How Truck and Rail Compete in Commodity Movement in the US?

Amir Samimi[a *], Abolfazl (Kouros) Mohammadian[b], Kazuya Kawamura[c]

[a] Ph.D. Candidate, Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 W. Taylor Street, Chicago, IL 60607-7023. Tel: (312) 996-0962. Fax: (312) 996-2426. Email: asamim2@uic.edu.

[b] Associate Professor, Department of Civil and Materials Engineering, University of Illinois at Chicago, 842 W. Taylor Street, Chicago, IL 60607-7023. Tel: (312) 996-9840. Fax: (312) 996-2426. Email: kouros@uic.edu.

[c] Associate Professor, College of Urban Planning and Public Affairs, University of Illinois at Chicago, 412 S. Peoria Street, Chicago, IL 60607-7064. Tel: (312) 413-1269. Fax: (312) 413-2314. Email: kazuya@uic.edu.

[*] Corresponding Author.

Paper submitted for Presentation in the Third International Conference on Innovations in Travel Modeling (ITM) of the Transportation Research Board (TRB), in Tempe, Arizona during May 9-12, 2010.

Word Count: 2,499

ABSTRACT

Mode choice is one of the most critical parts of any freight demand modeling framework. However, the amount of work on this issue is surprisingly modest mainly due to the absence of suitable data. This study introduces the binary logit and probit models that explain how truck and rail are chosen as the mode by the shippers, third party logistics, or receivers. The analysis of the data obtained from a nationwide establishment survey revealed that some shipment-specific variables, such as distance, weight and value of commodity, along with some mode-specific variables, namely haul time, and shipping cost have significant effects on the modal selection behaviors. Rail shipments were found to be more sensitive to the shipping cost, while road shipments are more responsive to the haul time. Different fuel price scenarios were also analyzed and revealed a low level of mode choice sensitivity to the fuel price.

Key Words: Freight, mode choice, behavioral model.

INTRODUCTION

Freight shipment decisions have been changing rapidly during the past two to three decades in response to the needs for leaner, more efficient supply chain systems. The complexity of today's logistics decision-making process presents a serious challenge for the freight demand modelers to provide reliable analysis tools to the policy makers and practitioners. This problem can be mainly attributed to the lack of appropriate disaggregate freight data, which prevents researchers to develop realistic behavioral models. It is obvious that furthering the understanding of the modal selection behaviors and having a more reliable analysis tool will facilitate the development of the broad strategies.

Mode choice is one of the most critical parts of any freight demand modeling framework. However, the amount of literature on this issue is surprisingly modest mainly due to the absence of suitable data. A direct comparison of shipment costs was the primary method in the most early freight mode choice models (Cunningham, 1982). However, reliability, flexibility, safety, and some other non-cost factors entered the analysis when the random utility models emerged (Norojono and Young, 2003). New supply chain concepts (e.g. just-in-time) were adopted by many companies, which subsequently influenced the shipping preferences (Hensher and Figliozzi, 2007) and required fundamental revision in the models. Based on a review of previous studies, the dominant factors impacting freight mode choice in the literature can be summarized as: accessibility, reliability, cost, time, flexibility, and past experience with each mode. This study introduces the binary logit and probit models that explain how truck and rail are chosen as the mode by the shippers, third party logistics (3PLs), or receivers. These models specifically look into transportation cost, distance, weight and value of commodity, and past experience with each mode.

DATA AND MODEL

Any disaggregated data on freight activities are difficult to obtain due to their scarcity and the concern for violating the confidentiality of the businesses that participated in the survey. Thus, it is not surprising that there is no disaggregate freight data at the national level in the U.S. that are publicly available. Therefore, our effort to develop a freight mode choice model had to begin with a data collection effort. An online survey was conducted at the University of Illinois in April and May 2009, providing information on 881 domestic shipments in the United States (Samimi et al., 2010). Basic information about each establishment along with data on five recent shipments, including origin, destination, transportation mode, type, value, weight, and volume of the commodity, cost and time of the entire shipping process, were obtained. Low response rate of such surveys could diminish or even nullify the credibility of the collected data, if not appropriately addressed. Therefore a comprehensive analysis of non-random selection bias was performed (Samimi et al., 2010) and revealed that size, location, and industry type of the firms have not significantly affected the probability of participation.

A proper choice model is sensitive to attributes of both decision-maker and choice alternatives. While characteristics of the decision-maker do not change across alternatives, the attributes of choice alternatives vary significantly from one alternative to the other (e.g., shipping time) and are typically collected only for the observed choice. One of the critical challenges in modeling freight modal selection is to obtain information on non-selected choices. In our case, shipping cost and time for using either truck or rail was obtained for each shipment in the survey. Using those data, the specifications of the non-selected choice were imputed in a machine learning module. Machine learning methods have been implemented in the field of transportation planning before (Mohammadian and Miller, 2003), and a more complete discussion on the topic can be found in the literature (Principe, 2000). Two artificial neural networks were constructed with two hidden layers and trained by NeuroSolution 5.07 package (Neuro Dimension Corporate, 2009). The first network was trained by the data for the rail shipments to impute the unobserved shipping times and costs for road shipments, while the other network used truck shipments for training to estimate the aforementioned information for rail shipments.

The most common framework used for choice behavior analysis in recent years has been discrete choice modeling approach. Two widely used forms of the discrete choice models are logit and probit models. Limdep econometrics software (Greene, 2002) was used for modeling purpose and final probit and logit models that estimate the probability of choosing between truck and rail are summarized in Table 1. Akaike and McFadden values are among many fit measures offered for binary choice models, which were used along with the chi-squared values for model selection (Train, 2003). All the estimated parameters in the final models turn out to be significant with a p-value of less than 0.05, and most of them are significant with 99% confidence interval. Wald, Likelihood Ratio, and Lagrange Multiplier tests, known as Neyman-Pearson tests (Greene, 2002), were also carried out to show the overall significance of the final models. Both models have pseudo R-squared values of more than 57%, and correctly predict 95% of the observations. Percentage of correctly predicted observations is usually high in binary choice models that predict a rare event, and in many cases this number could be misrepresented as the general explanatory power of the model. When the two possible outcomes are either rare or common event, binary models tend to over predict the latter, resulting in high rates of correct predictions at the expense of largely ignoring the rare event outcomes. For example, if 99 out of 100 choices are common and only 1 is a rare event, the model can attain 99% accuracy by simply predicting all cases to be common. Thus the percentage of rare events that are correctly predicted is a more valuable measure of predictive power for such models. In our case, choosing rail over truck could be considered as a rare event with only around 9% chance of occurrence in this data. Both models predicted more than 72% of rail shipments correctly, which is quite impressive especially for a freight mode choice model.

Since the shipping cost and time of unobserved modes were imputed in a machine learning module, it seemed necessary to control for potential multicollinearity between explanatory variables. Although collinearity is unlikely to be a serious issue when all the coefficients were statistically significant in a binary choice model, very large off-diagonal values were searched in the variance-covariance matrixes as the primary effect of multicollinearity. Variance inflation factors (VIF) were also estimated for all the independent variables. Kutner et al. (2004) suggested a VIF of 5 as the threshold that indicates a presence of serious multicollinearity. For our models, none of the variables had the VIF in excess of 3.5 (Table 1).

TABLE 1 Mode choice models

Item		PROBIT MODEL		LOGIT MODEL		VIF
		Value	t-ratio	Value	t-ratio	
Coefficient	CONSTANT	-5.902 *	-6.050	-10.808 *	-5.696	-
	DISTANCE (miles)	0.237E-03 **	2.273	0.452E-03 **	2.156	2.776
	WEIGHT (lbs)	0.310E-04 *	4.293	0.569E-04 *	4.195	1.564
	TRUCK-TIME (Shipping time by truck in days)	0.622 *	5.019	1.110 *	4.815	1.648
	RAIL-TIME (Shipping time by rail in days)	-0.094 *	-2.579	-0.176 **	-2.295	2.387
	TRUCK-COST-INDEX (Ln (TRUCK-COST(in USD) / (TRUCK-TIME * VALUE (in USD)))	0.388 **	2.532	0.670 **	2.361	3.408
	RAIL-COST-INDEX (Ln (RAIL-COST(in USD) / (RAIL -TIME * VALUE (in USD)))	-0.659 *	-3.474	-1.188 *	-3.331	1.099
	POTENTIAL-INTERMODAL (1: truck-rail intermodal is considered always or often as a potential transportation mode / 0: otherwise)	1.214 *	3.468	2.270 *	3.265	2.776
Fit Measures	Log likelihood	-47.141	-	-47.780	-	-
	Model Chi-squared	128.577	-	127.300	-	-
	Akaike I.C.	0.296	-	0.300	-	-
	Pseudo R-squared	0.577	-	0.571	-	-
	Correctly Predicted (%)	95.430	-	95.699	-	-
	Correctly Predicted (%) – only rail	72.727	-	72.727	-	-

* Significant at 99% confidence interval.
** Significant at 95% confidence interval.

DISCUSSION

Distance, weight, truck shipping time, rail shipping time, truck cost index, and rail cost index turned out to be significant in the final models. *DISTANCE* has a positive sign indicating that rail is more likely to be chosen as a transportation mode for long hauls. This finding is intuitively interpretable and was also confirmed in former studies (Oum, 1979). One explanation for this trend is that rail shipments have a higher base price compare to truck, which is diminished in the long hauls. Weight of the shipment is another significant variable in the models with a positive coefficient, indicating that larger shipments are more likely to be transported on rail. This observation is also in line with past studies. As indicated by Ever et al. (1996), past experiences with each mode plays a determining role in the selection of mode. *POTENTIAL-INTERMODAL* variable shows such effect in the models with positive coefficients, indicating that firms that always or often consider truck-rail intermodal as a possible option are more likely to select rail mode. This finding may seem trivial at the first glance, but from the modeling perspective the inclusion of such variable makes other coefficients more meaningful. For instance, shipping behavior of a firm preferring truck over rail may be mistakenly attributed to the differences in cost and/or haul time, while the real reason may have been that the shipper is unfamiliar with the rail. Therefore excluding such variables that capture the effects of shippers' knowledge or prejudice from the models may result in erroneous interpretation of the coefficients.

Cost and haul time of each transportation mode are other significant factors in mode selection. Having such mode-specific indicators enhances the explanatory power of the model, especially when modeling freight transport behaviors. A comparison between the coefficients of truck and rail transit time reveals that the choice probability for truck is more sensitive to velocity than for rail. An analysis of the elasticities of truck and rail haul time also indicated that the effect of truck travel time is almost 20 times greater for the truck mode. This shows that the time is a crucial issue especially when truck is preferred to rail. The cost index, which is defined for each mode as the log of shipping cost divide by the product of haul time and value of shipment shows that the choice of rail is sensitive to the cost. Rail shipments' sensitivity to the cost index is around 1.7 times greater than that for truck shipments. An interesting observation in the coefficients of time and cost variables is that shippers preferring truck are mainly concerned about the shipping time, and in general, less sensitive to the cost. On the other hand, the decisions on rail shipments are more sensitive to the cost, but not to time. This suggests that rail shipments are generally quite sensitive to the cost and easily react to changes in price.

Impact of fuel cost fluctuations on mode choice behaviors were also looked into. The results of the analysis suggest that freight modal decisions are very much inelastic to the fuel cost and do not change significantly with even a 50% increase in fuel cost. When the fuel price doubles, however, shippers start shifting to rail mode when fuel cost accounts for a large portion of the total cost. This may happen in long haul shipments in respectively low level of congestion. Two other scenarios explore 150 and 200 percent increase in fuel cost. Analysis of such scenarios is enlightening and essential for future decisions. In these scenarios, around 7 percent of total shipments are expected to shift to rail when the fuel price is a major component of total shipping cost. However, even when the fuel cost is not a large factor, a significant shift of around 3% is expected.

ACKNOWLEDGMENTS

The authors would like to deeply appreciate the National Center for Freight, Infrastructure, Research and Education (CFIRE) at the University of Wisconsin-Madison and Illinois Department of Transportation for funding this study.

REFERENCES

Cunningham, W.H.J. Freight Modal Choice and Competition in Transportation: A Critique and Categorization of Analysis Techniques. *Transportation Journal*, 21(4), pp. 66-75, 1982.

Greene, H.W. LIMDEP Version 8.0 User's Manual. Econometric Software Inc., Plainview, NY, USA, 2002.

Hensher, D., and M. A. Figliozzi. Behavioural insights into the modelling of freight transportation and distribution systems. *Transportation Research Part B*, 41, pp. 921–923, 2007.

Kutner, M. H., C.J. Nachtsheim, and J. Neter. Applied Linear Regression Models, 4th edition, McGraw-Hill Irwin, 2004.

Mohammadian, A. and E.J. Miller. Nested Logit Models and Artificial Neural Networks for Predicting Household Automobile Choices: A Comparison of the Performance. *Transportation Research Record,* No.1807, pp. 92-100, 2003.

Neuro Dimension Corporate. NeuroSolutions: Premier Neural Network Development Environment. <http://www.neurosolutions.com>, accessed July 15, 2009.

Norojono, O., and W. Young. A Stated Preference Freight Mode Choice Model. *Transportation Planning & Technology*, 26(2), pp. 195-212, 2003.

Oum, T.H. A Cross Sectional Study of Freight Transport Demand and Rail-Truck Competition in Canada. *Bell Journal of Economics*, 10, pp. 463-482, 1979.

Principe, J.C., N.R. Euliano, and W.C. Lefebvre. Neural and Adaptive Systems: Fundamentals through Simulations. John Wiley & Sons, New York, 2000.

Samimi, A., A. Mohammadian, and K. Kawamura. An Online Freight Shipment Survey in US: Lessons Learnt and a Non-response Bias Analysis. Submitted to the 89th Annual Transportation Research Board Meeting, Washington, D.C., 2010.

Train, E. K. Discrete Choice Methods with Simulation. Cambridge University Press, 2003.