's no surprise that today's electric vehicles cost more and perform worse than their gasoline counterparts. Gasoline cars have benefited from a century of intensive development; electric cars have been virtually ignored for over seventy-five years. Even today, gasoline cars profit from billions of dollars of research every year while electric vehicles receive a tiny fraction of that."[27]

Prior to California's ZEV mandate, less than $10 million a year was being spent on battery research aimed specifically at the vehicle market, "and most of that went to improving the primitive techniques employed in building battery packs for golf carts and slow-speed industrial vehicles."[28] James Womack, director of MIT's International Motor Vehicle Program, claims that R&D funds issued to automakers between 1969 and 1991, and aimed at alternative concepts such as EVs, were "a complete and total waste," because "the industry had no real interest in these ideas; they just took the money."[29] Jacoby and Steinbruner (1973), note that this lack of R&D is not surprising, given that the focus of automaker laboratories has been on qualities the market tends to value - speed, acceleration, fuel economy, and smooth and reliable performance: "It has never been concerned with developing advance technologies for the purpose of controlling emissions."[30]

---

[26] See for example, OECD, 1992, pp.194-198.

[27] Sperling, 1995, p.36.

[28] Sperling, 1995, p.49.

[29] Nadis and MacKenzie, 1993, p.163.

[30] Jacoby and Steinbruner, Clearing the Air: Federal Policy on Emissions Control Ballinger

Although it is somewhat speculative to assume that EV technology would have advanced dramatically had more R&D occurred, the case is strengthened considerably when one considers the advancements that have taken place since automakers learned that California law would soon force them into mass-producing EVs. In direct response to California's ZEV mandate, the USABC was launched in 1991. With funding provided by the federal government (50 percent), the Big Three (23 percent), battery companies (20 percent), and electric utilities (7 percent), USABC has an R&D budget of $230 million to be spent by 1995 and $100 million more per year into the early years of the next century.[31] Similar battery consortia were also formed in Japan and Europe after the creation of USABC. These R&D projects have paid off handsomely for EV technology. For example, nickel-metal-hydrade and lithium-based battery technologies have advanced steadily. This technological progression was duly noted by the California Air Resources Board (CARB), who responded after three months of public debate and a two-day hearing in Los Angeles in May of 1994, in which Ford voiced their opposition to the ZEV mandate:

> "We heard from no one who claimed the mandate had not accomplished its stated objective of stimulating technological development and innovation. While [EV] and battery technology may not have advanced much between the turn of the century and the 1980s, there is no doubt that tremendous advancements have occurred since we adopted the [ZEV mandate]. We heard over and over again that the mandate caused or contributed to these advancements."[32]

As noted above, the allocation and use of funds is also an important factor. While the ZEV mandate incited substantial R&D, Sperling (1995) argues that the central goal, of the USABC and the Japanese MITI consortium, is the development of a battery that would allow EVs to compete head-to-head with ICEVs. `The Economist' notes that USABC research and development is limited to "medium- and long-term work on batteries," to the exclusion of research on current lead-acid technology.[33] This emphasis on high power batteries has diminished R&D on low

---

Publishing Company, Cambridge, Mass., 1973, p.52. See also Business Week, May 30, 1994, p.110.

[31] Sperling, 1995, p.49.

[32] Sperling, 1995, pp.40-41.

[33] The Economist, September 19, 1992, p.101.

cost batteries, a factor which has delayed the introduction of competitively priced EVs.

Consumer Resistance

Kemp and Soete (1992), assert that the decision, on the part of industry, to generate a clean technology is determined in part by the potential market demand. The market demand, in turn, depends in part on the willingness of consumers to change their attitudes and modify social norms. Because consumers appear as diffuse sources of pollution, it is not always possible for government regulations to directly influence behaviour. Consumer acceptance of `green' technologies, therefore, relies heavily upon structural social change.

The environmental movement of the 1970s and 1980s represents a structural social change. While heightened societal consciousness of environmental issues has paved the way for California's ZEV mandate, however, Sperling (1995) cautions that consumers require more than an ecological justification for purchasing EVs. Detracting from the attractiveness of environmentally-sound EVs, is the fact that consumers have not been able to overcome some ingrained misperceptions, and some very real concerns about the EV. A number of surveys quantify consumer resistance to the EV. Of the 4,512 respondents to a recent study, for example, 36 percent said they "would definitely not consider purchasing an EV," and another 36 percent indicated they "would probably not consider an EV for their next automobile purchase."[34] A 1990 survey by the Ford Motor Co. found that only 1 percent of the American population would buy EVs.[35] The history of the EV, reveals the difficulties in attempting to change consumer perception of the automobile's function. The ICEV won out over the EV in the early 1900s, because consumers were enamoured with the concept of `touring'.[36] This perception remains a near intractable quandary for the EV in the 1990s: "The crisis before us is largely cultural and perceptual; we may simply have to change our image of what a car looks like and does."[37] John R. Dabels,

---

[34] CQ Researcher, July 1993, p.583. Note that 19 percent of the remainder of respondents were lukewarm to the idea of EVs, while only 6 percent indicated that they would definitely consider purchasing an EV.

[35] Sperling, 1995, p.58.

[36] See Schiffer, 1994.

[37] Nadis and Mackenzie, 1993, p.91.

director of market development for GM Electric Vehicles, has recognized that acceptance of the EV has been hindered by the consumers' inability to overcome a widely held fallacy - that automobiles are used *primarily* for `touring':

> "A common misconception among drivers is that they drive for long distances at a time. Our research shows that 84 percent of the drivers studied drove less than 75 miles a day, a total daily mileage well within the range of an electric vehicle.[38]"

EV endorsement has not only suffered because of a misperception of travelling distance, but also because of a misperception of travelling speed. The US Federal Highway Administration estimates that the typical home-to-work commute takes place at an average speed of 31 mph, far below the average speeds associated with `touring'.[39]

Schiffer (1994) cautions that the single most important lesson that the history of the early electric vehicle has to offer, is that the choice of a car technology is influenced by the *extreme*, not the average, anticipated use:

> "Middle-class Americans in the teens flocked to the gasoline car, even though [EVs] could meet all of their urban transportation needs, because they anticipated touring every once in a while. In the showroom, today and tomorrow, ordinary Americans will not be thinking about the 20-mile trek back and forth to work each day, but about the few uncommon occasions...when an extra-long range is essential.[40]

Misperceptions over driving habits explains only part of the consumer resistance to EVs. Drivers have also expressed legitimate concerns over price, quality, safety, and infrastructural requirements.[41] For example, a recent survey found that consumers were willing to pay a median price of $14,200 (US) for an EV, far below the projected $25,000 for GMs two-seat Impact.[42] Concerns over

---

[38] CQ Researcher, July 1993, p.581.

[39] Nadis and MacKenzie, 1993, p.75.

[40] Schiffer, 1994, p.187.

[41] Laurie Michaelis, "The Abatement of Air Pollution from Motor Vehicles," Journal of Transport Economics and Policy January 1995, p.81. See also, Freund and Martin, 1993, pp.148, 175.

[42] CQ Researcher, July 1993, p.583.

the necessity of `home-recharging' has also thwarted consumer acceptance of the EV.[43] Regardless of the origin of the misperceptions and concerns, it is evident that consumer resistance has helped thwart the diffusion of EV technology.

Factors with Mixed Evidence

All of the factors discussed above have been decisive obstacles to the diffusion of EV technology. The literature on environmental technological trajectories also contains some factors in which the evidence is less clear-cut.

Mode of Regulation

North America's rightward shift in politics during the 1980s and 1990s reignited a long-standing debate pertaining to regulations and markets. This debate has manifested itself in the realm of environmental regulation, pitting those in favour of the traditional command-and-control (CAC) approach to regulation against those who prefer the harnessing of market forces.

Analysts focusing on environmental technologies, have joined the fray. A reliance on CAC regulation has been pointed to as a prime impediment to technological innovation. Critics argue that a CAC approach, with its rigidities and emphasis on `best-available-technology', stymies the incentive for producers to reduce pollution beyond the required level, and impedes both the production and diffusion of innovative pollution-reducing techniques.

The evidence concerning EVs is mixed. Sperling (1995) contends that US regulations for automobiles are "fundamentally flawed", because they focus only on `downstream' tailpipe emissions, whereas the emissions from EVs will largely be `upstream' - from electricity-producing plants. Moreover, Sperling emphasizes the automakers' fear that any technological innovation will be used against them: "that those specific improvements will be either explicitly required by regulators or used as justification for further tightening of standards."[44]

Although the evidence suggests that a reliance on the CAC approach to regulation hindered EV technology, a great deal of this evidence is conjectured. What is not speculative, is that the 1990 California zero emission vehicle (ZEV) mandate, a mandate drawn up within a CAC framework, has been the catalyst for the diffusion of EV technology: "Thanks largely to California's pending zero emissions deadline, which has been adopted by several other states, the nation's EV population is due to explode by the end of the decade."[45] Even Sperling, a

---

[43] CQ Researcher, July 1993, p.585.

[44] Sperling, 1995, pp.118-119.

[45] CQ Researcher, July 1993, p.592.

proponent of regulatory reform, recognizes the power of the CAC approach in the case of EVs:

> "No imaginable set of incentives and subsidies could ever have achieved in such a short time what the ZEV mandate has achieved...Almost entirely because of the mandate, every major automaker in the world has now invested in electric vehicle development."[46]

Market Structure

Another factor that can have an important effect upon the diffusion of an environmental technology, is market structure. Kemp and Soete (1990), profess that "the diffusion of pollution abatement technology depends on the market structure of the `polluting' sector."[47]

The evidence regarding EVs is mixed. The market structure of the automobile industry has evolved significantly since Jacoby and Steinbruner remarked in 1973 that the classic oligopolistic nature of the auto industry facilitates "tacit arrangements" between firms to limit competition.[48] Limited competition between World War Two and the mid-1970s, discouraged automakers from taking the lead in introducing major technological or design changes. With the emergence of Japanese companies in the late 1970s, however, competition intensified. General Motor's ambitious `Impact' EV program in the early 1990s reflects this competition, and indicates an elevated willingness on the part of automakers "to innovate and take risks in pursuing promising new technologies."[49]

While competition in the auto sector has escalated in the last two decades, there is reason to suspect that the market structure has remained aligned in such a fashion that has impeded EV technology. An attestation to this assertion is the inability and reluctance of small EV companies to expand their market share. Solar Electric Engineering, Sebring Auto-Cycle, and Solar Car Corporation, for example, are EV companies that have struggled to survive under the shadows of the major

---

[46] Sperling, 1995, pp.135, 140.

[47] Kemp and Soete, "Inside the `Green Box': on the economics of technological change and the environment," in New Explorations in the Economics of Technological Change 1990, p.253.

[48] Jacoby and Steinbruner, 1973, p.53.

[49] Sperling, 1995, p.133.

automakers. These companies and others like them, will be "shoved further into the background as the leviathans of the auto industry belatedly introduce EVs, later in the 1990s."[50] Rather than challenge the major automakers, these companies have to satisfy themselves with co-operating with the automotive behemoths. A vice-president from the EV company U.S. Elictricar, summed up this sentiment: "When you're dancing with the giants, you do anything wrong and they'll crush your toes."[51]

Clusters of Technological Trajectories

Kemp and Soete (1992) note that economic growth is likely to be characterized by clusters of economically interrelated technological trajectories. These technologies reinforce one another, and facilitate `locked-in' development.

The evidence concerning EVs is once again mixed. The ICEV has evolved hand-in-hand with moderately priced and adequately abundant oil-based energy.[52] Nevertheless, the fuel that would energize EVs - electricity, has also been relatively inexpensive and in abundant supply for nearly three-quarters of a century. Furthermore, evidence from independent researchers as well as private and public institutions suggest that there is no need to build new electricity capacity to support large numbers of EVs.[53]

While the availability of electricity may not have been an impediment to the diffusion of EV technology, the infrastructural requirements for this energy source have provided a barrier. Although Sperling (1995) maintains that EVs would not require a new fuel distribution network, others disagree. Auto industry analyst Maryanne Keller, for example, asserts: "For electric power to become anything more than an automotive oddity, we will need convenient recharging stations along the road...but we don't have that infrastructure at this point."[54]

---

[50] Schiffer, 1994, p.178.

[51] BusinessWeek, May 30, 1994, p.114.

[52] The exception is of course the oil shortages during the 1970s. Not surprisingly, alternative fuels gained cogency during this era.

[53] Sperling, 1995, pp.58, 64; see also CQ Researcher, July 1993, pp.584-585.

[54] CQ Researcher, July 1993, p.585.

Conclusions and Policy Implications

The findings suggest that no single answer can fully explain the failure for the dispersion of EV technology. The `conspiracy theory' does not account for many factors, such as the costs to industry, and consumer resistance. Similarly, the `consumerist theory' neglects the impact of appropriability conditions. The `technological constraint' theory probably offers the single most robust explanation; but it too neglects important factors such as the mode of regulation. Policy-makers focusing on EV technology, must recognize the multitude of impediments in order to deal effectively with the issue.

The material contained in this paper, is part of a larger debate pertaining to the role of technology and the environment, and has implications that extend beyond the case of the EV. This debate has its roots in an historical dialectic between `pessimists' - `founded' by the reverend Thomas Malthus, and `optimists' - inspired by classical economists such as Adam Smith. The contemporary polemic pits Neo-Malthusians such as Elrich (1968, 1990) and Meadows et. al. (1972), against the neo-classical economists such as Cole et al. (1973), Simon (1981), and Rosenberg (1994).

As a central point of contention between these two sides revolves around the role of technology, the debate has manifested itself in literature regarding the political economy of technological change. The rift between the positions of Heilbroner (1974) and Onuf (1984), and that of Freeman (1992), is illustrative of how the debate has manifested itself in this realm. While the latter two authors have a pessimistic/neo-Malthusian outlook, a central point in Freeman's thesis is that changes in technoeconomic paradigms, which are a major feature of Schumpeterian long waves, possess the capacity for positively altering the relationship between humans and the environment.

The findings of this paper sound a note of caution to the optimist camp. The failure of the EV innovation to become widely dispersed suggests that technological trajectories are not "an autonomous force with their own internal trajectory" (kemp and Soete, 1992), and `green' technoeconomic paradigms do not necessarily accompany the crest of a long wave. EV technology has been existent for over a century, consistently failing to solidify its position in society, despite numerous opportunities throughout the last century. Concern over the environment, the most recent and perhaps most pronounced opportunity for the diffusion of EV technology, has once again met with great resistance, California's ZEV mandate notwithstanding. The failure for the diffusion of EV technology indicates that policy-makers, interested in promoting `green' technologies, will have to focus their attention on the larger socio-political framework in which environmental technological trajectories are nurtured.

References

Anonymous:

Eclectic Vehicles: Technology, Performance and Potential, International Energy Agency, 1993.

"Let the Market Fight Auto Pollution," editorial, *BusinessWeek*, May 30, 1994.

"The Electric Car's Achilles Axle," *The Economist*, September 19, 1992.

Bates, Bradford, *Electric Vehicles: A Decade of Transition*, Society of Automotive Engineers, 1992.

Freeman, Christopher, *The Economics of Hope*, Pinter Publishers, London and New York, 1992.

Freund, Peter, and George Martin, *The Ecology of the Automobile*, Black Rose Books Ltd., Montreal, 1993.

Gordon, Deborah, *Steering a New Course: Transportation, Energy, and the Environment*, Union of Concerned Scientists, Island Press, Washington, 1991.

Hazelton, Lesley, "Really Cool Cars," *The New York Times Magazine*, March 29, 1992.

Jacoby, Henry D., and John D. Steinbruner, *Clearing the Air: Federal Policy on Automotive Emissions Control*, Ballinger Publishing Company, Cambridge, Mass., 1973.

Kemp, Rene, and Luc Soete, "The Greening of Technological Progress: An Evolutionary Perspective," *Futures*, 24:5, June 1992.

Kemp, Rene, and Luc Soete, "Inside the `green box': on the economics of technological change and the environment," *New Explorations in the Economics of Technological Change*, Pinter Publishers, London and New York, 1990.

Michaelis, Laurie, "The Abatement of Air Pollution from Motor Vehicles," *Journal of Transport Economics and Policy*, XXIX:1, January 1995.

Daniel Schwartz

Nadis, Steve, and James J. MacKenzie, Car Trouble, World Resources Institute, Beacon Press, Boston, 1993.

OECD, The Technology/Economy Programme, Technology and the Economy: The Key Relationships, chapter 9, 19??.

Perrin, Noel, Life with an Electric Car, Sierra Club Books, San Francisco, 1994.

Schiffer, Michael Brian, Taking Charge: The Electric Automobile in America, Smithsonian Institution Press, Washington and London, 1994.

Shacket, Sheldon R., The Complete Book of Electric Vehicles, Domus Books, 1981.

Sperling, Daniel, Future Drive: Electric Vehicles and Sustainable Transportation, Island Press, Washington DC, 1995.

Sperling, Daniel, "Electric Vehicles: Performance, Life-Cycle Costs, Emissions, and Recharging Requirements," Transportation Research 23A:3, May 1989.

Stix, Gary, "Electric Car Pool," Scientific American, May 1992.

Wakefield, Ernest H., History of the Electric Automobile, Society of Automotive Engineers, 1994.

Wang, Quanlu, Mark A. Deluchi and Daniel Sperling, "Emission Impacts of Electric Vehicles," Journal of the Air & Waste Management Association 40:9, September 1990.

Woodruff, David, et. al., "Electric Cars: Will They Work? Who Will Buy Them?," BusinessWeek, May 30, 1994.

Worsnop, Richard L., "Electric Cars," CQ Researcher (formerly Editorial Research Reports), July 9, 1993.

Wright, J. Patrick, On A Clear Day You Can See General Motors, Sidgwick & Jackson, London, 1979.

Daniel Schwartz

Snake: A New Information Fusion Approach To Augment GPS Using Map-Derived Information[1]

Chuang Tao
Dept. of Geomatics Engineering, The University of Calgary

BIOGRAPHY

Chuang Tao is a Ph.D. candidate at the Department of Geomatics Engineering, The University of Calgary. He is experienced with the integration of GIS, GPS, photogrammetry and remote sensing technologies. His current research emphasis is placed on the development of multi-sensor (GPS/INS/Video/Laser) integrated digital mapping and inventory systems. He is a recipient of a number of international awards, such as the Robert E. Altenhofen Memorial Award by American Society for Photogrammetry and Remote Sensing (ASPRS) and the ISPRS Best Young Author Prize by the International Society for Photogrammetry and Remote Sensing (ISPRS), and the SPIE Nikon International Scholarship by the International Society for Optical Engineering.

ABSTRACT

A physically-based information fusion approach which greatly differs from the commonly used probabilistic approach is introduced. In this approach, a road is represented as a deformable model composed of an abstract elastic material. The combination of GPS measurements with map-derived information is accomplished by using a physically-based least action principle. The optimal information fusion can be achieved by solving energy-minimizing equations of the motion. Tests on simulated data show that this approach opened a new avenue for integrating map-derived information to augment GPS for vehicle navigation. The feasibility, effectiveness, and distinctive features of this new approach are demonstrated in this paper.

---

[1] Jim Davey Award for Best Doctoral Papers'96 (second place prize) by CTRF.

INTRODUCTION

The combination of navigation technologies and computer technologies has advanced the development of Intelligent Vehicle Highway Systems (IVHS) in recent years. Positioning and locating are the two distinct requirements in land-based vehicle navigation. Positioning involves determining the positions of vehicles in a coordinate framework (e.g., geodetic or Cartesian system, etc.), while locating provides a vehicle's location relative to land features (e.g., roads or intersections, etc.) [8]. The primary objective of land-based vehicle navigation systems is to combine navigational information with route knowledge to provide drivers and/or navigators with real time positions and locations of the vehicles, and to assist dispatchers manage traffic flow and congestion. Map information plays an important role in these applications, and currently all techniques for vehicle locating and route planning make use of maps.

Currently, a total of 280 land navigation systems have been identified in an IVHS navigation system database [9]. Seventy percent of them have the capability of using GPS as a positioning source. Since stand-alone GPS is not adequate for continuous navigation, many technologies, such as DGPS, Hybrid-GPS, GPS/Dead-reckoning, GPS/Map, have been developed to augment GPS in terms of accuracy, reliability and coverage [7,10,12,13]. Among them, map-aiding has been recognized as an efficient and low-cost means for augmenting GPS navigation [2]. The distinct advantage of map aiding is that no additional hardware is necessary compared to DGPS or dead-reckoning. This feature greatly facilitates the implementation of a portable navigation system [1]. Moreover, digital map databases are already available in most land-based navigation systems. For these reasons, the use of map-derived information in navigation systems is increasing.

The major approaches to map aiding can be classified into three categories: (1) Displaying and monitoring: the map matching technique is applied to snap the computed vehicle coordinates (derived from navigation sensor data) onto the road network in order to show the driver or dispatcher the current road of travel; (2) Drift correction: the accurate map matching is required to extract the accurate positional information of roads (e.g., intersections) to correct drift errors arising from navigation sensors; and (3) Integration of road information: the further integration of road information with other sensor data is a trend to the utilization of map-derived information. The Kalman filtering based integration and map-

aided optimal estimation approaches have been developed [1,10,13,14]. However, existing approaches appear to be incapable of utilizing the map-derived information efficiently with other information sources (such as GPS data). It is considered that a mathematical framework of integrating GPS measurements with map-derived information is still lacking. The exploitation of map information and, the optimal combination of different information sources has not been achieved yet. This leads us to explore a new theory on representing and integrating different information sources, so as to further improve map aided GPS navigation.

This paper presents a novel approach addressing the combination of GPS measurements with map-derived information. Following the introduction, the problem statement and a detailed analysis of previous approaches to the utilization of map-derived information are given. A new information fusion theory and its mathematical formulation are introduced. The implementation of a practical algorithm, *Snake*, for map aided GPS navigation is described. An experimental evaluation of the new approach is also presented.

PROBLEM STATEMENT AND ANALYSIS

The key to reliable vehicle navigation is the optimal fusion of information from a variety of sources. In land-based GPS navigation systems, three information sources are included: vehicle dynamics, GPS measurements, and map-derived information. In the previous approaches to the utilization of map-derived information, two processes are generally involved: (1) Road identification: determining the road on which the vehicle is traveling; and (2) Road information fusion: integrating the road information with GPS measurements.

Road Identification

The nearest point (NP) searching algorithm is intuitively used to perform road identification. It outputs a nearest point to the GPS computed point through searching the road network. If a wrong road is chosen, the performance of the information fusion may degrade dramatically. This is a reoccurring problem in current map aided GPS systems.

Recent considerations are focusing on searching a road segment rather than a road point (such as a NP). The method uses the results of a Kalman

filter to determine a vehicle trajectory and to match this trajectory with a road segment in a road network [6]. To make use of this method, an accurate navigation result is required. The paradox is that the road information should be used at first so that an accurate navigation result could be produced. Solutions to this paradox are still being sought.

Road Information Fusion

Kalman filtering based information fusion methodology has been predominantly used in vehicle navigation. It has been proven that the Kalman filter has great capability of integrating multiple sensor measurements with the vehicle dynamics via a sequential solution. The understanding of the statistics of multiple information sources govern the performance of a Kalman filter. In [1], the map-derived information is treated as an observable and introduced directly into a Kalman filter with the other observables. The weighted point and slope-intercept constrained observation equations are incorporated into the Kalman filter for the navigation solution.

Another approach to augment GPS using map-derived information is given in [14]. The approach is based on the premise that the correct roads of travel have been identified. Due to the various noise sources affecting GPS measurements, the measured positions of the vehicle would not lie on the road. The approach takes into account the measurement noise statistics of GPS and optimally translates the raw position measurements onto the road network. A maximum a posterior (MAP) estimator is proposed to determine the translation using the spatial correlation of the measurement errors. This is an alternative approach of combining the GPS measurements and map-derived information. The advantage of the approach is that the NP searching is avoided. The performance of this approach depends heavily on the statistic model of GPS measurement errors. Temporally correlated GPS errors and biases introduced by Selective Availability (SA) may violate an implicit assumption of the MAP estimator [13].

The common characteristics of the above approaches can be summarized as: (1) Both the map-derived information and GPS measurements are investigated by means of the probabilistic approach and combined using statistical criteria. (2) Only discrete road points are used and their positional information is treated as constraints in the estimation model. (3) Road identification is critical. Incorrect road identification would greatly

affect the results.

Problematic Domain Analysis

Absolute positioning systems, GPS, generally provide an accuracy of 30-100 m in a stand-along mode. Due to the various noise sources, the curve of vehicle positions determined by GPS alone is not smooth, and even vibrated when the vehicle is traveling in tree canopy or urban canyon areas. The spikes on the trajectory are caused by satellite signal blockage or the constellation changes. The multipath error is about 150 m commonly in urban canyon areas [1]. The above errors are not able to be modeled and corrected successfully in the GPS stand-alone mode. This fact implies that the GPS measurements are not a stable information source compared to maps.

Since the vehicle is always traveling on roads, map information seems to be reliable and stable to the vehicle navigation. The absolute accuracy in maps is smooth (usually less than 30m [8]), and the internal relative accuracy is very high. This is an important structure feature of maps. In the previous approaches, only the positional information of road points is used rather than the road segments or the road network. In fact, a road should not be decomposed into independent points. It should be treated as a whole body in order to consider its structure feature. In this sense, the optimized combination of GPS measurements with road information has not been achieved by using the previous approaches, in particular, the structure feature of maps has not been applied efficiently. This is a motivation for this research.

PHYSICALLY-BASED INFORMATION FUSION

The physically-based information fusion approaches greatly differ from the commonly used probabilistic approaches. The approach is fundamentally dynamic and governed by a least action principle (Hamilton's motion principle). The Lagrangian equations of motion form the basis of the mathematical framework of the approach [4,17]. Recently, the physically-based approach has been of increasing interest in image information processing [11,15,16]. The approach is applied to infer and reconstruct objects of interest from images. Due to the existence of large ambiguities of the object information in images, the physically-based approach shows great advantages on combining the geometric, physical

and dynamic constraints. It is essentially a novel information fusion approach.

Conceptual Considerations

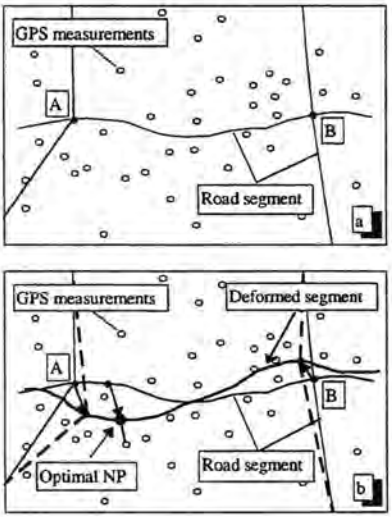


Figure 1. The physically-based combination of GPS measurements with map-derived information.

As can be seen in Figure 1a, from a geometry point of view, information fusion of GPS measurements and road information can be treated as a combination of the linear feature information and the point feature information. If using a probabilistic approach, both lines and points must be modeled as random objects associated with a probabilistic distribution. If we resort to a physically-based approach, we may treat this case as a force field in which the forces coming from different sources are acting upon each other. A curve composed of an abstract elastic material is placed in the force field. Each point of this force field acts a force onto the curve and deforms the curve (Figure 1a). The curve has its internal energy determined by the composed elastic material. This internal energy maintains the natural shape of the curve and resists the external deformation forces. According to Hamilton's least action principle [4], the motion of an elastic body, under the influence of shape constraints and deformation forces, follows a path from an initial state to a final state that minimizes the action energy. It shows that the deformation of the curve under the action of internal and external forces indicates an optimal energy combination. Figure 1b illustrates a deformed curve after an energy combination.

The combination of GPS measurements and road information can be accommodated by using this physically-based energy action theory. The road represents the curve while the GPS measurements represent the

points. Since a road segment in maps has a high relative accuracy and a smooth absolute accuracy, it nicely fits to an elastic body associated with an internal energy. Thus a road segment becomes a whole body rather than the separated points. The advantage of this road modeling is that the structure feature of road information can be represented. The combination of GPS measurements with road information can be modeled as a deformation action, and such an action can be formulated by using the characteristics of different information sources. After an energy action process (information fusion), a deformed road model is obtained. This final road model reaches a state in which an optimal combination of the internal and external energies is accommodated. At this state, the nearest point to the GPS computed point can be considered as an optimal position of the vehicle, because this NP is determined by a dynamic synthesis of the road segment information and the surrounding GPS measurements. As illustrated in Figure 1b, the optimal NP of the GPS computed point *P*

Deformable Road Model (Snake)

A road network is composed of nodes and segments. Each segment is connected by two nodes. The physical meaning of a node is an intersection in maps. We treat each road segment as a deformable curve, termed a *snake*, composed of an abstract elastic material. In Figure 1, the road segment connected by nodes A and B are defined as a snake(A-B). The connecting relationship between the segments and the nodes are always maintained. Therefore, no matter how big the deformation of a road segment is, the topological relationship in the road network can never be destroyed (see Figure 1b). The parametric representation of a snake can be described as $v(u) = (x(u), y(u), z(u)), (u \in [0,1])$. Its total energy can be written as:

$$E = \int_0^1 [e_1 E_{int}(v(u)) + e_2 E_{ext}(v(u))] du \qquad (1)$$

where E_{int} is the internal energy of a road segment, and E_{ext} is the external energy. e_1 and e_2 are constants for weighting these two energies. During the minimization of the total energy E, the snake is deformed to reach an optimal compromise between the constraints introduced by the internal and external forces.

Internal Energy of the Snake

The internal energy represents the structure of roads. It maintains the local continuity and connectivity of the snake and controls the snake from deforming from its natural state. E_{int} can be described by a spline energy [17]:

$$E_{int}(v(u)) = \alpha(u)|v'(u)|^2 + \beta(u)|v''(u)|^2 \qquad (2)$$

The first term is a measure of the distance discontinuities (stretching effectiveness), the second term is a measure of the orientation discontinuities (bending effectiveness). This spline energy imposes a smoothness force onto the snake. Adjusting the weights α and β controls the relative balance of stretching and bending forces. The structure features of the roads (e.g., absolute/relative map accuracy and road shape) are used to determine the weights α and β , and to partially control the magnitude of the deformation.

External Energy of the Snake

Another factor affecting the snake deformation is the external energy defined by the GPS measurements and the correlation to the roads. The GPS measurements are the estimates obtained by Kalman filtering, DOP fix, least squares or other methods [1]. In order to quantify this energy, a gravity-type force field is introduced. The reason is that the closer the distance between the GPS measurement point and the road, the greater the force. Shown in Figure 2, a gravity-type spring is imposed between the measurements and the road. To avoid the singularity when the distance is approaching zero, we employ the following function in defining the external energy:

$$E_{ext} = f(D_p / r), \qquad f(x) = \begin{cases} x^2, & x < 1 \\ x^{-1}, & x \geq 1 \end{cases} \qquad (3)$$

where D_p represents the distance between the measurement P and the curve points. The coefficient r controls a singularity of the action. The corresponding external force can be derived :

$$F_{ext} = \nabla E_{ext} = f'(D_p / r) / r \qquad (4)$$

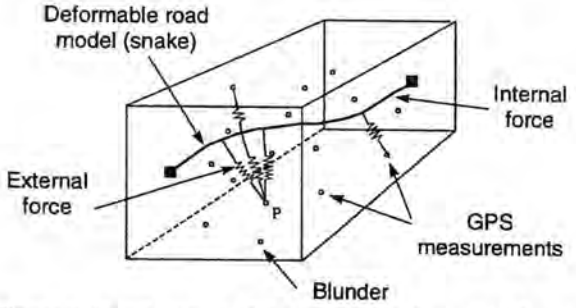


Figure 2. Combination of the internal and external forces

It should be noted that the accurate error model of GPS measurements and its spatial correlation model with the roads can also be used to define the external energy. The reason of the use of a gravity-type energy definition is that the above-mentioned models are not available or not accurate to be applied. On the other hand, the current model has a strong robustness. The effects of blunders and points distant from the snake are minimized.

ALGORITHMIC IMPLEMENTATION

B-Spline Based Snake Modeling

The solution of an energy-minimizing function E in Equ. (1) involves a large number of unknowns. In order to implement an efficient algorithm for navigation, 3D cubic B-splines are applied to represent the snake. Thus, the model of the snake (*B-Snake*) becomes :

$$\begin{aligned} Q(u_j) &= (x(u_j), y(u_j), z(u_j)) \\ &= (\sum_{i=0}^{m-1} V_i B_i(u_j)), \quad j=0,1,...,n-1 \end{aligned} \tag{5}$$

where B_i are the basis functions of B-splines, $V_i=(X_i,Y_i,Z_i)$ are a set of control vertices to this B-spline curve, n and m are the numbers of sampling points on the curve and control vertices, respectively.

Snake Model Based Vehicle Navigation

The GPS measurements are updated when the vehicle is traveling on a road segment which has been modeled as a snake. Substituting Equ. (5)

and (2) into Equ. (1), the discrete version of the motion equation is obtained,

$$E = \sum_{j=0}^{n-1} \{e_1[\alpha(u_j)(\sum_{i=0}^{m-1} V_i B_i'(u_j))^2 + \beta(u_j)(\sum_{i=0}^{m-1} V_i B_i''(u_j))^2] + e_2 F_{ext}[Q(u_j)]\} \quad (6)$$

To obtain the minimum of E, the solution of a variation extreme is involved. The final form of the solution ends up with the following matrixes [15]:

$$\begin{aligned} X &= (A+\gamma I)^{-1}(\gamma X - F) \\ Y &= (A+\gamma I)^{-1}(\gamma Y - F) \\ Z &= (A+\gamma I)^{-1}(\gamma Z - F) \end{aligned} \quad (7)$$

where γ is an Euler step size, and A is a banded matrix (m×m). It should be noted that the dimension of the inversion is m instead of n. The computation of the inversion by LU decomposition only needs a linear time O(m).

After solving the motion equations, the deformed snake is determined. And then, an optimal vehicle position at the current time epoch is obtained which is defined by a nearest point on the snake to the latest GPS estimate (see Figure 1b). If a new GPS measurement is added, the balanced force field would be damaged and the new position of the snake is to be determined via solving the updated motion equations.

EXPERIMENTAL EVALUATION

Tests on simulated data have been conducted to evaluate the feasibility and viability of the new approach. A road network database is established in which the roads are spaced at about 100-200 m (Figure 3a). Most urban areas have similar road spacing. The bold line in Figure 3a represents a simulated vehicle trajectory. A number of random points along the vehicle trajectory are generated for simulating GPS measurements. These random points serve as zero-mean uncorrelated Gaussian distributions. The standard deviations of these point sets are randomly varied from 0 to 100 m, the actual accuracy of the stand-alone GPS mode.

In this test, 12 nodes (intersections) are involved. A total of 37 snakes are generated on this travel, e.g., snake(1-2), snake(2-3), snake(1-13), etc.. The average length of involved road segments is 158 m. The relative accuracy of the road network is assumed to be 15 m (the claimed accuracy of Etak maps). The spacing of the control vertices in the snake model is set up as 15 m. Thus the number of unknowns of the motion equations ranges from 7 to 15. Without using the B-spline representation, the number of unknowns would be 35-75 (assuming the spacing of the sampling points along the road is 3 m). In this test, since the accuracy of positions derived from the road database is better than that of GPS, the constants $e_1=0.6$ and $e_2=0.4$ are chosen for balancing the action of the internal and external energies. The weights $\alpha=0.7$ and $\beta=0.5$ are used according to the experiences. The singularity coefficient $r=15$ m is determined in this case.

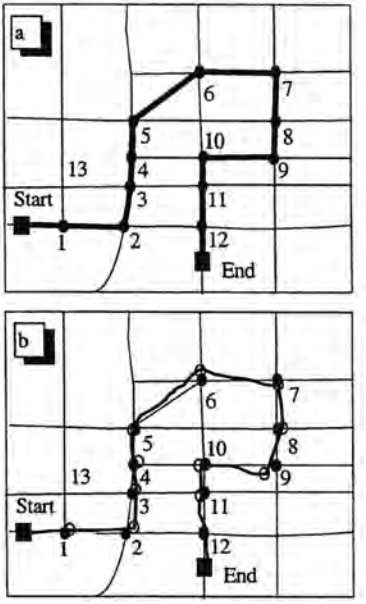


Figure 3 (a) A simulated vehicle trajectory (bold line) and a road network; (b) The final results of the positions (bold line) combining both GPS measurements and road information.

The final result is illustrated in Figure 3b. The marks "o" represent the positions of the deformed intersections. The effectiveness of this physically-based approach is demonstrated. The obtained trajectory is very smooth. The common problem of "Jumping from road to road" is alleviated, because the snake model is always attached to a current road segment. The wrong road identification could be avoided even though the neighboring road is very close to the current road (e.g., the road B is far close to the road A in Figure 3b). In order to evaluate the performance of the intersection identification, we placed five branches at the node 5. As can be seen the correct road was identified using the designed candidate

election algorithm. At node 9, the neighboring nodes are very close and it is usually very easy to snap a wrong road. Since the end point of the snake was deformed together with the whole body, the wrong road will not be snapped when the vehicle makes a turn. In fact, the common problem coming from the use of a NP searching to identify the road is completely avoided in this approach.

CONCLUDING REMARKS AND FUTURE WORK

The physically-based approach for information fusion demonstrates its advantages and distinctive features to augment GPS using map-derived information. The following characteristics of the new approach can be summarized:

(1) The synthesis of multiple constraints can be implemented by a physically-based least action principle. The geometric, physical and dynamic constraints can be unified into the motion equations. The information fusion becomes a physical problem of solving the motion equations.
(2) A road segment is modeled as a deformable model (snake). It allows us to efficiently use the significant structure features of maps rather than the positional information of the independent road points.
(3) The deformation mechanism is endowed with the snake so that the constraints coming from the GPS measurements and map-derived information can be combined through an optimal and robust solution of energy minimization.
(4) The common problems of the wrong road and wrong intersection identification are resolved.
(5) The numeric advantage of using B-splines based snake technique is significant.

Practical tests for this new approach are obviously necessary. This defines our ongoing work. The accommodation of the model of GPS measurement errors and its correlation to roads into the snake energy model needs further research. In fact, the statistics of information sources can be incorporated implicitly into this physically-based approach. However, the further combination of the physically-based approach with the probability approach appears to be important. Thus, a mathematically more rigorous approach can be obtained.

ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr. Ron Li, for his kind support and guidance. Special thanks go to Dr. E. J. Krakiwsky who initiated my interest and encouraged me to pursue this research. His valuable discussions are appreciated.

REFERENCES

[1] Bullock, J.B., 1995. "A Prototype Portable Vehicle Navigation System Utilizing Map Aided GPS." Proceedings of ION GPS-95, pp. 319-329.

[2] Buxton, J.L., S.K. Honey, W.E. Suchowerskyi, and A. Tempelhof, 1991. "The Travelpilot: A second-generation automotive navigation system." IEEE Transaction on Vehicular Technology, Vol. 40, No. 1, pp. 41-44.

[3] Collier, W.C, 1990. "In-vehicle route guidance systems using map matched dead reckoning," Proceedings of IEEE Position, Location and Navigation Symposium (PLAN'90), pp. 359-363.

[4] Courant, R. and D. Hilbert, 1953. Methods of Mathematical Physics, Vol. I, London: Interscience.

[5] French, R.L., 1989. "Map Matching Origins, Approaches and Applications." Proceedings of Land Vehicle Navigation. July 4-7, Koln, Germany, pp. 91-116.

[6] Harris, C.B., 1989. "Prototype For A Land Based Automatic Vehicle Location And Navigation System." M.Sc. thesis, Dept. of Geomatics Engineering, The University of Calgary, Canada.

[7] Krakiwsky, E.J. and J.F. McLellan, 1995. "Making GPS Even Better With Auxiliary Devices." GPS World, November 1994, Vol. 6, No. 3, pp. 46-53.

[8] Krakiwsky, E.J. and J.B. Bullock, 1994. "Digital Road Data: Putting GPS On The Map." GPS World, May, Vol. 5, No. 5, pp. 43-46.

[9] Krakiwsky, E.J., M.A. Abousalem, D.J. Szabo, and J.B. Bullock, 1995. IVHS Navigation System Database, Version 4.5, Dept. of Geomatics Engineering, The University of Calgary, Canada.

[10] Krakiwsky, E.J., C.B. Harris, and R.V.C. Wong, 1988. "A Kalman Filter For Integration of Dead Reckoning, Map Matching, and GPS Positioning." Proceedings of IEEE Position, Location and Navigation Symposium (PLAN'88), pp. 39-46.

[11] Pentland, A. and S. Scaroff, 1991. "Closed-form Solution for Physically-based Shape Modeling and Recognition." IEEE Transaction on Pattern Recognition and Machine Intelligence, Vol. 13, No. 7, pp. 715-729.

[12] Schwarz, K.P., M.E. Cannon, and R.V.C. Wong, 1989. "A Comparison of GPS Kinematic Models for the Determination of Position and Velocity Along a Trajectory." Manuscripta Geodetica, Vol. 14, No. 5, pp. 345-353.

[13] Scott, C.A., 1994. "Improved GPS Positioning for Motor Vehicles Through Map Matching." Proceedings of ION GPS-94, pp. 1391-1400.

[14] Scott, C.A. and C.R. Drane, 1994. "Increased accuracy of motor vehicle position estimation by utilizing map data, vehicle dynamics, and other information sources." Vehicle Navigation Information Systems Conferences (VNIS'94), August, 1994.

[15] Tao, C. 1996, "Road Centerline Reconstruction from Sequential Images Based on Shape From Sequences." SPIE Proceedings on Visual Communication and Image Processing (VCIP'96).

[16] Terzopoulos, D.A. and A. Witkin, 1988. "Physically Based Models with Rigid and Deformbale Components." The Visual Computer, Vol. 4, No.6, pp. 306-331.

[17] Tezopoulous, D. A. Witkin, and M. Kass, 1988. "Constraints on deformable models: Recovering 3D shape and nonrigid motion", Artificial Intelligence, Vol. 36, pp. 91-123.

[18] Yamaguchi, F., 1988. "Curves and Surfaces in Computer Aided Geometric Design." Springer-Verlag, New York.

[1]A Proposed Framework for Optimized Road Maintenance

Curtis F. Berthelot
Ph.D. Student, Texas A&M University

Abstract

Highway agencies are faced with managing the deterioration of road infrastructure assets. This paper presents a road maintenance framework that can be used to assist highway agencies optimize road maintenance expenditures based on the integration of state of the art road engineering science and the sound principles of road asset management. This type of integrated approach will enable highway agencies better meet the specific needs of the agency and the users of the road infrastructure.

1.0 Background

The road infrastructure is a depreciating asset with its remaining service life (i.e. value) written off over time as it deteriorates. However, the value of the road infrastructure may increase if maintenance and rehabilitation is performed to improve the level of service and/or extend the service life of the road. It is the responsibility of highway agencies to determine optimal road maintenance strategies that will provide the maximum value in terms of reduced life cycle maintenance costs and/or improved service to the road users.

Recent years of shrinking highway budgets has made it increasingly difficult for highway agencies to manage the road infrastructure in an optimized manner. Therefore, a management framework that can accurately quantify costs and benefits associated with alternative road maintenance and rehabilitation strategies and identify to whom the costs and benefits accrue, is an essential tool for modern highway agencies.

---

[1] The author would like to acknowledge Saskatchewan Highways and Transportation and the Transportation Association of Canada for the assistance provided for the author's Ph.D. program.

2.0 Define Agency Maintenance Goals and Objectives

The first step to optimize road infrastructure investments is defining clear goals and objectives for maintaining the road infrastructure. Well defined agency goals and objectives provide two things: 1) direction for road maintenance efforts and, 2) the means for which to measure the success or failure of maintenance expenditures. Recent years of shrinking highway budgets has resulted in agency objectives that focus on minimized short term agency costs. Although minimizing short term costs reflect well in deficit reduction efforts, they often result in stop gap maintenance which ultimately increases future maintenance expenditures. Because the road infrastructure is a long term asset, road maintenance goals and objectives should consider the life cycle of the road, thereby minimizing the bias of short term benefits associated with stop gap maintenance strategies.

3.0 Road Performance Models

In order to optimize the design and preservation of roads, road asset management must include road performance models that explicitly account for road type, road condition, traffic, environment, climate, and alternative maintenance treatments. The term "model" refers to a set of information or mathematical relationships that represent some physical aspect of the real world. In the context of road performance, models provide a measure of effectiveness of different maintenance treatments in terms of future road condition.

Road performance prediction models require: 1) definition of critical road performance measures, 2) identification of the variables that determine road performance and, 3) determination of the functional relationships between the critical variables and road performance. There are essentially two approaches that can be employed to predict road performance: an empirical-phenomenological modeling approach, or a mechanistic modeling approach (1).

3.1 Empirical-Phenomenological Road Performance Modeling

In the past, road engineers have relied primarily on empirical-phenomenological road performance models. Figure 1 illustrates the general empirical-phenomenological road performance modeling framework.

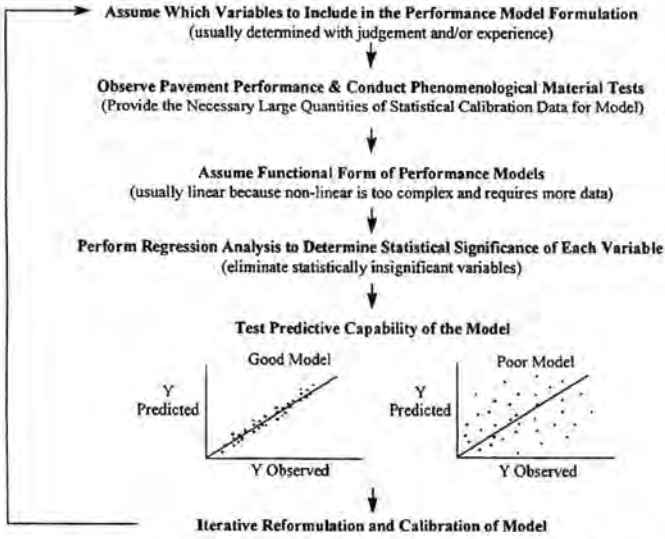


Fig 1 Empirical-Phenomenological Road Performance Modeling

As can be seen in Figure 1, empirical road performance models are heavily reliant on judgment and/or past experience for model formulation and identification of performance variables. As a result, empirical road performance models are limited in their applications and dependability because empirical models offer no sound scientific basis from which the functional form and variables of the performance prediction relationships may be determined (2,3).

3.2 Mechanistic Road Performance Modeling

Road performance is directly related to the mechanical response of the road structure (4). As a result, recent road performance modeling efforts have focused on employing mechanics and materials science to identify the important performance variables and formulate the constitutive behavior of roads based on the natural thermomechanical laws of science. The benefits associated with mechanistic road performance modeling include the following:

- fundamental principles of thermomechanics are universal and facilitate a multidisciplinary approach to modeling complex road behavior;
- thermomechanical laws provide a stable platform for continuous

improvements over time without discarding previous modeling efforts;

- mechanistic road performance modeling provides a direct mapping to road distress predictions based on thermomechanical behavior therefore reducing expensive model calibration and validation;
- thermomechanical material testing is independent of the test apparatus, test procedure, or test condition, and replace phenomenological "test-dependent" material characterization methods that are not sufficient to predict road performance (5) and standardizes all road material testing for research, design, quality control/quality assurance, and forensic analysis purposes;
- mechanistic based road performance models employ the same modeling framework for all road types, under any traffic or environmental conditions.

A general framework for mechanistic based road performance modeling is illustrated in Figure 2.

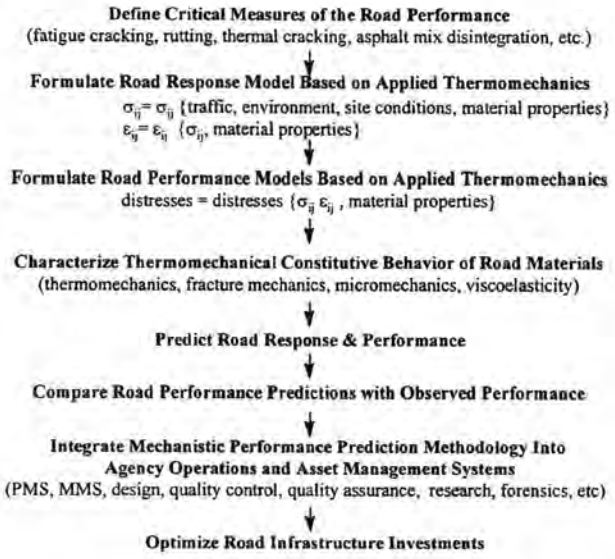


Fig 2 Mechanistic Road Performance Modeling Framework

As can be seen in Figure 2, mechanistic road performance modeling incorporates fundamental thermomechanical material characterization, a mechanistic road response model, and a distress prediction model based

on the primary response of the road structure. As a result, mechanistic based road performance modeling requires a thorough understanding of engineering thermomechanics as applied to material constitutive characterization. Once adopted, mechanics can form the foundation for all agency operations, including: 1) end-product, performance and ultimately warranty based specifications, 2) quality control/quality assurance testing, 3) road materials selection and allocation, 4) road material job specifications, 5) mix design and structural design, and 6) research and forensic investigations (6). Therefore, mechanics is the foundation highway agencies can employ to optimize agency operations and the taxpayer's investment in the road infrastructure.

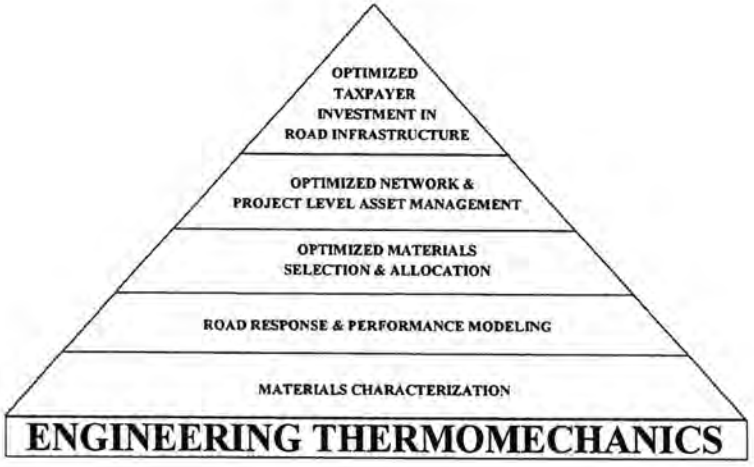


Fig 3 Thermomechanics Foundation for Road Asset Management

3.3 Probabilistic Road Performance Models

Road performance and the variables that affect road performance are inherently stochastic. Probabilistic modeling frameworks offer the advantage of accounting for the inherent variability of individual performance variables. Figure 4 illustrates a deterministic and probabilistic road distress predictions.

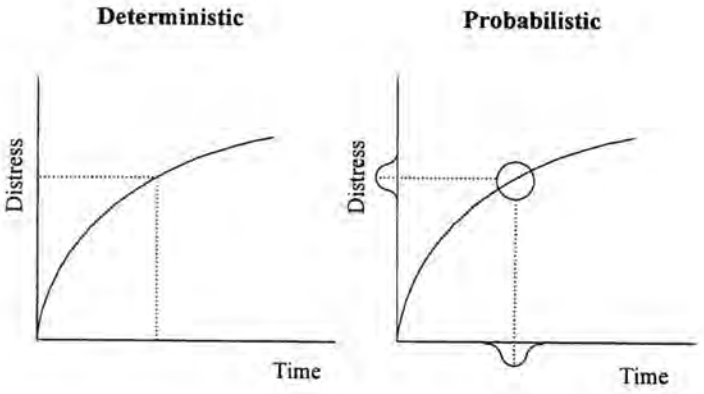


Fig 4 Deterministic vs. Probabilistic Performance Predictions

As can be seen in Figure 4, a deterministic performance models yields one value of performance for a given single value input. A probabilistic model employs distributions of input which is more representative of what is actually known. Therefore probabilistic models yield a distribution of possible performance outcomes. By formulating the inherent variability of the parameters that influence road performance, maintenance managers can quantify the impact variability will have on the road performance predictions. This assists in concentrating data collection efforts on the parameters with high variability in an attempt to reduce random and systematic errors of the performance predictions.

4.0 Road Asset Information

To optimize road infrastructure maintenance, highway agencies must have reliable up-to-date information regarding the inventory and condition of road infrastructure assets in order to establish needs based maintenance budgets. Because road condition information is time sensitive and expensive to collect, road condition data can be detrimental unless it is collected at the right place, at the right time, and used effectively within a road maintenance optimization framework. Too much data can "bog down" the maintenance optimization process and obscure the important and relevant issues of maintaining the road infrastructure. Therefore,

more road condition data does not necessarily mean improved maintenance decisions. Therefore, road asset data collection should only be performed if the information corresponds to the current thresholds of maintenance decisions as dictated by the needs of the road asset management decision process.

Quality road infrastructure information should contain the following:

- be well defined and only collected if the maintenance management framework can employ the data in the decision process;
- be collected in stages to the level of detail and accuracy required by the management decision process;
- be fully auditable;
- be efficiently collected and compiled into a format that can directly support management decisions;
- be cost justified (i.e. maintenance savings that result from changed decisions as a result of the data collected should be greater than the cost of collecting and compiling the data);
- be collected at the correct time taking into account seasonal variations in the road infrastructure, lead time required to collect, process and analyze the data.

Road maintenance managers must implement effective data collection that meets the needs of the maintenance management decision framework. This requires maintenance managers to be aware of the costs, accuracies, and limitations of alternative data collection sources and technologies.

Data Management Systems

Given that road infrastructure condition data needed for scheduling road maintenance is relatively volatile, up to date information must be provided to the asset management decision framework to prevent maintenance strategies that are based on outdated information. Therefore efficient data management systems are commonly employed by highway agencies to assist in efficient road infrastructure asset management decisions. Geographic Information Systems (GIS's) are commonly used to store, compile, and analyze the vast amounts of spatial and non-spatial information typically required by highway agencies. The ability to efficiently update, summarize, and present road infrastructure data makes the GIS an extremely powerful analysis and communications custodial tool for information regarding the highway infrastructure.

Agency Construction and Maintenance Records
Highway agencies often have readily available historic information regarding the original construction and maintenance of road segments (i.e. as built drawings, materials testing, etc.). Maintenance records can provide valuable "first hand expert" information that can be readily used in the maintenance decision process with several benefits:

- road maintenance personnel understand road maintenance activities and are in direct contact with the road network;
- training of road maintenance personnel to identify road distresses would be relatively easy because of their expertise and background;
- experienced road maintenance personnel can provide explanations of road deterioration mechanisms occurring in specific locations;
- involving road maintenance personnel directly in the maintenance planning and management framework promotes a Total Quality Management orientated approach to optimized road maintenance.

To utilize road maintenance crews for road condition assessments, maintenance crews must have an efficient method for recording and downloading data without interfering with routine maintenance duties. Hand held electronic data loggers and GPS location referencing devices (7) can spatially reference and directly download information efficiently into a central GIS database.

Road User Complaints and Accident Records
Because most roads are publicly funded roads, the level of service being provided by the road infrastructure should be considered in the maintenance planning framework. Therefore, road user complaints and accident reports can be a valuable source of feedback regarding the public's perception of the level of service being provided by the agency.

Road Condition Surveys
Road condition surveys are performed to determine the approximate condition and rate of deterioration of the road infrastructure. Network level condition surveys are usually performed at regular intervals (i.e. every 1 to 3 years) depending on the age of the road, and the maintenance operations pending, to provide a cursory differentiation between road segments that require maintenance and the probable type of maintenance required. Previous research (8) determined the percentage of the road surface to be surveyed at the network level as a function of the intended use for the data.

- 3% to determine the general overall condition of the network;

- 10% to determine the general condition within one category of a particular asset;
- 30% for maintenance planning purposes.

Project level data collection is usually only performed on the road segments that: 1) have been identified in the network level survey and post maintenance analysis as being in need of maintenance in the very near term (either preventative or corrective maintenance) and 2) if the optimal maintenance to be performed on individual road segment is debatable and additional information is required to support the maintenance decision. Project level data collection often involves detailed walking inspections of the road and surrounding conditions (i.e. distress surveys, drainage, etc.) and may also employ structural assessment with ground penetrating radar, deflection measurements, or coring to obtain additional road structural information to support road maintenance planning decisions.

Recent advances in semi-automated data collection systems and technologies offer the road maintenance managers with the advantage of collecting road information at high speeds with semi-automated data compiling and downloading to a GIS system. Semi-automated data collection vehicles typically employ non-contact measurement devices such as lasers and ultrasonic sensors to simultaneously collect a wide range of pavement measurements including: transverse profile, rut depths, longitudinal roughness, surface texture, concrete pavement faulting, crack identification, etc. Video and film logging of the right of way and pavement surfaces has been widely used to provide pavement and road side appurtenance condition. Video logs may be integrated directly into storage systems for direct retrieval from the data base and as digital storage capacities are improved, entire video logs may be able to be stored directly into a GIS data base. Complimentary post processing systems are also being refined to provide on screen surveying for inventory data collection or automated distress identification and evaluation directly from the video logs. To accurately reference semi-automated data collection, spatial referencing is usually collected with the condition data. Global positioning systems (GPS) employ the theory of triangulation off high altitude orbiting satellites to provide absolute coordinate locations in global 2-D or 3-D space. Typical GPS accuracies from standard global positioning are typically ±30 to 100m lat.-long and ±50 to 200m elevation with respect to the absolute surface grid of the earth. If a fixed base station is used to differentially correct the GPS readings, positioning accuracy can reach sub meter. To compliment GPS, aircraft grade orientation systems employing aircraft grade accelerometers

and gyroscopes may also be employed for improved road measurement and road surveying accuracy.

Structural Integrity

The structural integrity of the road structure is a critical input to road performance modeling. Road structural information can be obtained from coring, ground penetrating radar and deflection measurements. Ground penetrating radar may be used to provide high speed pre-surveillance for deflection measurements and coring by identifying potential weak spots in the road structure that require investigation.

Traffic Data Collection

In the past, traffic data has been collected in terms of equivalent standard axle loads (ESAL's). However, the ESAL is based on overall serviceability index. Mechanistic road performance models that predict the propagation of specific distresses require accurate traffic data collection in terms of specific vehicle types and axle weights because different distresses influence various levels of traffic, therefore, each distress has its own mechanistic based equivalency. The type of equipment used to collect traffic information will depend on the type of road, type of traffic, and the intended use for the data. The four most common weigh in motion technologies and their respective weighing accuracies may be summarized as follows (9):

- piezoelectric sensors: ±15%
- capacitance sensors: ±15%
- bending strain WIM scales: ±10%
- load cell WIM scales: ±5%

Climate and Environmental Data Collection

Climatic and environmental conditions can significantly effect the load carrying capacity and distress propagation of road structures. Because many maintenance treatments are performed to repair or mitigate against environmental damage, climate monitoring is essential to accurately predict the performance of road structures.

5.0 Maintenance Prioritization and Optimization

Road maintenance may be categorized into three distinct groups:

- preventative maintenance: maintenance performed to slow the rate of deterioration before it occurs;

- repair and rehabilitation: maintenance that corrects road deterioration after it has occurred;
- reconstruction: replaces a severely deteriorated road with a new road.

Prioritization of road maintenance refers to the identification and ranking of viable maintenance projects identified at the facility/network level. Prioritization is often used as an initial cut at identifying viable maintenance projects across the network. Road maintenance optimization refers to the decision processes used to identify optimal life cycle maintenance treatment strategies across a selected number of viable projects. The objective of road maintenance optimization analysis is to identify the optimal combination of the following:

- candidate road segments to receive maintenance treatments;
- maintenance treatments to be performed;
- timing of maintenance treatments.

Several optimization techniques have been employed in the road industry (10), linear and nonlinear programming, integer programming, survivor curves, and the Markovian and semi-Markovian optimization techniques. Of these optimization techniques, the Markovian and semi-Markovian optimization methods explicitly recognize the probabilistic nature of road infrastructure deterioration. The Markovian optimization framework establishes a road condition transition matrix to predict the future condition of the road at set intervals (usually 1 year). The Markovian transition matrix assigns a probability to road condition transitions that may occur from time period to time period. As a result, road structures may have several possible condition states, several possible condition transition paths, and several corresponding transition probabilities as a function of climate, traffic, maintenance treatments, etc.

The semi-Markov optimization approach employs the same probabilistic transition theory as the Markovian approach, however, the holding time between condition states is the dependent variable, and the transition probability is applied to the condition holding time (11). If the sequence of consecutive condition states over the life cycle of a road can be predicted, the semi-Markov formulation method only requires one node (expected transition time and time probability distribution per transition). The semi-Markov optimization methodology offers several advantages over the Markovian framework including: reduced computational dimensionally of any life cycle maintenance analysis; logical and convenient formulation framework for planning road maintenance;

definable condition states with time as the dependent variable which better correlates with road maintenance activities; and maintenance activity trigger points can be directly integrated into the analysis framework based on the maintenance manager's logic for planning road maintenance.

6.0 Life Cycle Cost Analysis

Once the optimal maintenance strategies have been identified, road agency and user costs and benefits associated with each maintenance strategy should be calculated and compared in terms of equivalent uniform annual costs and benefits at a given discount rate over different analysis time frames. There is an increasing need for road user costs and benefits to be incorporated into road asset management due to the increase in partnering relationships between highway agencies and the transportation sector. The intent of these partnering relationships is to facilitate increased transportation efficiency through larger and heavier truck configurations while still protecting the road infrastructure. As a result, road user cost models must be employed so as to accurately quantify the costs and benefits of such partnering agreements and to determine to whom the costs and benefits will accrue.

7.0 Feedback and Model Calibration

Once the road maintenance optimization analysis has identified and implemented the optimal maintenance strategy, the specific maintenance treatments and their respective performance should be monitored as part of the routine data collection process and entered into the historic road maintenance data base on a segment by segment basis. The performance information obtained from each road segment can be used to calibrate and update the road performance models and the maintenance condition transition matrices, thereby improving future road maintenance decisions.

8.0 Summary and Conclusions

In Canada, it is estimated that almost half of the pavements are below a minimal level of acceptable service, requiring approximately $11 billion to bring the road infrastructure up to an acceptable standard (12). To do this, highway agencies must adopt modern scientific methods and management techniques for the design, preservation,

and management of roads. The procedures outlined herein outline an integrated approach of fundamental road engineering science and sound road management principles that will assist highway agencies optimize road design and preservation strategies to better meet future road infrastructure performance objectives.

References

1 J. Uzan, D.G. Zollinger, and R.L. Lytton. The Texas Flexible Pavement System (TFPS), Vol. II - Mechanistic/Empirical Model. Research Report No. 455-1, Texas Transportation Institute, Texas A&M University, 1990.

2 Lytton, R.L. Concepts of Pavement Performance Prediction and Modeling. Proceedings, Volume 2, Second North American Conference on Managing Pavements. 1987.

3 Rauhut, J.B., Gendell, D.S. Proposed Development of Pavement Performance Prediction Models for SHRP/LTPP Data. Proceedings Volume 2, Second North American Conference on Managing Pavements. 1987.

4 Lytton, R. L., et al. Clusters of Terms Relevant to Pavement Performance Prediction. SHRP Contract 89-P-020 Unpublished Technical Memorandum. 1992.

5 E.V. Edris, and R.L. Lytton. Dynamic Properties of Subgrade Soils, Including Environmental Effects. TTI Research Report No 164. 1976

6 Lytton. R.L. Mid-America Transportation Centre Distinguished Lecture, Pavement Performance Modeling. 1997.

7 Roads and Traffic Authority of New South Wales, Australia. Asset Control Technology. 1993

8 Mahoney, J.P., Lytton, R.L. Measurements of Pavement Performance Using Statistical Sampling Techniques. Research Report 207-2, Texas Transportation Institute, Texas A & M University, College Station Texas, 1984.

9 Bergan, A.T., Berthelot. C.F. Effect of Weigh In Motion Accuracy on Weight Enforcement Efficiency. Second International Conference on Intelligent Transportation Systems. 1995.

10 Lytton, R.L. Optimization Techniques. Third International Conference on Managing Pavements. 1994.

11 Nesbitt, D. M., Sparks, G. A. Neudorf, R. D. A semi-Markov Formulation of the Pavement Maintenance Optimization Problem. Canadian Journal of Civil Engineering. 1992.

12 Transportation Association of Canada. Canada's Roadway Infrastructure Selected Facts and Figures. 1990.

An Analysis of Canadian Air Transport Policy

I - Executive Summary - This paper analyzes the Canadian federal government's policy with respect to economic regulation of the airline industry. The paper opens with a short introduction which details why airline policy is important to Canadians. A concise history of Canadian airline regulatory policy is then presented to give the reader an appreciation for how present airline policy has evolved. Present airline policy is then reviewed and critiqued in detail and a set of specific recommendations for change is offered. Each of these later sections are organized by the three major regulatory classifications of domestic market, international market and trans-border market.

The paper argues that the economic structure of the Canadian airline industry has characteristics which lead to industry consolidation and provide the potential for abuse of dominant market position. Accordingly, a review of Canadian competition policy and its impact on airline markets is provided. This competition policy section concludes that Canadian competition policy is ineffective in preventing or mitigating the effects of competitive market failures in the airline industry.

Having established the present regulatory context from both an aviation policy perspective and the more general perspective of basic competition policy, the paper delivers a critique of the existing airline policy package. It is argued that the present pro-competitive package, which favours de-regulation in all markets, fails to define a sustainable role for Canada's major carriers. The policy package strongly favours the short term interests of consumers of transportation over the interests of providers of the service. It is argued that a short term focus on stimulating competition in all market areas weakens the domestic airline industry, hurts the long term interests of Canadian consumers and limits the ability of Canadian carriers to carve out a strong role in the global market for aviation services. It is also argued that certain political and national interests are advanced by maintaining a viable domestic airline industry and that such interests would not be serviced by foreign carriers.

The paper concludes with a set of normative policy recommendations which attempts to establish a more equitable balance between the long term interests of consumers and the long term interests of carriers. A pro-competitive stance is advocated for the domestic and trans-border

markets while elimination of competition between Canadian carriers on international routes is recommended. It is also proposed that a more proactive role for airline regulators be established in order to prevent abuse of dominant market position. In summary, the revisions which I propose recognize that the transportation system exists to service the needs of consumers, but they also recognize that a healthy air travel industry is consistent with the long term interests of consumers.

II - Introduction - Government policy with respect to transportation is of vital importance in a country such as Canada. Canada is one of the largest countries in the world in terms of area, yet we have a relatively small population base which is scattered along the U.S. border from the Atlantic to the Pacific. Our unique geography has always been a major obstacle to the formation of a cohesive economic union. From a purely economic point of view, Canada is an unnatural economic unit and our history is full of examples of government intervention in transportation markets in attempts to offset the effects of geography. The most notable example of government intervention was in the construction of the Canadian Pacific Railway. As Michael Bliss points out, this initiative was not justified on the basis of developing the economic potential of the western region, but it was successful, at least to date, in cementing the nation as a political unit (Bliss, 5).

In more recent times, there has been a shift in Canadian trade patterns from east/west flows to north/south flows. As Courchene points out, this shift has diminished the role of the railway as an economic policy factor binding the country together (1992, 763). The railway's role was further diminished by the emergence of aviation as the dominant transportation mode for the movement of people between regions. Today, aviation links play a crucial role in shrinking the size of Canada and facilitating the continued operation of our federal system of government.

Canada is also a nation which relies heavily on international trade and tourism for its economic well being. The airline industry plays an important role in facilitating trade and developing tourism. In a country the size of Canada, there are no practical alternatives to air travel in such applications. Air transport policy is, therefore, a matter of vital importance to the nation and it should be a high priority issue for all levels of government.

Stakeholders - The Stakeholder groups affected by air transport policy are effectively identified in Transport Canada's international air transport policy document (1994, 2).

> Canadian's look at air transport from variety of perspectives: as a means to travel abroad...for business and leisure purposes; as a means to attract commercial activity to our communities; as a means to import and export high value and perishable commodities; as a means of employment; as a generator of foreign currency and wealth; as a driver of technological change; and as a source of civic and national pride.

The policy document goes on to classify this "Community of interests" into four interest groups - the business sector, travel and tourism, airlines and consumers. It also goes on to identify the formulation of "clear and coherent objectives, strategies, and policies that best meets the overall national interest" as being the appropriate role for government. It is difficult to find issue with the government's assessment of the stakeholder interests or the government's self defined role.

III - Brief History of Canadian Aviation Regulation and Policy

In order to assess current policy, it is useful to have an appreciation for how the industry has evolved to its present status. A useful review of the development of regulatory policy is provided by Christopher in a paper which was issued during the development period in advance of the National Transportation Act of 1986 (1982, 1-17). Supplementary information from other sources has been added to Christopher's review to produce the following summary:

A. The Early Days 1936 - 1942 - Government involvement in aviation in Canada began in 1936 with the formation of Civil Aviation Branch of the Department of Transport. The potential role for aviation in national transportation was recognized by C.D. Howe, the Minister for Transport and he established Trans-Canada Airlines (TCA) as a government airline with a monopoly in all areas of service. By 1942, TCA had established transcontinental service linking all major cities in Canada. The monopoly protection given TCA was instrumental in allowing airline service to expand throughout the infant national market (Christopher, 1982, p 3).

B. Regulatory Protection 1942 - 1959 - Canadian Pacific Airlines (CP) was formed in 1942 from several small bush pilot operations which had been allowed to develop around the periphery of the TCA monopoly. CP pressed for rights to compete with TCA but met with little success. The only rights available to CP were routes not wanted by TCA. For example, in 1948, CP applied to Howe for rights to service Sydney, Hong Kong, and Tokyo. As TCA was the designated international carrier, Howe approached TCA to see if they were interested. TCA's president, Gordon McGregor, is reported to have stated, "Who would want to lose money on those routes? Let him (Grant McConachie, CP's President) lose it all and get out of our hair." (Goldenberg, 1994, 14). Thus was born Canada's "policy" of dividing the world between TCA and CP based on Pacific (CP) and Atlantic routes (TCA). Constant efforts to secure competitive domestic rights finally resulted in CP being granted rights to provide one daily flight in each direction between Vancouver and Montreal (Christopher, 1982, 4).

C. Growth of Competition 1960 - 1980 During this period, CP gradually became a more serious contender to TCA in the domestic market and several regional carriers developed. In 1966, Canadian policy was adjusted to recognize 5 regional markets served by 5 different carriers - Pacific Western Airlines, Transair, Nordair, Quebecair, and Eastern Provincial Airways. Policy was also adjusted to allow international charter operations by companies such as Wardair, but care was taken to ensure that such operations, as well as the operations of the regionals, were kept separate from the national carriers' routes (Goldenberg, 1994, 150). In 1967, the government allowed CP to increase its share of trans-continental service to 25%. By 1980, CP air was allowed to compete head to head with Air Canada[1] in the domestic market.

[1] TCA was renamed Air Canada in 1964

IV - Present Regulatory Policy

A. Domestic Policy - Development of the present regulatory environment can be linked to the introduction of airline deregulation in the United States in 1978. According to Tretheway, the gradual introduction of deregulation in the U.S. between 1978 and the end of 1984 led to the following developments:

1. An increase in the number of airlines.
2. A decrease in passenger concentration.
3. A decrease in the general level of fares.
4. A decrease in average operating expenses.
5. Neutral impact on surviving carrier profits and employment
6. Fewer labour disturbances and a general increase in employee productivity (1986, 299).

These changes were, of course, closely observed in Canada. The new Canadian air policy that emerged in May of 1984, provided for substantial deregulation, including the elimination of the distinct roles of national, regional and local carriers. The policy proposed that any carrier be allowed to operate any service within a defined "liberalized zone" and it proposed the elimination of fare controls. It is interesting to note that the policy of deregulation was opposed not only by the major Canadian carriers, but also by the aviation bureaucracy in the Canadian Transport Commission (Goldenberg, 1994, 50). The transport minister of the day, Lloyd Axworthy, felt that the regulatory body had become too closely aligned with the interests of the industry which it was trying to regulate.[2]

The air policy of 1984 was not enacted due to the fall of the liberal government. The conservative government which followed recognized the popularity of the deregulation issue. As a result, the "Freedom to Move" policy, which was issued in 1985, also promoted deregulation of the domestic industry. The move towards deregulation was put into law

---

[2] Axworthy's comments are consistent with the Capture Theory of Economic Regulation. This theory states that regulation is initially provided to meet public interest objectives but that over time regulatory agencies come to be dominated by the industries they regulate. This theory, as well as a variety of other theories of economic regulation, is effectively detailed and analyzed by Richard Posner (1974, 335-357).

with the passage of the National Transportation Act of 1987. The act contains only 3 basic requirements for entry into domestic markets. The agency shall issue a license if:

(1) the applicant is a Canadian[3]; and
(2) the applicant holds a Canadian aviation document; and
(3) the applicant has the prescribed liability insurance.

In March of 1996, Bill C-14 received first reading. It proposes to revise the NTA of 1987, but the only substantive changes to domestic policy are the addition of a fourth test which would ensure that an applicant meets "prescribed financial requirements" and extension of the zone of de-regulation to include the northern region.

The NTA was built around the principle that the transportation system should exist to serve the needs of shippers and travelers. Accordingly, present domestic airline policy is heavily weighted in favour of consumers of air travel. This was a large departure from previous systems which were designed to promote and advance the development of designated national carriers.

B. International Policy - Present international policy recognizes the fact that international route rights can only be gained through government to government negotiations. Canada secures aviation rights which it then allocates as it sees fit to designated carriers. Historically, TCA/Air Canada was the designated carrier. Subsequently, the government adopted a policy of splitting the world between carriers according to defined regions. In 1994, however, Canada adopted a new policy with respect to international air transport - the "use it or lose it" policy. Substantive changes associated with the new approach were:

1. Elimination of the allocated territories policy.
2. Designation of 37 countries for assignment to new carriers.
3. Designation of secondary carriers on high volume routes.
4. Provision for foreign carriers to serve routes without concurrent Canadian carrier service.
5. Establishment of financial hurdles for new charter entrants.
6. Creation of regulations to eliminate computer reservation

---

[3] Canadian is defined in the NTA as being a minimum of 75% Canadian ownership.

system bias in favour of the host airline.

The new policy extends the trend towards deregulation into the international sphere. The policy document clearly states, "...the time has come to focus on the interests of travelers and shippers and airport communities." Accordingly, competition is promoted between Canadian carriers in the international market. The policy swing in favour of users as opposed to carriers is made apparent throughout the document. In fact, the international air policy is stated to be subservient to the government's earlier policy with respect to transfer of airports to community based organizations (Transport Canada, 1994, 2).

C. Trans-border Canada's international policy does not apply to the U.S. market. This area was covered by a series of bilateral agreements which provided for a total of 83 specific point to point routes. Canadian carriers had sole rights on 26 city pairs, U.S. carriers had sole rights on 38 city pairs, and Canadian and U.S. carriers competed on the remaining 18 city pairs. Under this agreement, Canadian carriers carried 38% of the traffic while U.S. airlines carried 62% of the traffic. Due to the fact that Canada has its population concentrated in a few major cities, U.S. carriers had access to 90% of Canada's population base while Canadian carriers could only access 30% of the U.S. market (Lewis, 1995, 5).

On March 10th, 1995 the government announced the completion of the "Open Skies" agreement with the U.S. The main points covered under this agreement are as follows:

1. Canada gains immediate access to all U.S. points.
2. U.S. carriers gain immediate access to all Canadian points except Toronto, Montreal and Vancouver. Increased U.S. service to these destinations to be phased in over a 3 years.
3. Restrictions on cabotage[4] remain in both countries.
4. 24 new landing slots are provided to Canada at Chicago and New York to improve Canadian carrier access.[5]

---

[4] Cabotage refers to the carriage of passengers by a foreign airline between two points in another country's domestic market.

[5] Landing slots at U.S. airports are governed by the municipal level of government. Local airports sell gate rights to carriers. U.S. federal negotiators had previously taken the position that provision of landing

The new air services agreement governing trans-border routes completes the government's deregulation package. It opens up options for additional routes into the U.S., increases competition from U.S. carriers and it encourages competition between Canadian carriers. Both Canadian and Air Canada will, for example, be competing for passengers on routes into Chicago and New York.

V -Canada's Competition Policy - Its Impact on Airline Markets

As clearly established under section IV, Canada has pursued a policy of deregulation within its airline industry. Removal of airline regulation entrusted the marketplace to competitive discipline and the generic rules of Canada's competition policy. A review of competition policy and the economic characteristics of the airline industry is, therefore, in order.

The economic structure of the airline industry in Canada is not conducive to perfect competition. While economies of scale are reported to be negligible, there are substantial economies to be achieved with respect to route density (Tretheway, 1991, 8). This economic characteristic leads to the formation of hub and spoke route networks as the most efficient structure for the industry. Under such a structure, regional routes consolidate traffic into a central hub and the consolidation of density leads to cost efficiencies in travel between hubs. As a result, marginal increases in traffic generated by the feeder system generate large amounts of profit on the main line route (Madore & Shaw, 1993, 5).

Economies of traffic density create the need for an airline to establish strong feeder networks which have schedules closely coordinated with those of the trunk lines. Additionally, there are strong consumer preferences to fly with a large carrier and to avoid interlining where possible. Such consumer preferences are strengthened through airline loyalty programs such as frequent flier miles. These characteristics combine to produce a strong incentive to achieving network economies through mergers and acquisitions.

In the U.S., several competing hub and spoke structures have developed as different carriers developed hubs in different cities. Additionally, the

---

slots was outside of their jurisdiction and could, therefore not be part of the "open skies" agreement.

density of traffic in certain areas of the U.S. permits viable competition between carriers at common hubs (Windle & Dressner, 1993, 29). In Canada, however, the long thin geographic structure and the concentration of our relatively small population in only a few cities, limits the choices for hub locations; consequently, Canadian carriers compete head to head at the same locations. Given the small size of the Canadian market, economic arguments can be made that the natural consequence of complete deregulation of the airline industry in this context will be the formation of a domestic market monopoly.

As indicated above, the economic characteristics of the airline industry produce strong tendencies towards industry consolidation. These same forces can be assumed to create strong incentives for anti-competitive practices. Airline policy, however, is designed to minimize regulation of the industry. In this situation, competition policy is relied on to keep the industry competing fairly. The evidence would suggest that Canadian competition policy has not been a substantial barrier to industry consolidation to date. The entire history of Canadian Airlines is one of growth through acquisition and merger. Air Canada has also pursued growth in its feeder system through this strategy. Under section 96 of the Competition act of 1986, gains in efficiency are recognized as a legitimate reason to allow a merger if the competition tribunal finds that efficiency gains will offset or be greater than the effects of the lessening of competition. Section 93 (b) also indicates that one of the factors to be considered is whether the party being acquired is likely to fail without the merger. These sections provide a great deal of flexibility within the competition act to allow concentration within the industry.

In the area of anti-competitive practices, the issue of predatory pricing is of particular relevance to the airline industry. Given the large economic rewards to be derived from eliminating the competition and achieving associated increases in route density, there is a strong incentive for a carrier to engage in predatory pricing as a strategy to drive out a weaker competitor. Competition law allows for such practices to be addressed in two ways. Under the civil law section covering abuse of dominant position, the use of fighting brands to discipline or eliminate a competitor is prohibited. Predatory pricing is also dealt with under section 50 of the act which deals with criminal offenses (Brander, 1995, 310).

For practical purposes it is extremely difficult to successfully prosecute a case dealing with predatory pricing in the airline industry. The difficulties associated with proving a case are as follows:

1. Predatory pricing has been defined by the Director of Investigation and Research of the Bureau of Competition Policy as, "a price which is below the average variable costs of the firm" (Odette & Shaw, 1993, 8). In the airline industry, average variable cost (AVC) is not easy to define. If it is defined as direct operating costs exclusive of charges for aircraft ownership or use, there is tremendous scope for pricing flexibility. Additionally, employment of marginal cost pricing strategies can be justified as a legitimate pricing practice in certain airline applications.

2. The issue is further complicated by the industry's adoption of sophisticated yield management systems. As a result, the number of fare classifications has proliferated and there are many different fare classes offered on any given flight. The question of whether a given fare is predatory is sufficiently complicated to lead one industry expert to comment, "...it is impossible to prove ... unless the competitor or anti-trust authority has a written proof (e.g. internal memo) which shows the intention to predate..." (Oum, 1996).

3. The requirement to prove guilt beyond a reasonable doubt eliminates practical use of the criminal provisions of the competition act. As stated above, even the civil law provisions are difficult to prove on a balance of probabilities basis. These two factors combine with the relatively minor costs of conviction to make anti-competitive behavior an attractive business strategy.[6] Even under the new act, fines have rarely exceeded the costs of restitution. The situation is best described as a retroactive license to engage in anti-competitive behavior. Under such an environment there is little incentive to avoid anti-competitive practices such as predatory pricing.

It should be pointed out that the costs of predation can be significant. Game theory in an oligopolistic setting (Dixit & Nallebuff, 1991, 95) predicts that such action would only be profitable for the predator if there

---

[6] Fines have rarely exceeded $1 million for competition policy offenses. Such fines are minor in relation to the $7 billion combined revenue of Air Canada and Canadian Airlines.

is a substantial difference in the economic strength of the competitors. Where competitors are evenly matched predatory pricing would result in mutual destruction. Unfortunately, actions taken by the federal government in the privatization of Air Canada set the stage for predation. In order to make the sale attractive, Air Canada was privatized with a low debt position and a market imbalance was created.

The Gemini case provides an interesting sidelight to the area of competition law and the airline industry (Brander, 1995, 333). As pointed out by Brander, the ruling by the competition tribunal to allow Canadian to exit the Gemini reservations partnership came down to a matter of airline policy. Did Canada want to maintain two domestic carriers or only one? The tribunal ruled in favour of preserving Canadian Airlines but this was a rather backhanded approach to determining government policy on a fundamental issue. The inescapable conclusion is that competition policy is, at best, only able to address such issues on a reactive basis and only when the situation is already in crisis.

VI - Critique of Government Policy

A. Domestic Policy - A critique of Canadian airline policy must begin with the fundamental values underlying the National Transportation Act. In assessing this legislation it is important to recognize that the value which underpins all Canadian transportation policy is that the transportation system exists to service the needs of shippers and travelers. Canada has made the choice to favour consumers of transportation service over the interests of Canadian providers of such services. Given that transportation is largely a derived demand and that Canada is highly dependent on efficient transportation service, this fundamental value makes sense. Canada's industries must have access to efficient and cost effective transportation systems in order to remain viable.

Recognizing the fundamental value, we must now determine to what extent the maintenance of a viable Canadian air transport industry is tied to the interests of Canadian shippers and travelers. Are consumers of travel better off with a strong domestic industry which is focused on servicing the Canadian market, or would we be better served by lower cost foreign based carriers? If Canada were to drop its Canadian content rules tomorrow, it is unlikely that there would be any great rush by foreign carriers to provide service domestically across Canada. Our

market is simply too small to be of much interest. Instead, we would probably see the major Canadian cities incorporated as spokes into major US hubs. Travel between Canadian points and from Canadian points to international destinations would, in many cases, be routed through U.S. hubs.[7] While the cost of service would likely fall, the services provided would not be targeted to the needs of Canadians. Additionally, there would likely be serious political issues to contend with if our air system was totally integrated into the U.S. system. As stated earlier, the transportation system plays a role in binding the country together. If travel between Canadian regions becomes dependent on U.S. airline hubs, it does not bode well for Canadian unity.

If we accept the premise that a strong Canadian airline industry is still in the national interest, the next question becomes, "Is it in our national interest to have more than one domestic trans-continental carrier?" Here there is a trade-off to be made between the efficiencies to be gained from traffic density and the efficiencies to be gained from carrier competition. A monopoly structure would likely be more efficient from a carrier perspective, but it would be difficult to sell to the Canadian public. The prospect of a "Mapleflot"[8] holds little interest for Canadians due to the potential for such a carrier to abuse its dominant position. Reverting back to the fundamental value, such a trade off decision should be made in the interests of travelers. Thus it appears that the preferred approach is to maintain competition between two balanced[9] domestic carriers. The conclusion that Canada should maintain two major domestic carriers appears to be unstated government policy. Unfortunately, the requirement for balance does not seem to be recognized.

The growth of Canadian Airlines to become a legitimate competitor to Air Canada, was achieved through acquisitions (Barone et al, 1986, 421). The expansion was required in order to achieve needed network efficiencies, but it left Canadian in a weakened financial position. In

---

[7] This type of structure is referred to as "Hub Cabotage"

[8] The term Mapleflot was coined by Canadian Airlines employees who were opposed to the proposed merger of Canadian and Air Canada. It is intended as a derogatory term comparing the merged Canadian airline to Aeroflot the former Soviet Union's national carrier.

[9] As demonstrated in Section V balance is required for the competitive system to be sustainable as imbalance creates incentives for predation.

1991 the effects of the Gulf War, the bombing of a Pan Am flight over Lockerby, Scotland, and a worldwide recession all contributed to poor performance within the airline industry. Because of its highly leveraged position, Canadian was particularly hard hit. Losses of $162 million in 1991, $543 million in 1992 and $292 million in 1993 drove the carrier to the brink of bankruptcy. Air Canada fared little better recording losses of $318 million in 1991 and $197 million in 1992 before breaking even in 1993. Collectively the two Canadian carriers lost more than $1.5 billion in just three years.

While airline losses were incurred worldwide during this period, losses by Canada's carriers represented 10% of the total industry loss. Given that travel to, from and within Canada on both domestic and international carriers only represents 2% of the world market (Madore & Shaw, 1993, 3) the Canadian share of the loss is disproportionately high. By 1993, Air Canada's shareholders had lost 47% of their investment's net worth and Canadian's shareholders had lost 95% (Madore & Shaw, 1993,13-15). An investment analyst likened the battle between Canadian and Air Canada to, "...two people jumping from an office tower - PWA[10] from the 15th floor and Air Canada from the thirtieth, thumbing its nose because PWA would hit the ground first." (Goldenberg, 1994, 81).

There is strong circumstantial evidence that Air Canada attempted to drive Canadian Airlines out of existence during this period (Madore & Shaw, 1993, 9). Having failed twice in 1991 in its effort to merge with Canadian or acquire its international routes, Air Canada began an aggressive campaign of discounting fares and adding capacity. It also arranged to increase its credit by the sum of $400 million which was, coincidentally, equivalent to Canadian's cash and accounts receivable. Internal correspondence from Air Canada refers to a "scorched earth" policy with respect to the Gemini negotiations and it is clear that Air Canada continued to express interest in a merger, "...because as long as we profess interest, it helps prevent an AMR[11]/PWA deal." (Goldenberg, 1994, 127).

---

[10] PWA Corp was the holding company which owned Canadian Airlines International Ltd. In 1994 the name of the company was officially changed to Canadian Airlines as part of the corporate restructuring.

[11] AMR Corp is the parent company to American Airlines.

It has been argued that PWA/Canadian's management was negligent in allowing their liquidity position to deteriorate to the point where they became vulnerable to such tactics (Madore & Shaw, 1993, 24). While there is some merit to the claim that PWA's management was negligent, the argument indicates that companies should be responsible for their own protection when it comes to anti-competitive behavior. It also fails to recognize the fact that PWA had to extend itself in order to form a network which was capable of supporting competition with Air Canada and it ignores the government's role in contributing to the imbalance through the favourable terms associated with the privatization of Air Canada.

If one accepts Madore's position that PWA was the author of its own misfortune, PWA's shareholders paid a large price for their management's mistakes. Additionally, the sale of a 30% equity interest in the airline to AMR Corp. was made on very unfavourable terms. While AMR Corp. did not gain full control of the airline it did gain an equal voice with respect to such matters as capital expenditures, sale of assets and acquisitions (Smith, 1993, 4). The shape of the domestic industry was changed dramatically, but it was done out of a private company's desperation rather than through conscious strategic policy on the part of industry or government.

Throughout much of the course of these events the government chose to sit on the sidelines. When they did get involved their position was often inconsistent. For example, in July of 1992 the government stated that it would purchase 3 Airbus 310 aircraft from PWA for $150 million if PWA agreed to break off talks with AMR and resume merger talks with Air Canada (Goldenberg, 1994, 103). At other times the government appeared to favour the AMR alliance, as was demonstrated by the competition tribunal's Gemini decision and the provision of a $50 million interim loan.[12] The result of the lack of decisive action in this crisis was the financial decimation of the domestic industry. It is hard to see how the government's lack of action was in the interests of Canadian travelers. Short term benefits may have been provided in terms of

[12] Part of the inconsistency was related to the change in governing party from Conservative to Liberal in 1993. The liberals took a more proactive role in attempting to settle the dispute but neither government acted in a consistent decisive manner.

reduced airfares, but the cost to the industry reduced its ability to provide future service. When you are focused on keeping enough cash in the bank to keep creditors at bay, innovation and market development take a back seat.

The government's reluctance to act decisively was very costly, however, it is somewhat understandable given the structure of decision making in Canadian government. Aviation policy is a tough issue which presents no easy solutions. Policy decisions are further complicated by the division of the industry along regional lines with Air Canada being based in Quebec and Canadian being based in the west. Bakvis and MacDonald argue that Canada's system of cabinet decision making limits the government's ability to deal with broad issues of national scope. They argue that the decision structure constrains government to a select set of issues deemed to be strategic for the government's agenda (Bakvis and MacDonald, 1993, 75). As aviation policy was not a clearly defined strategic issue for either the liberal or conservative governments at the time of the crisis, each government had difficulty in dealing with the situation.

B. International - The competition tribunal's decision to allow PWA to withdraw from Gemini only removed part of the obstacles to PWA's recovery. PWA was still faced with several Air Canada lawsuits which totaled billions of dollars. Mysteriously, Air Canada withdrew its opposition to the AMR/PWA deal on January 26th, 1994. On January 27th the Canadian government rewarded Air Canada for its new stance with new international route rights to Osaka, Japan. This was subsequently followed up with second carrier rights to Hong Kong. Rather than being censured for anti-competitive behavior, Air Canada was rewarded with additional international rights. The course of events clearly indicates that the government finally took action to resolve the impasse by buying off Air Canada with additional route rights. The fact that the government choose to use a carrot instead of a stick is indicative of Air Canada's lobbying strength and strong government ties. One must question, however, the power balance in the regulatory system when a carrier can dictate terms to the government.

The award of rights to Air Canada in the heart of Canadian's international marketplace is inconsistent with the government decision to preserve two competitors in the domestic marketplace. The bulk of Canadian's profits are generated on its north pacific international routes. It simply makes no

sense to open up a weak carrier's most profitable routes to additional competition if you wish to allow the carrier to recover some semblance of fiscal strength. The market timing of the decision was also in question. The addition of Air Canada's service to Japan and the concurrent addition of a second Japanese carrier, took place at a time when traffic to and from Japan had fallen by 25%. It is noteworthy that Canadian has still not been able to return to profitable operation.

A lack of competition between Canadian carriers on international routes does not mean that the route is not competitive. International carriers provide aggressive competition on all major international routes. Additionally, Canadian carriers are often at a competitive disadvantage due to high taxes, conservative depreciation rules, and high Canadian interest rates. Dispersion of Canadian route rights only produces an additional competitive disadvantage for Canadian carriers. Again the situation is one where additional competition between Canadian carriers will produce short term savings for consumers, but the end result could be the failure of Canadian carriers and a loss of competition in both the domestic and international markets. Even if this event does not occur, if Canadian carriers continue to fight each other they will have little left with which to take on foreign competition.

C. Trans-border The demand for trans-border air transportation has increased greatly since 1966. The 1966 agreement was outdated and no longer serviced the needs of either country. The new agreement has been welcomed by users of the transportation system as it will make travel to and from the U.S. much easier. From a carrier perspective the agreement is predicted to provide more benefit to U.S. carriers than to Canadian carriers (Lewis, 1995, 10). U.S. carriers already enjoy 62% of the market and they have a significant advantage with respect to operating costs. Additionally, U.S. carriers often have the benefit of carrying Canadian traffic beyond the initial leg of the trip to a secondary destination. This is rarely the case for traffic to Canada.

Canada has tried to correct the imbalance by securing a phase in period for additional American service to Canada's three major markets. Canada's negotiators were relatively unsuccessful in this task as the protected period is only 2-3 years. Canada was somewhat more successful in obtaining landing slots and gate rights at major U.S. airports.

Overall, the open skies policy recognizes the economic realities of airline travel to the U.S. Canadian carrier interests have been made subservient to improved services for travelers and economic development for airport communities. Given the network and cost advantages of U.S. carriers and the economic importance of efficient transportation links to Canada's largest trading partner, this policy appears to be rational. It is interesting to note, however, that the direction taken under the Open Skies agreement is contrary to the recommendations of the Ministerial Task Force on International Air Policy which recommended that carrier interests should take precedence (Lewis, 1995, 10). The Canadian negotiating position with respect to the bilateral agreement reflected the perceived need to use the agreement to ensure the viability of the domestic industry. This approach delayed the agreement and is yet another example of inconsistency within Canadian air policy.

D. Conclusions Regarding Present Policy - The present package of Canadian airline policy is a collection of individual policies which have been patched together over a period of approximately 12 years. The policies are generally clear in their stated objectives to put the needs of travelers ahead of the needs of carriers, but there is no vision expressed for the role that Canadian carriers should play in the system. Additionally, there is little recognition of the interdependencies between domestic markets and international markets. This is a crucial failure as approximately 40% of the revenue generated by Canada's airlines is generated on international routes and growth in air travel markets is concentrated in the international and trans-border markets.

It can also be argued that the present policy does not reflect the economic structure of the Canadian industry. Canada's approach to airline deregulation was borrowed from the U.S. without adequate recognition of the differences in structure between the two countries' airline industries.

Finally, the timing of recent changes to Canadian air policy has been poor. Additional competition is being created at a time when the domestic industry is weak and vulnerable. Carriers are being encouraged to extend themselves into new international markets when their balance sheets are extremely weak. As the industry is highly subject to business cycles, the results could be disastrous if the industry meets with another recession before Canadian carriers can improve their financial position.

VII - Where to from Here?

A. International Market - There is conflicting evidence as to whether strategic national industrial policies work. Those who oppose creation of such strategies claim that "...governments are not particularly successful in picking winners, but losers are incredibly adept at picking governments." (Sylvia Ostry as quoted in Courchene, 1992, 777). Because of the uncertain track record, some have suggested that governments should not attempt to set industrial strategies. The fact remains, however, that international aviation routes have to be negotiated between states. This means that the government cannot avoid involvement in establishing strategy for Canadian air carriers.

I have argued that a coherent air policy incorporating all market areas is essential for Canada's economic development. The air transport system is a vital part of the infrastructure required to connect our businesses and travel destinations to global markets. I have also argued that a strong domestic airline industry is desired, not necessarily as a economic objective for the nation,[13] but as a means to deliver transport services tailored to the specific needs of Canadians and to continue the role of domestic carriers in providing political linkages across Canada.

Our airline policy must reflect the relative strengths and weaknesses of our domestic carriers within the global environment. Michael Bliss in commenting of Martin Frobisher's adventures in Canada remarked that, "No amount of faith, subsidy or aid, can produce gold where it does not exist." This comment should also be recognized with respect to Canada's airlines. The role carved out for domestic carriers within our air policy must be consistent with their capabilities. It is unfortunate that these capabilities have recently been greatly reduced by destructive

---

[13] While the economic contribution of the domestic airline industry to Canada is argued to be subordinate to the economic interests of the industries which it serves, the contribution is still very substantial. The industry employs in excess of 50,000 people in relatively high skill and wage positions. Given the present liberal government's strategic issue of job creation, strong policy arguments could be advanced to support a level of protection for the domestic industry on the basis of job security. I have, however, chosen to focus on other aspects of the policy debate.

competition, but this is now the economic reality. We must define a sustainable role within this environment.

Tae Oum has defined a vision for the role which Canada might play in the world aviation market. He argues that globalization of the airline industry is inevitable and that it is only a matter of time before the industry forms global networks. Each network is expected to be created from the alliance of "anchor" carriers from each of three major continental regions - North America, Europe and Asia. Oum has also indicated that the world airline market will only be able to support 5-6 networks and that Canadian air policy should be designed to prepare and strengthen our industry to assume an anchor carrier role within one of the global networks. In short Oum has advocated the consolidation of Canada's carriers into one stronger international player (1993, 14-28).

> " This article advocates that the governments should set policies in such a way as to maximize the long run combined consumer and producer surpluses plus external economic benefits. Severe economic penalties are associated with not being a major partner in the global network formation process. The adoption of a pro-competitive policy especially when it is motivated by short run consumer surplus maximization, may significantly reduce the future ability for a nation's carriers to earn surpluses or for that nation to enjoy the increased economic activities and employment arising from senior partner status....The strategic industrial policy usually involves trading off short run benefits for much larger gains in the long run....With respect to the Canadian situation, unless some action is taken it is likely that global airline networks will bypass Canada altogether" (Oum, 1993, p27)

Oum is clearly an advocate for "forcing the pace" of international development of Canada's airline industry, but are his views realistic? [14] While there is a strong trend towards globalization within the international industry we are a long way from the global structure which is presented.

---

[14] It should be noted that Oum's views on consolidation of the Canadian airline industry may reflect his position as a consultant to the industry. Positions he expresses in his recent papers are in conflict with positions expressed in prior work authored jointly with M. Tretheway.

The world aviation market is highly complex and still heavily regulated. It is my view that we will continue to move towards the structure presented by Oum but that consolidation will be slow and that there will be many regional niches that do not conform to the general industry structure.

Even if one accepts Oum's view, is it realistic to believe that a Canadian carrier could be attractive as an "anchor" alliance partner? Even if the two Canadian carriers were combined, the combined carrier would still be small in comparison to the major U.S. carriers, ranking only 7th in revenue-tonne-kilometers. In terms of revenue, Canadian's international business ranked 23rd amongst the world's airlines while Air Canada's ranked 24th. On the domestic revenue side Air Canada ranks 15th while Canadian ranks 18th (IATA, 1994).

More important than relative size is the fact that neither Canadian carrier has direct access to the major part of the continental market. U.S. traffic to Asia is 11 times greater than Canadian traffic (Oum, 1993, 21). Similarly, U.S. traffic to Europe is 6 times greater than Canadian traffic. These facts dictate that the anchor carrier will be based in the U.S. Air Canada's initiative to acquire a 25% interest in Continental Airlines recognized this reality.[15] Such approaches may allow Canadian carriers to gain a greater share of the benefits from emerging global alliances and therefore should be encouraged, however, it must be recognized that the benefits of this approach are limited because U.S. policy limits foreign ownership of U.S. airlines to 25%. The analysis leads to the logical conclusion that Canada can not hope to create an anchor carrier in a global network. The best that we can hope to become is a minority interest partner. To restructure Canada's airlines to achieve anchor carrier status is beyond the power of government. As Kroeger remarks in reference to King Canute, "There are some things that governments just can't do." (Kroeger, 1).

Having rejected the anchor carrier role for our carriers we are still left with the dilemma of finding a sustainable role for Canadian carriers to play which also serves the interests of Canadian travelers. Both Canadian

---

[15] Air Canada's recent decision to divest itself of this share is, in my view, a financial decision related to their weakened position and their need to refinance a major portion of their debt in the coming year.

and Air Canada have now established alliances with U.S. carriers. Our policy should recognize and support these alliances. It would seem that a rational approach would be to focus on making the two carriers as attractive as possible as complementary second tier carriers.

One of the advantages that Canada can contribute to an alliance is an extra set of bilateral license agreements. Division of route rights amongst Canadian carriers should be made in recognition and support of the alliance strategy chosen by that carrier. This means that route awards should be concentrated in certain regions of the world and that care should be taken to minimize direct competition between Canadian carriers on international routes. Instead of trying to serve all areas of the world, Canadian carriers should be encouraged to focus their efforts on specific regions. Given the decimated financial position of the industry, focus is desperately needed.

One of Canadian Airlines major attractions is its strength on north-pacific routes. This is an area of weakness for Air Canada and it is one of the factors which attracted AMR to form an alliance. On the negative side, Canadian's alliance with AMR ultimately cost it other alliance positions with European carriers Lufthansa and Air France. Air Canada on the other hand is much stronger than Canadian in Europe and this strength complements Continental Airlines' relative weakness. While both Continental and Air Canada are weak in the North Pacific, it makes more sense for Continental and Air Canada to fill this hole through an Asian partner such as Cathay Pacific. If international route designations were tailored to support each Canadian carrier's alliance strategy, the restriction of direct competition between Canadian carriers on international routes would strengthen each domestic carriers position.[16] At the same time competition on the route would still be provided for the benefit of Canadian travelers.

The recommended approach is not simply a return to the division of the world policy. It contemplates that carriers would bid for licenses and that the awards would be based on who could present the best strategic

---

[16] While Oum proposes the adoption of multiple carrier designations on international routes, he recognizes that, "...awarding rights to a second Canadian carrier may prevent either carrier from developing a profitable service." (1995, 22)

argument on how they could exploit the route. Transport Canada would assume the role of determining which carrier could make best economic use out of the license. Periodic reviews and adjustment would be made to allow the system to adjust to changing traffic patterns.

B. Domestic Market - Returning again to the domestic market, the international route award policy which I have suggested might also be used to deal with extreme abuse of dominant position in the domestic market. If a carrier engaged in what the regulatory agency regarded as anti-competitive behavior, they could be punished by the revocation of international route rights. This ability exists today, but the regulatory agency has been reluctant to use it. In my view, such a proactive approach by the agency could be an effective way to police the industry. The agency would, of course, have to walk a fine line between encouraging true competition and preventing abuse. Additionally, some mechanism would need to be found to protect the system from special interest group pressure. A better solution would be to ensure that imbalances in the financial strength of the carriers are not created. As pointed out earlier, predation is only profitable when there is a large difference in the relative strength of the competitors.

New challenges for policy makers continue to arise in the domestic market. In 1996 WestJet introduced regional service throughout western Canada and Greyhound announced its intention to enter the market as a low cost national carrier. Given the industry's history of poor financial performance, one must question the wisdom of these ventures. It would appear that the romance of the industry has clouded sound business judgment and it will be interesting to see if these ventures will succeed. From a policy perspective, it would appear that WestJet's service is consistent with domestic objectives. WestJet is a regional carrier which is focusing on expanding the market by offering low cost fares. As their target market is people who now choose to drive rather than fly, their operations should not bleed substantial revenues from the national carriers. In short, regional competition will provide substantial benefits to consumers without significantly weakening the major carriers.

The Greyhound situation is more complicated. While Greyhound is also trying to offer low cost service they are targeting their efforts at the national market. Of equal importance is the fact that Greyhound is trying to skirt the existing regulations with respect to the 75% Canadian

ownership requirement for domestic airlines.[17] The National Transportation Agency has dealt firmly with Greyhound's actions by threatening to shut down the carrier if it should begin service. These actions are consistent with existing policy and, as detailed in other sections of this paper, they are also consistent with the long term interests of consumers of air transportation in Canada.

Greyhound has proposed to restructure its ownership to meet the Canadian content requirement. If they are successful in this initiative, present policy requires that they be granted a license. Introduction of additional domestic market competition on national routes will lead to fare wars which, given its already weakened position, may result in the failure of Canadian Airlines. In recognition of this possibility the government should implement a temporary moratorium on new licenses for trans-continental service. Once balance is restored to the industry the moratorium should be lifted as it is important to maintain contestable trans-continental markets in order to ensure that the trans-continental carriers continue to provide competitive fares.

In equilibrium then, the domestic market would rely on competition between equally matched major carriers. The costs associated with predatory pricing would make the practice unattractive and open access to markets would help to ensure that the majors do not collude to charge monopoly fares. In the event that further intervention becomes required, the government would have access to its discretionary power with respect to international route awards to police the system.

C. Trans-Border - The proposed international policy is consistent with the domestic policy of maintaining competition between the major carriers on trans-continental routes but it must also be consistent with policy in the trans-border sector. The Open Skies agreement opened up the ability to connect any point in the U.S. with any point in Canada. This is consistent with my recommendations for international policy as it allows the formation of multiple links between Canadian carriers and their U.S. alliance partners. In addition, it allows Canadian carriers to take advantage of certain niche markets. For example, Air Canada is well

---

[17] Greyhound is 67% owned by the Dial Corporation of Phoenix, AZ. Their business plan involves the contract use of aircraft provided and flown by Kelowna Flight Craft which is Canadian owned.

positioned to take advantage of passenger's distaste for overcrowded hubs by use of an "overfly" strategy on selected markets. As a side benefit to Canada, Air Canada would utilize Canadian built Canadair jets on such services.

On the other hand, the Open Skies agreement continues to restrict cabotage and both countries have preserved their 25% limit on foreign ownership of their airlines. These restrictions are consistent with the domestic policy of maintaining Canadian carriers. If Canada were to drop these restrictions Canadian air travel would be integrated into a continental hub and spoke system. It is argued that such a system is politically unacceptable.[18]

Finally, I believe that further consideration needs to be given to removing the disadvantages Canadian carriers face with respect to fuel costs and the tax structure. Canadian carriers face a difficult business environment in the global market for air travel. The government should endeavor to provide them with a cost base which is equivalent to their international competitors.

In summary, my recommendations for an integrated air transport policy are:

1. Canada should explicitly recognize that a balanced system of domestic competition between two major carriers is in the national interest.

2. Open access to domestic markets should be maintained as contestable markets will help to ensure that domestic markets remain competitive. However, a temporary moratorium on new licenses for the trans-continental market should be implemented until the industry's financial position improves.

3. Canadian international policy should discourage competition between Canadian carriers on international routes and it should support competition with international carriers.

---

[18] Restrictions on foreign ownership of domestic airlines have traditionally been based on national security reasons. In the event of war or natural disaster governments wish to have access to a fleet of aircraft. Such arguments appear to be weak in the Canadian context.

4. Canada should encourage its domestic carriers to form alliances with complementary international carriers.

5. International route awards should be competed for on the basis of which national carrier can make better use of the rights within its alliance structure.

6. The government should assume a more proactive role in the health of its aviation industry. International route awards should be used to police the domestic industry in cases of extreme abuse.

7. The competitive approach established in the open skies agreement should be maintained as is.

8. Existing restrictions on foreign ownership of Canadian carriers should be preserved.

9. Action should be taken to ensure that Canadian carriers fuel cost and tax structure is competitive with major foreign carriers.

These policy recommendations are designed to establish a more equitable balance between the long term interests of consumers and the long term interests of carriers. A largely pro-competitive stance is maintained in the domestic and trans-border markets while competition between Canadian carriers on international routes is eliminated. In summary, the revisions which I propose recognize that the transportation system exists to service the needs of consumers, but they also recognize that a healthy air travel industry is consistent with the long term interests of consumers. If these recommendations were to be implemented, a stronger domestic industry would emerge - one that would be better able to service the needs of Canadian travelers into the next century.

References

Barone, S.S., Javidan, M., Reschenthaler, G. B., & Kraft, J.H. "Deregulation in the Canadian Airline Industry: Is There Room for a Large Regional Airline?" The Logistics and Transportation Review. Vol 22, #4, (1986):421-448

Bliss, Michael. "Forcing the Pace: A Reappraisal of Business-Government Relations in Canadian History." Unpublished Course notes - University of Toronto , Theories of Business-Government Relations 105 (undated): p. 5

Brander, J.A., Government Policy Toward Business. Toronto: Wiley & Sons, 1995

Christopher, John. "Domestic Airline Policy in Canada." Library of Parliament Current Issue Review. 82-14 (1982): p 1-17

Christopher, John, "Canadian Airline Industry" Library of Parliament, Current Issue Review. 89-2E, (1993): 1-15

Courchene, Thomas J. "Mons pays, c'est l'hiver: reflections of a market populist." Canadian Journal of Economics XXV, No. 4 (1992): p. 759 - 791

Dixit, A.K., Nallebuff, B.J. Thinking Strategically. Norton & Co. New York (1991)

Goldenberg, Susan. Troubled Skies. Whitby: McGraw-Hill-Ryerson Limited, (1994)

International Air Transport Association. World Air Transport Statistics. (1994)

Kroeger, A.., "Governing in the Millenium." Course readings package

Lewis, Ira. "United States-Canada Air Services: The role of Alliances in a Future Bilateral Agreement." Transportation Journal. Vol 34 Spring (1995): p 5-11

Madore, O., Shaw, D.J., "The Canadian Airline Industry: Its Structure, Performance and Prospects." Library of Parliament, Background paper BP-329E (1993): p 1-26

Oum, Tae. Personal E-Mail correspondence (1996)

Oum, T., Taylor, A.J., Zhang, A. "Strategic Airline Policy in the Globalizing Airline Networks." Transportation Journal. Vol 32 #3 (1993): p 14-30

Oum, T., Taylor, A.J. "Emerging Patterns in Intercontinental Air Linkages and Implications for International Route Allocation Policy." Transportation Journal. Vol 34 #4 (1995): p 5-27

Posner, Richard A. "Theories of Economic Regulation." The Bell Journal of Economics and Management Science. (1974): p 335-357

Smith, M. "Canadian Airlines and AMR Proposal: National Transportation Agency Decision." Library of Parliament Background Paper, BP-337E, (1993): p 1-9

Transport Canada. Canada's International Air Transportation Policy. December, 1994: p 1- 9

Tretheway, M.W. "The Characteristics of Modern Post-Deregulation Air Transport." Report prepared for the Ministerial Task Force on International Air Policy, (1991)

Windle, R., Dressner, M. "Competition at "Duopoly" Airline Hubs in the U.S." Transportation Journal. Vol 33 #2 (1993): p 18-30

Acknowledgments

I would like to thank the following individuals for taking the time to submit to interviews in the course of researching this subject:

Professor Garland Chow, University of British Columbia
Professor Tae Oum, University of British Columbia
Michael Tretheway, Vancouver Airport Authority

ISSN#
1183-2770

Designed and printed by Printing Services • University of Saskatchewan • CUPE 1975