MODELING BORDER WAIT TIMES OF CANADIAN TRUCKS IN ONTARIO: EVIDENCE FROM RECENT GPS DATA FOR 2012/2013

Rahaf HUSEIN, Kevin GINGERICH, Hanna MAOH, William ANDERSON

Introduction

Canada and the United States share the largest international border in the world. This border facilitates the movement of more than $2 billion in daily goods and services. Among the Canadian provinces, Ontario is highly dependent upon international trade attributing imports and exports to 31% of the gross domestic product (Anderson, 2012). The United States is Ontario's most significant trade partner with 2012 exports and imports valued at $148.2 billion and $147.5 billion, respectively (Gauthier, 2014).

While the Canada-US border is over 5,000 miles long, approximately 59% of Canada's total trade with the United States flows through three major Ontario-US crossings. These include the Ambassador Bridge between Windsor and Detroit ($91.9 billion yearly), the Peace Bridge ($60.3 billion yearly) between Fort Erie and Buffalo, and the Blue Water Bridge between Sarnia and Port Huron ($55.5 billion yearly) (Ontario Chamber of Commerce, 2005). Potential delays to the supply chain at the border increases the cost of international trade for businesses. This effect is further magnified for industries where time reliability becomes more important than the crossing time itself.

This paper analyzes the crossing time trends occurring at the three major Ontario-US crossings over the course of one year from September 2012 to August 2013. The data used for this analysis is derived from global positioning system (GPS) information pertaining to 850 Canadian owned carriers with trucks travelling throughout Canada and the United States. Based on the crossing time trends that were derived from the GPS data, multivariate regression models are estimated using the average hourly crossing time for a given month. The objective of the models is to explain the spatio-temporal variation in crossing times. The analysis in this paper provides novel results regarding border crossings trends in Ontario. This is possible by utilizing a fairly large truck movement dataset to analyze crossing time variations across daily and monthly time intervals and between several heavily utilized border crossing locations.

The rest of this paper is organized as follows. An overview of literature on cross-border delay modeling is discussed first. Next, the data used in this study are summarized, followed by an analysis of the crossing time trends. Regression models are then formulated based on these trends and the implications of the findings are discussed before providing conclusions in the last section.

Background

Efficient transportation flows across the Canada-US border are a necessity for continued economic stability and growth in both countries. Border delays and uncertainty reduce the profits of firms and erode the benefits of international trade (Anderson and Coates, 2010). With trade increasing at an annual rate of 10 percent (U.S. Census Bureau, 2014), research on border traffic becomes a necessity.

While academic research on cross-border wait times is limited, several different methods have been proposed to model traffic delays at the border. For instance, stochastic queuing theory can be applied to model wait times and optimize booth configurations based on arrival rate, service rate, and the number of open servers (Kim, 2009). Queueing models can be developed using traffic counts to represent the behavior of traffic flows as a function of determinants such as vehicle speed, environmental impact, and so on. Alternatively, a hazard-based duration model (Stathopoules and Karlaftis, 2002) can be used to model border delays such as those estimated by Paselk and Mannering (1994).

An application of border-delay models is the prediction of impending queues and higher wait times on a real-time basis, enabling vehicles to avoid delays where possible. Several studies have fashioned original attempts to use queueing analysis for a real-time prediction of traffic queues (Chan et al. 2003; Lin et al. 2014). Intelligent technologies and methodological advances for developing efficient, reliable, and cost effective methods have emerged in recent years to automatically estimate queueing and delay at border crossings. This includes the development of Artificial Neural Network (ANN) models utilizing sensor data as a real-time input for delay predictions (Khan, 2010), and an enhanced Spinning Network (SPN) method inspired by human memory (Lin et al. 2014).

Despite the constant development in technology, model prediction and validation is still a challenge due to the limited availability of accurate field data for border crossing wait times. Three different methods have been used to collect/generate data including: the physical collection of data through on-site surveyors or driver surveys (Paselk and Mannering, 1994; Goodchild et al., 2008); passive technology such as global positioning systems (GPS) (Goodchild et al., 2010), and synthesized data (Lin and Lin, 2001; Khan, 2010; Lin et al., 2014). When considering these methods, it is worth noting that data generated from on-site surveys typically span a short period of time or result in a limited number of observations. In contrast, GPS data like the one used in this study could overcome the limitations inherited in on-site physical surveys.

The GPS data used in this paper covers a one year period with a sufficiently large amount of data. To the author's knowledge, such a large dataset has not been used in the past to model delays at North American border crossings. In addition, this study provides novel results by comparing the crossing delays at three major North American crossings.

In a nutshell, the existing literature on border crossings is noticeably limited compared to the large impact these transportation facilities have on the economy.

Methodology

GPS Truck Data

The primary dataset used for this analysis consists of roughly 1 billion Global Positioning System (GPS) pings provided by Transport Canada. These data records are generated from 56,000 trucks belonging to 850 Canadian-owned carriers occurring over a one year time period from September 2012 to August 2013. For each GPS ping, a record is generated to provide an ID for the vehicle, the location of the truck (latitude and longitude), and a timestamp for when the ping occurred to the nearest second.

This study focuses on the crossing times at three of the busiest border crossings connecting Ontario and the U.S: the Ambassador Bridge; Blue Water Bridge; and the Peace Bridge. The crossing time for each bridge is calculated as the time required for a truck to cross a boundary geofence, as shown in Figure 1. Due to the time interval between successive GPS pings, a linear interpolation was employed to estimate the entry and exit times as a given truck crosses the geofence. To limit potential errors arising from this interpolation, data points located outside the geofence must be located within a 15 km by 15 km square zone to be considered in the calculation.

As can be seen in Figure 2, hourly variations exist in the average wait time of trucks at the three crossings. However, this variation is much more pronounced for vehicles traveling to the U.S although the hourly trends for the three borders are relatively similar. This is expected given that the inspections are all performed by the Canada Border Services Agency (CBSA) for vehicles traveling towards Canada and the Customs and Border Protection (CBP) agency for vehicles traveling towards the U.S.

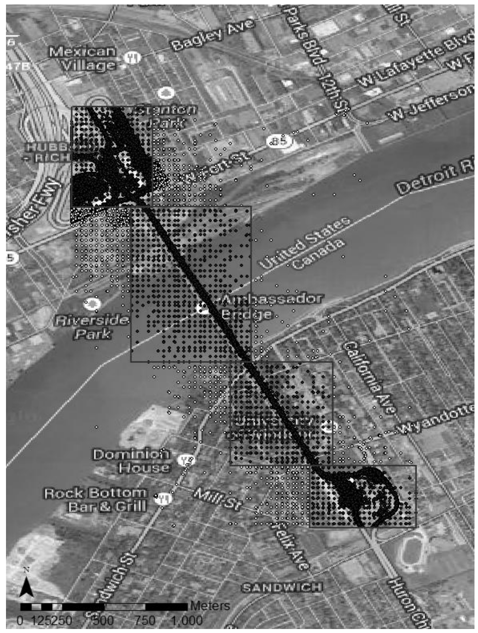

Figure 1: Border crossing geofence (Ambassador Bridge)

Model Specification

Regression models were specified and estimated to explain the crossing time trends in our border crossing data. Separate models were estimated for the direction of travel. Each of the regressions takes the following form $Ln(t) = \beta_0 + \sum_{i=1}^{N} \beta_i X_i + \varepsilon$, where the log-transformed t is the dependant variable representing the average crossing time for a given border crossing at a particular month and hour of the day. Subsequently, the models were estimated using 864 records representing each hour of the day for twelve individual months at three separate border crossings (24 × 12 × 3). Among the independent variables, β_0 is the constant intercept, β_i is the beta coefficient pertaining to variable X_i, N is the total number of variables, and ε is the error term capturing unknown random variables (assumed to be independent and identically distributed (iid)).

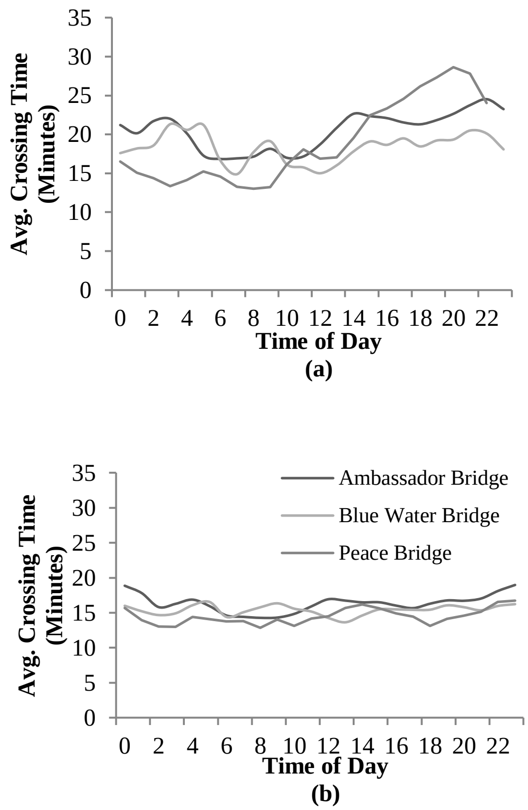

Figure 2: Average wait time for three border crossings by time of day for (a) U.S. bound traffic and (b) Canadian bound traffic

Explanatory Variables

The explanatory variables for the model are summarized in Table 1. A variable, *Arr/Dep*, was introduced to measure the ratio of trucks arriving and departing from the geofence for the given hour. This variable is introduced as a proxy for the level of service (i.e. amount of queuing and congestion occurring) at a given border crossing. When the arrival to departure ratio is greater than 1, the inspection booths are oversaturated indicating that more trucks are arriving to the geofenced area compared to outgoing trucks. Therefore the ratio reflects the level of service (i.e. service rate) for the different crossings by the hour of day. The parameter associated with *Arr/Dep* is expected to exert a positive influence on the crossing time since a larger queue inside the geofence will typically result in a larger wait time.

In order to isolate the wait time for the trucks at each border crossing, explanatory variables were included for each of the three border crossing locations (*Ambassador*, *Peace*, and *BlueWater*). The Ambassador Bridge was used as the base reference category and has the highest average crossing time of the three border crossing locations with overall averages of 21.23 minutes and 16.75 minutes for trucks heading to the US and Canada, respectively. In addition to the location variables, 24 indicator variables are included to account for each hour of the day ($Hour_h$). Hour_8, representing trips crossing the border between 8 AM and 9 AM, was chosen as the reference category.

Besides to the hour of the day, the season of the year is expected to have some impact on the crossing times. In the models estimated here, the *Summer* variable was chosen as the reference category. Based on our data exploration, the other three seasonal variables (*Fall*, *Spring*, and *Winter*) are expected to have negative coefficients for the Canadian bound models. However, these are expected to have positive coefficients in the U.S. bound model since our exploration showed that the fall and spring seasons have higher average crossing compared to the summer season. In comparison, the winter season is expected to have a negative coefficient based on the observed lower average crossing times. Finally, some interaction terms were also developed to control for noticeable outliers in the trends developed by the hour of the day and season.

Table 1 Variables used in regression model

Variable	Description	Expectation
Arr/Dep	The ratio of arrival to departure volume of trucks crossing the border in a given time period	+
Ambassador	1 if the observation on crossing time pertains to trucks crossing the Ambassador Bridge; 0 otherwise (Reference Category)	+
Peace	1 if the observation on crossing time pertains to trucks crossing the Peace Bridge; 0 otherwise	-
BlueWater	1 if the observation on crossing time pertains to trucks crossing the Blue Water Bridge; 0 otherwise	-
Summer	1 if observation on crossing time pertains to the following months: June, July and August; 0 otherwise (Reference Category)	+
Fall	1 if observation on crossing time pertains to the following months: September, October and November ; 0 otherwise	+/-
Winter	1 if observation on crossing time pertains to the following months: December, January and February; 0 otherwise	-
Spring	1 if observation on crossing time pertains to the following months: ; March, April and May; 0 otherwise	+/-
$Hour_h$	1 if observation on crossing time pertains to a specific hour *h* of the day (*h* = 0, 1, …, 23); 0 otherwise ($Hour_8$ is the Reference Category)	-
$Inter_1$	1 if observation on crossing time pertains to trucks crossing the Peace Bridge during the months of	+

	July and August during the hours 14 – 20 (inclusive); 0 otherwise	
$Inter_2$	1 if observation on crossing time pertains to trucks crossing the Blue Water Bridge during the months of July and August during the hours 14 – 21 (inclusive); 0 otherwise	+
$Inter_3$	1 if observation on crossing time pertains to trucks crossing the Peace Bridge during the months of July and August during the hours 19 – 23 (inclusive); 0 otherwise	+
$Inter_4$	1 if observation on crossing time pertains to trucks crossing the Blue Water Bridge during the months of July and August during the hours 1 – 5 (inclusive); 0 otherwise	+
$Inter_5$	1 if observation on crossing time pertains to trucks crossing the Blue Water Bridge during the months of July and August during the hours 16 – 20 (inclusive); 0 otherwise	+
$Inter_6$	1 if observation on crossing time pertains to trucks crossing the Ambassador Bridge during the months of June, July and August during the hours 2 – 6 (inclusive); 0 otherwise	+
Out_1	1 if observation on crossing time pertains to a residual between 5 and 10; 0 otherwise	+
Out_2	1 if observation on crossing time pertains to a residual less than -5; 0 otherwise	-
Out_3	1 if observation on crossing time pertains to a residual greater than 10; 0 otherwise	+

Results and Discussion

The results of the estimated regression models are summarized in Table 2. The estimated parameters were able to explain an acceptable percentage of the observed variability in trucks' average crossing time. While not all the parameters in the models were significant, they were kept in the models for comparability advantage. For instance, the *Spring* variable was not significant in the U.S. bound model, but was significant for the Canadian bound model. This suggests that, other things being equal, the trucks' average crossing time in the spring season is no different than in the summer season for the U.S. bound.

The *Arr/Dep* variable showed a positive and significant relationship to the *Ln(Average Crossing Time)* in the two models, as expected. Furthermore, the seasonal variables (*Winter*, *Spring*, *Fall*) behaved as expected for the most part while holding the *Summer* variable as the reference category. Also, they were mostly significant suggesting that the seasons of year do have an effect on the average crossing time.

As expected, the choice of bridge crossing variables (*Bluewater Bridge* and *Peace Bridge*) showed a negative relationship when holding the *Ambassador Bridge* variable as the reference category. The significance of the crossing location parameters in most cases suggests that the average crossing time is also affected by the choice of bridge crossing. This result could be attributed to a number of reasons including, but not limited to, the difference in the geometry of each bridge, the number of open lanes, and the truck volume served by the bridge.

The use of interaction terms proved to be highly beneficial as these compound variables improved the performance of the models. Additionally, after running the regression models, some outlier terms were created to control for bias in the dataset. The models' explanatory power of average crossing times was acceptable but not astounding (Adj. R-square is 0.73 and 0.43 for the US-Bound and Canada-Bound models, respectively). However, the predicted crossing times are fairly acceptable. This is presented when taking the

Table 2 Regression Parameter Estimation Results

Variable	US-Bound		Canada-Bound	
	Beta	t-stats	Beta	t-stats
Constant	2.54	43.25	2.64	53.57
Arr/Dep	0.27	5.82	0.14	3.18
Winter	-0.04	-2.98	-0.08	-7.27
Spring	-0.01	-0.41	-0.06	-5.63
Fall	0.04	3.50	-0.06	-5.34
Bluewater Bridge	-0.14	-14.18	-0.07	-7.48
Peace Bridge	-0.16	-14.94	-0.14	-16.05
$Hour_0$	0.19	5.66	0.15	5.90
$Hour_1$	0.15	4.55	0.08	2.89
$Hour_2$	0.19	5.48	0.01	0.43
$Hour_3$	0.26	7.22	0.01	0.48
$Hour_4$	0.20	5.68	0.09	2.94
$Hour_5$	0.14	3.80	0.04	1.24
$Hour_6$	0.02	0.60	0.00	0.03
$Hour_7$	-0.05	-1.32	0.00	0.10
$Hour_9$	0.03	1.03	0.03	0.87
$Hour_{10}$	0.04	1.12	-0.01	-0.19
$Hour_{11}$	0.09	2.99	0.03	0.91
$Hour_{12}$	0.09	3.18	0.03	1.14
$Hour_{13}$	0.15	4.92	0.05	1.46
$Hour_{14}$	0.24	7.82	0.09	3.22

$Hour_{15}$	0.30	9.78	0.08	3.26
$Hour_{16}$	0.27	8.26	0.07	2.59
$Hour_{17}$	0.28	8.54	0.03	1.07
$Hour_{18}$	0.30	8.70	0.03	1.09
$Hour_{19}$	0.33	10.39	0.06	2.37
$Hour_{20}$	0.36	11.07	0.06	2.47
$Hour_{21}$	0.43	14.02	0.08	3.08
$Hour_{22}$	0.39	12.44	0.14	5.22
$Hour_{23}$	0.29	9.25	0.17	6.78
$Inter_1$	0.65	20.51	-	-
$Inter_2$	0.05	2.10	-	-
$Inter_3$	-	-	0.22	8.22
$Inter_4$	-	-	0.11	2.47
$Inter_5$	-	-	0.12	5.34
$Inter_6$	-	-	0.05	1.47
OUT_1	0.34	32.16	0.35	13.73
OUT_2	-0.35	-17.88	-	-
OUT_3	0.54	15.35	-	-
Number of Observations	864		864	
Total Trucks Crossing	197, 570		190, 205	
R^2	0.741		0.454	
Adj. R^2	0.730		0.431	

correlation between the predicted and observed crossing times based on the estimated models. The results are promising with a correlation coefficient of 0.89 and 0.69 in the US-Bound and Canada-Bound models, respectively.

Conclusion

Regression models were estimated by analyzing the crossing time trends for trucks at three major Ontario-U.S. crossings. These models were created using a detailed GPS data provided by Transport Canada that spanned a study period of one year. Models by direction of travel were estimated to compare the movements from and to Ontario. The estimation results showed that the average crossing time of trucks differs by the choice of bridge crossing, the hour of day, and the season of year. Also, the level of service which depends on the arrival and departure rates is also significant. Furthermore, the consideration of the outlier variables in the models had a significant influence on the overall fit. For instance, the models improved their *Adjusted* R^2 when controlling for the identified outliers. As it stands, the estimated models are able to predict reasonable trucks' average crossing times for the three different crossings when controlling for the observed outliers. However, the U.S. bound model showed fairly better predictions compared to the Canadian bound model.

Our study suffered some limitations that are worth noting for future research. First, the analyzed data is based on a sample of carriers; hence, there might have been different proportions of the companies using the various analyzed crossings. Second, the estimated models did not have information about the number of open booths at each crossing. While the ratio of arrivals to departures was used as a proxy for the level of services, the number of available booths by hour of the day can improve the predictive ability of the estimated models. Lastly, the utilized average crossing times did not differentiate between the classes of border crossing trucks (i.e. trucks moving through the Free and Secured Trade (FAST) program, trucks going through primary and secondary inspections). Such information, which is currently lacking, is expected to have an influence on the modeled crossing times. As such, the above limitations should be addressed when data become available in the future.

Acknowledgements

The authors wish to thank Louis-Paul Tardiff and Andrew Carter of Transport Canada for providing the GPS data used in the analysis. The first author is also thankful to the Natural Science and Engineering Research Council of Canada (NSERC) for providing her with a USRA scholarship to conduct the research presented in this paper.

References

Anderson, W. (2012) "The Border and the Ontario Economy." Retrieved from: http://www1.uwindsor.ca/crossborder/system/files/The_Border_Report2012.pdf

Anderson, W., & Coates, A. (2010) "Delays and Uncertainty in Freight Movements the Canada-US Border Crossings." *Transportation Logistics Trends and Policies: Successes and Failures, Proceedings of the 45th Annual Conference of the Canadian Transportation Research Forum*, pp. 129-143.

Chan, T., Lee, C., & Hellinga, B. (2003) "Real-Time Identification and Tracking of Traffic Queues Based on Average Link Speed." Retrieved October 1, 2014, from: http://www.civil.uwaterloo.ca/bhellinga/publications/Publications/TRB 2003 Queue Identification.pdf

Gauthier, A. (2014) *Ontario's Merchandise Trade with the World,* Library of Parliament, Ottawa, Canada

Goodchild, A., Globerman, S., & Albrecht, S. (2008) "Service Time Variability at Blaine, Washington, Border Crossing and Impact on Regional Supply Chains." *Transportation Research Record: Journal of the Transportation Research Board,* 71-78.

Goodchild, A., Leung, L., & Albrecht, S. (2010) "Free And Secure Trade Commercial Vehicle Crossing Times At The Pacific Highway Port Of Entry." *Journal of Transportation Engineering, 136*, 932-932.

Khan, A. (2010) "Prediction and Display of Delay at Road Border Crossings." *The Open Transportation Journal, 4,* 9-22.

Kim, A. (2009) "The toll plaza optimization problem: Design, operations, and strategies." *Transportation Research Part E*, vol 45, pp. 125-137.

Lin, F., & Lin, M. (2001) “Modeling Traffic Delays at Northern New York Border Crossings.” *Journal of Transportation Engineering, 127,* 540-545.

Lin, L., Wang, Q., & Sadek, A. (2014) “Border Crossing Delay Prediction Using Transient Multi-Server Queueing Models.” *Transportation Research Part A: Policy and Practice, 64,* 65–91-65–91.

Lin, L., Wang, Q., Huang, S., & Sadek, A. (2014) “On-line prediction of border crossing traffic using an enhanced Spinning Network method.” *Transportation Research Part C: Emerging Technologies*, 158-173

Ontario Chamber of Commerce (2005) *Cost of Border Delays to the United States Economy,* Ontario Chamber of Commerce, Ontario, Canada. Retrieved from: http://www.occ.ca/Publications/Cost-of-Border-Delays.pdf

Paselk, T., & Mannering, F. (1994) Use of duration models for predicting vehicular delay at a US/Canadian border crossing. *Transportation, 21(3),* 249-270. Retrieved September 1, 2014, from: http://link.springer.com/article/10.1007/BF01099213#page-1

Stathopoulos, A., & Karlaftis, M. (2002) Modeling Duration of Urban Traffic Congestion. *Journal of Transportation Engineering*, *128*(6), 587-590.

U.S. Census Bureau (2014) *Trade in Goods with Canada.* Retrieved December 2, 2014, from: https://www.census.gov/foreign-trade/balance/c1220.html#2000