MODELING FREIGHT TRAVEL DEMAND USING A STRUCTURAL EQUATIONS MODELING FRAMEWORK

Ram M. PENDYALA
Professor
Department of Civil and Environmental Engineering
University of South Florida, ENB118
4202 East Fowler Avenue
Tampa, FL 33620-5350
Tel: +1-813-974-1084
Fax: +1-813-974-2957
Email: pendyala@eng.usf.edu

Abstract: The modeling of freight travel demand has gained increasing attention in the recent past due to the importance of efficient and safe freight transportation to regional economic growth. Despite the attention paid to the modeling of freight travel demand, advances in modeling methods and the development of practical tools for forecasting freight flows have been limited. This paper attempts to make a contribution in this context by developing a structural equations modeling framework that can be applied to the modeling of freight travel demand using data contained in readily available commercial databases in the United States context. The model system is estimated at the level of the individual zip code and can be used to estimate freight flows by commodity by major modes between origin-destination pairs. The paper describes the database development, model framework, sample model estimation results, and directions for further model development in the future.

Key Words: Freight transportation, freight travel demand modeling, structural equations, freight flow data, multimodal freight logistics

1. INTRODUCTION

Freight transportation is a critical element in the overall demand-supply chain of commodities and services in a region. Freight transportation is critically tied to the economic growth and well-being of a region that freight transportation planning has become a major focus of transportation planning around the world (Cambridge Systematics, Inc., 1997). An important part of freight transportation planning is the ability to quantify and predict freight flows between origin-destination pairs in a region of interest (Cambridge Systematics, Inc. *et al.*, 1997). With the traveling public and policy makers concerned about the congestion, safety, and pollution implications of freight movement, it is imperative that models of freight flows be developed to assist in freight transportation planning efforts.

Major advances have been made in the development of freight transportation modeling methods and frameworks (Pendyala *et al.*, 2000; Regan and Garrido, 2001; Shankar and Pendyala, 2001). However, many of these methods have not seen application in practice, partially due to the lack of adequate data to support their estimation and application to forecasting. Freight transportation data has been traditionally difficult to collect due to the proprietary nature of the data. The absence of freight transportation data is particularly critical at the disaggregate (spatial and temporal) level, making the estimation of disaggregate models of freight transportation demand a challenge that needs to be addressed. While there are aggregate level freight transportation data sets such as the commodity flow survey (CFS) data in the United States, these data sets are generally insufficient to develop disaggregate models of origin-destination freight flows by mode and commodity.

The objective of this paper is to describe the development of a structural equations modeling framework for estimating origin-destination freight flows. The model development effort described in this paper uses freight flow information by mode and commodity at the zip code level for the state of Florida. This data, available from a commercial source called Reebie Associates (2005), was combined with population data from the 2000 U.S. Census (U.S. Census Bureau, 2005) and with employment data available from another commercial source, InfoUSA (InfoUSA.com, Inc., 2005), to form a comprehensive socio-economic and freight flow database for freight transportation modeling. All databases were available for the year 2000, thus ensuring consistency across databases.

This paper is organized as follows. The next section provides a brief overview of the model construct and relates it to the previous literature on freight transportation modeling. The third section describes the modeling methodology. Section 4 provides a description of the database used in the study. Section 5 presents sample model estimation results. Finally, concluding remarks are provided in Section 6.

2. MODEL CONSTRUCT

Considerable progress in the estimation and application of freight transportation demand models has been made in the recent past (e.g., Friesz, 2000; Garrido and Mahmassani, 2000; Holguin-Veras and Thorson, 2000; List and Turnquist, 1994; Southworth and Peterson, 2000). It is not possible to offer a comprehensive literature review within the scope of this paper. The significance of freight movement and activity is increasing in terms of both its role in the economy and its potentially adverse impacts on safety and congestion on the transportation system (e.g., Ogden, 1991; Souleyrette *et al.*, 1998). It is not surprising that freight transportation planning has attracted the attention of researchers over the past several decades (e.g., Allen, 1977; Baumol and Vinod, 1970; Bronzini, 1980; Crainic, 1987; Harker, 1985; Harker and Friesz, 1986; Slavin, 1976; Winston, 1983).

Some of the models used to forecast freight flows are simple growth factor models while others are more complex and accurate autoregressive integrated moving average models (ARIMA), elasticity models, network models of logistics, direct demand and aggregate demand models, disaggregate demand models, and economic input-output models. Freight transportation models can be used to support a host of planning applications including facility planning, corridor planning, strategic planning, business logistics planning, and economic development. Within the context of urban freight travel demand analysis, various truck trip generation methods have been developed (Cambridge Systematics, Inc. *et al.*, 1997; List and Turnquist, 1994).

Network models of logistics constitute freight network equilibrium models that attempt to model shipper and carrier decisions through a market clearing process (Harker, 1985; Harker and Friesz, 1986). Network models that focus on strategic freight network planning employed primarily to forecast, months or years into the future, freight traffic over specific network links and routes and through specific network nodes and terminals have been developed (Friesz, 2000). Work has also been done in the development and application of integrated digital representation of multimodal and transcontinental freight transportation networks for freight transportation modeling (Southworth and Peterson, 2000).

Several policy studies have looked at the role of intermodalism, changing customer requirements, and competition of supply chains in a global marketplace (Regan and Garrido, 2001). Studies have also been done in the area of potential effects of research and technology on urban freight movement and the need to educate policy makers and transportation planners for the purpose of developing national freight transportation policies (Cambridge Systematics, Inc., 1997). Work has been done to present a framework for organizing and identifying planning goals, key issues, and predominant commodities for intercity freight transportation (Souleyrette *et al.*, 1998).

The model construct adopted in this paper is largely in line with paradigms and freight transportation demand-supply relationships identified in the literature. Figure 1 shows the overall model framework that guided the model development effort of this study. Origin and destination population and employment characteristics influence total freight flow and modal freight flows between origin-destination pairs. Also, modal level of service characteristics including travel distance, travel time, and travel cost influence freight flows by mode. The model framework is simple and practical and can therefore be easily estimated on a database that can be assembled by any public agency that has resources to purchase commercially available databases.

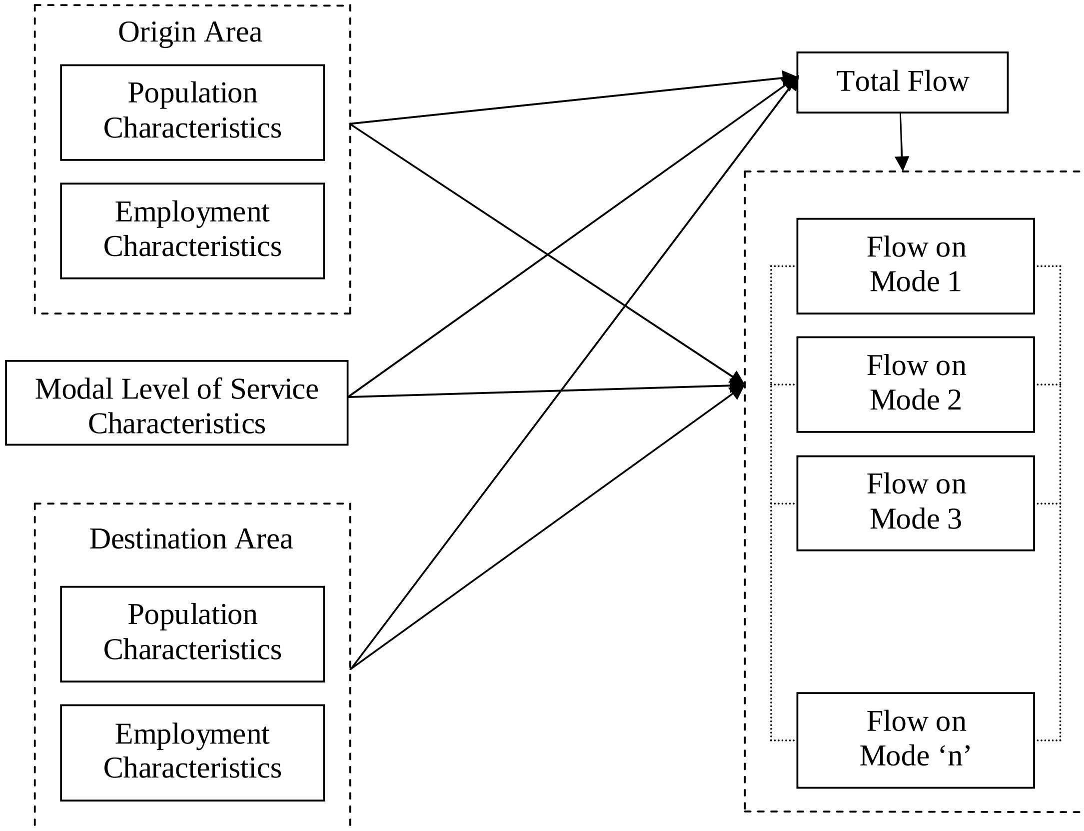

Figure 1. Conceptual Framework for Modeling Freight Transportation Movements

3. MODELING METHODOLOGY

The modeling methodology adopted in this paper is centered around the structural equations modeling framework that can be used to determine and model relationships among several dependent (endogenous) variables simultaneously. A typical structural equations model may be written as:

$$Y = BY + \Gamma X + \varepsilon \tag{1}$$

or

$$Y = (I - B)^{-1}(\Gamma X + \varepsilon) \tag{2}$$

where Y is a column vector of endogenous variables,
B is a matrix of parameters associated with right-hand-side endogenous variables,
X is a column vector of exogenous variables,
Γ is a matrix of parameters associated with exogenous variables, and
ε is a column vector of error terms associated with the endogenous variables.

Structural equations systems are estimated by covariance-based structural analysis, also called method of moments. In this methodology, the difference between the sample covariances and the covariances predicted by the model is minimized. The observed covariances minus the predicted covariances form the residuals. The fundamental hypothesis for the covariances-based estimation procedures is that the covariance matrix of the observed variables is a function of a set of parameters:

$$\Sigma = \Sigma(\theta) \tag{3}$$

where Σ is the population covariance matrix of observed variables,
θ is a vector that contains the model parameters, and
$\Sigma(\theta)$ is the covariance matrix written as a function of θ.

Equation 3 implies that each element of the covariance matrix is a function of one or more model parameters. The relation of Σ to $\Sigma(\theta)$ is basic to an understanding of identification, estimation, and assessments of model fit. The matrix $\Sigma(\theta)$ has three components, namely, the covariance matrix of *Y*, the covariance matrix of *X* with *Y*, and the covariance matrix of *X* (Bollen, 1989).

Let Φ = covariance matrix of *X* and Ψ = covariance matrix of ε. The unknown parameters in *B*, Γ, Φ, and Ψ are estimated so that the implied covariance matrix, $\hat{\Sigma}$, is as close as possible to the sample covariance matrix, S. In order to achieve this, a fitting function F(S, $\Sigma(\theta)$) which is to be minimized is defined. The fitting function has the properties of being scalar, greater than or equal to zero if and only if $\Sigma(\theta)$ = S, and continuous in S and $\Sigma(\theta)$.

The asymptotically distribution-free weighted least squares (ADF-WLS) method of estimation was used because the univariate distributions of the endogenous variables are non-normal in that there are substantial numbers of observations for each variable with zero value, which denotes no commodity flow between a zip code pair. For such distributions, coefficient estimates obtained using the traditional maximum likelihood estimation method will be consistent, but the estimates of parameter standard errors and the overall model χ^2 goodness-of-fit will likely be biased (Brown, 1984). Unbiased estimates of standard errors and goodness-of-fit can be generated using the ADF-WLS method (Brown, 1984).

The fitting function that is minimized in the ADF-WLS method is:

$$F_{WLS} = [s - \sigma(\theta)]' W^{-1}[s - \sigma(\theta)] \tag{4}$$

where s is a vector of censored correlation coefficients for all pairs of endogenous and exogenous variables, $\sigma(\theta)$ is a vector of model-implied correlations for the same variable pairs, and W is a positive-definite weight matrix. The best choice of the weight matrix is a consistent estimator of the asymptotic covariance matrix of s:

$$W = ACOV(s_{ij}, s_{gh}) \tag{5}$$

Under very general conditions:

$$W = \frac{1}{N}(s_{ijgh} - s_{ij}s_{gh}) \tag{6}$$

is a consistent estimator for the weight matrix, where s_{ijgh} denotes the fourth-order moments of the variables around their means, and s_{ij} and s_{gh} denote covariances. Browne (1984) demonstrated that F_{WLS} with such a weight matrix will yield consistent estimates, which are asymptotically efficient with correct parameter test statistics. These properties hold for very general conditions, and consequently such estimators are known as arbitrary distribution function, or asymptotically distribution free (ADF) estimators. ADF-WLS estimators are available in several structural equation model estimation packages including AMOS (Arbuckle, 2000) that was used for estimating the model presented in this paper.

4. DATABASE PREPARATION AND DESCRIPTION

Two widely known databases for freight transportation planning are the Bureau of Transportation Statistics Commodity Flow Survey (CFS) (U.S. Census Bureau, 2005) and the Reebie Associates' TRANSEARCH database (Reebie Associates, 2005). While the former is a free publicly available database, the latter is a commercially available database that may be customized to fit the needs of the planning agency. The Commodity Flow Survey data provides information on commodities shipped, their value, weight and mode of transportation, as well as the origin and destination of shipments at the national, state, and large metro-area levels. Thus, it is quite a detailed aggregate level dataset that can be used to study overall trends in commodity flows between major geographic areas and offers useful control totals on freight movement volumes.

In this study, as the intent is to model freight movements at a more detailed geographic level, the aggregate CFS data is not sufficient. Therefore, the model estimation effort reported in this paper utilizes the Transearch freight flow databases available at the zip code level (Reebie Associates, 2005). The zip code was considered an appropriate level of disaggregation where the data could be considered reliable and the presence of significant missing data could be avoided. The Transearch 2000 database for the state of Florida was used as the primary source of commodity flow information for this study. The database also includes commodity flows (by ton) into, out of, and through Florida. However, in order to keep the model estimation database tractable, only those commodity flows that originated and ended at zip codes within Florida were used for model estimation. The estimation database contained commodity flows at the zip code-to-zip code level with commodities classified at the level of the two-digit Standard Transportation Commodity Classification (STCC) code. In order to reduce the commodity groups to a more manageable level, the commodities at the two digit STCC level were collapsed further into 17 commodity groups as shown in Table 1.

It should be noted that the Transearch database is not necessarily a complete and comprehensive coverage of all freight transportation flows. There are certain types of commodity movements that are not captured in the Transearch database. Moreover, zip code level data was purchased by the Florida Department of Transportation for all pairs of zip codes that accounted for about 70 percent of the total statewide freight flows. Remaining freight flow data was available only at the larger county level of spatial geography. This study utilizes only the zip code level data for model estimation and thus captures only about 70 percent of the intra-state freight flow in Florida.

Commodity flows were broadly assigned to four modes: truck, rail, water, and air. The truck mode was further subdivided into full truck load, less than truck load, and private truck load. Full truck load was defined as a "for-hire" commodity flow on a truck with greater than 10,000 pounds, and less than truck load was defined as "for-hire" commodity flow on a truck with less than 10,000 pounds. A Private Truck was defined as any company truck that is part of a private fleet. The Rail mode was also subdivided into rail intermodal and rail car load modes. Thus, the estimation dataset consists of freight flows between pairs of zip codes in Florida by commodity (two-digit STCC code) and in annual weight (in tons) for each of the following modes of transport: full-truck-load, less-than-truckload, private truck, rail carload, rail intermodal, water, and air.

Modeling freight flows between zip codes required three more pieces of information as per the model construct presented in Figure 1. First, socio-economic information represented by population and employment characteristics was needed. All population information was derived from the 2000 Census databases. Census data was obtained at the zip code level and appropriately matched to the commodity flow database so that each record contained the population characteristics of the origin zip code and the destination zip code. Employment characteristics were derived from the InfoUSA 2000 database, a commercial database that contains information on every employer in the state of Florida. Aggregation of the employer database was performed to the zip code level and information on employment by standard industrial classification (SIC) category was matched into the database to develop a comprehensive database that contained freight flow information, population information, and employment information.

Finally, as per the model construct, detailed information on modal level of service variables was needed. For every zip code pair, it would be ideal to have travel time, distance, and cost information by all modes identified in the database. This intensive effort is currently ongoing and, as such, all modal level of service variables have not yet been merged into the database. At this time, models have been estimated using simple map distance (center of zip code to center of zip code) as a measure of impedance. However, in the future, modal level of service variables will be included in the model specification to make sure that the model is sensitive to modal level of service attributes. In its current form (as presented in this paper), the model is not sensitive to modal level of service variables. The importance of the role of distance and trip length in freight transportation modeling is well recognized (Holguin-Veras and Thorson, 2000).

The distribution of total annual flow (weight) by commodity group is shown in Table 1. As the database is focusing heavily on intra-state movements, the warehousing commodity group is found to account for more than 50 percent of the flows by weight. Other major commodity groups include other minerals, food, and clay, concrete, and glass. The overall mode share of total annual commodity flow by weight is also shown in Table 1.

The truck mode accounts for the major portion of overall freight flow, carrying 78 percent of the total annual commodity flows. The truck mode is a combination of two types of trucks, for-hire and private truck. For-hire truck accounts for 37 percent of the total annual commodity flow and about 48 percent of the total commodity flow carried by truck. Private truck accounts for 40 percent of the total annual commodity flow and 52 percent of the total commodity flow carried by truck. Rail mode accounts for about 20 percent of the total annual commodity flow. Air accounts for a very small share of commodity flow by weight at less than one percent while water accounts for a slightly higher share at 2.8 percent.

Table 1 also presents mode shares by commodity group. The table shows the 17 commodity groups and the mode share for each commodity in percent by weight. Thus coal is carried completely by rail car while warehousing is completely moved by truck. Differences across commodity groups with respect to modal share are quite important and noticeable. Therefore, structural equations models of freight flow should ideally be estimated separately for each commodity group.

Table 1. Distribution of Freight Flows within Florida by Commodity Group and Mode

Commodity Group		Freight Flow in Transearch Database		Commodity Flow by Mode (%)									
				Truck			Rail				Total		
Code	Name	Weight (tons)	Percentage	FTL	LTL	PVT	CL	IMX	Air	Water	Truck Total	Rail Total	Total
1	Agriculture	71,257	0.07	48.40	0.00	17.15	0.00	27.30	7.13	0.00	65.57	27.30	100
2	Coal	149,729	0.15	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00	100.00	100
3	Other Minerals	15,149,353	15.09	0.00	0.00	0.00	95.02	0.00	0.00	4.98	0.00	95.02	100
4	Food	4,505,852	4.49	32.82	0.48	63.64	1.85	1.20	0.02	0.00	96.94	3.04	100
5	Non-Durable Manufacturing	844,392	0.84	29.79	2.81	65.21	0.00	0.74	1.18	0.26	97.82	0.74	100
6	Lumber	1,815,565	1.81	45.87	0.81	47.67	5.65	0.00	0.00	0.00	94.35	5.65	100
7	Paper	542,104	0.54	30.82	6.96	54.66	4.62	2.33	0.61	0.00	92.45	6.94	100
8	Chemicals	6,035,141	6.01	55.26	0.52	0.07	43.32	0.75	0.08	0.00	55.85	44.07	100
9	Petroleum	2,396,887	2.39	5.43	0.13	16.30	0.03	0.15	0.01	77.96	21.85	0.18	100
10	Rubber Plastics	80,190	0.08	24.43	6.60	67.87	0.00	0.70	0.00	0.38	98.93	0.70	100
11	Durable Manufacturing	57,924	0.06	25.09	5.52	20.31	0.00	8.70	39.83	0.55	50.92	8.70	100
12	Clay, Concrete, Glass	14,582,327	14.52	30.16	0.07	67.26	1.17	1.31	0.00	0.02	97.50	2.49	100
13	Primary Metals	220,887	0.22	92.16	0.62	0.00	3.71	2.01	0.01	1.48	92.80	5.72	100
14	Fabricated Metal Products	455,010	0.45	42.37	2.07	50.08	0.20	0.11	4.39	0.77	94.53	0.31	100
15	Transportation Equipment	150,573	0.15	86.74	2.79	0.35	4.66	0.00	3.82	1.62	89.89	4.66	100
16	Miscellaneous Freight	1,917,071	1.91	0.00	0.00	0.00	3.78	87.07	2.45	6.69	0.00	90.85	100
17	Warehousing	51,427,628	51.22	48.72	2.05	49.24	0.00	0.00	0.00	0.00	100.00	0.00	100
	Total	100,401,890	100	36.10	1.21	40.25	17.56	2.00	0.12	2.75	77.56	19.56	100

Note:
FTL: Full Truck Load
LTL: Less Than Truck Load
PVT: Private Truck Load
CL: Rail Car Load
IMX: Rail Intermodal Load

As the commodities vary with respect to density, value, and time-sensitivity, there may be fundamental differences in the relationships among variables that can be used to predict their flows. In the context of the overall study, 17 separate structural equations models were estimated; in addition, one model that covered all commodity groups was also estimated. For the sake of brevity, only this model is presented in the paper as an illustrative example; the 17 commodity group-based models are presented in Pendyala (2004).

Although the Transearch freight database has its share of errors and omissions, many areas in the U.S. are investing in the purchase of this data to develop statewide freight travel demand models. In this context, it is appropriate to develop structural equations models of freight flow using the Transearch databases as the objective of this paper is to develop practical models of freight flow that utilize data available at many state and local agencies. However, readers should note the potential limitations of the database used in this study and interpret model results presented in the next section with appropriate caution.

5. MODEL ESTIMATION RESULTS

This section describes the model specification and estimation results for the illustrative structural equations model presented in this paper. The model employed a host of exogenous (explanatory) and endogenous (dependent) variables to model freight flows by zip code pair. Exogenous variables may be divided into three groups: population demographic characteristics of the origin and destination, employment characteristics of the origin and destination, and the impedance (distance) between the origin and destination. Endogenous variables are commodity flows between origin-destination zip codes by mode. The distributions of the endogenous variables are highly skewed and non-normal with a large number of zero observations. Nearly 97 percent of the observations are zero observations in the data set. Even within the context of the ADF-WLS estimation method, such a heavily zero-inflated distribution leads to computational intractability. Therefore, log transformations of the variables are used in the estimation process. A unit value is added to all observations to avoid having to deal with the logarithm of zero, which is undefined.

Structural equations models were estimated on all zip code pairs in Florida. Table 2 presents the structural equation model estimation results for all commodity groups combined. Figure 2 constitutes the path diagram corresponding to the model in Table 2. The model provides excellent goodness-of-fit measures with the χ^2 statistic indicating that the model can not be rejected with a high degree of confidence (95 percent or higher) and the goodness-of-fit index (GFI) equal to unity. Thus, the model is clearly capable of capturing the key relationships influencing freight flows, even within the context of a large database (more than 850,000 records) where endogenous variables are highly skewed, zero-inflated, and non-normal.

The indications provided by the model are quite consistent with expectations and plausible. The table shows the direct effects, indirect effects, and total effects that constitute relationships among variables. A direct effect is one where a variable directly affects another variable as depicted by a direct arrow linking the two variables in the path diagram. On the other hand, an indirect effect is one where a variable influences another variable through a mediating variable. For example, in Figure 2, one can see that origin employment does not directly affect the total freight movement by rail. However, origin employment affects both total flow and total truck flow. In turn, total flow and total truck flow affect total flow by rail.

Table 2. Structural Equations Model Estimation Results: All Commodity Groups

Endogenous Variable	Intercept	Effect	Distance	Destination Employment	Destination Population	Origin Employment	Origin Population	Total Flow	Total Truck	Total Rail	Full Truck Load	Private Truck
Total Flow	-0.370	Total	-0.012	0.100	-0.025	0.099	-0.026	0.000	0.000	0.000	0.000	0.000
		Direct	-0.012	0.100	-0.025	0.099	-0.026	0.000	0.000	0.000	0.000	0.000
		Indirect	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Total Truck Flow	-0.011	Total	-0.009	0.098	-0.025	0.097	-0.025	0.978	0.000	0.000	0.000	0.000
		Direct	0.003	0.000	0.000	0.000	0.000	0.978	0.000	0.000	0.000	0.000
		Indirect	-0.012	0.098	-0.025	0.097	-0.025	0.000	0.000	0.000	0.000	0.000
Total Rail Flow	0.009	Total	-0.011	0.018	-0.004	0.019	-0.006	0.192	-1.545	0.000	0.000	0.000
		Direct	-0.003	-0.001	0.001	0.000	-0.001	1.702	-1.545	0.000	0.000	0.000
		Indirect	-0.008	0.019	-0.005	0.019	-0.005	-1.511	0.000	0.000	0.000	0.000
Full Truck Load	0.003	Total	-0.007	0.084	-0.021	0.083	-0.021	0.845	0.886	0.028	0.000	0.000
		Direct	0.000	-0.001	0.000	-0.001	0.000	-0.069	0.930	0.028	0.000	0.000
		Indirect	-0.007	0.085	-0.021	0.084	-0.021	0.915	-0.044	0.000	0.000	0.000
Private Truck Load	-0.001	Total	-0.007	0.090	-0.023	0.089	-0.023	0.899	0.949	0.017	-0.106	0.000
		Direct	0.001	-0.001	0.000	0.000	0.000	-0.065	1.074	0.020	-0.106	0.000
		Indirect	-0.008	0.091	-0.023	0.090	-0.023	0.964	-0.125	-0.003	0.000	0.000
Rail Car Load	0.001	Total	-0.011	0.017	-0.004	0.018	-0.006	0.191	-1.505	0.999	-0.008	-0.006
		Direct	0.000	0.000	0.000	0.000	0.000	-0.039	0.052	0.999	-0.008	-0.006
		Indirect	-0.011	0.017	-0.004	0.018	-0.006	0.230	-1.558	0.000	0.001	0.000

Note:

N = 859,329; $\chi^2 = 1.064$ with df = 6; p-value = 0.983; CFI= 1; RMSEA = 0.000

All Variables Significant at 95% level

All Variables are in Logarithmic Form

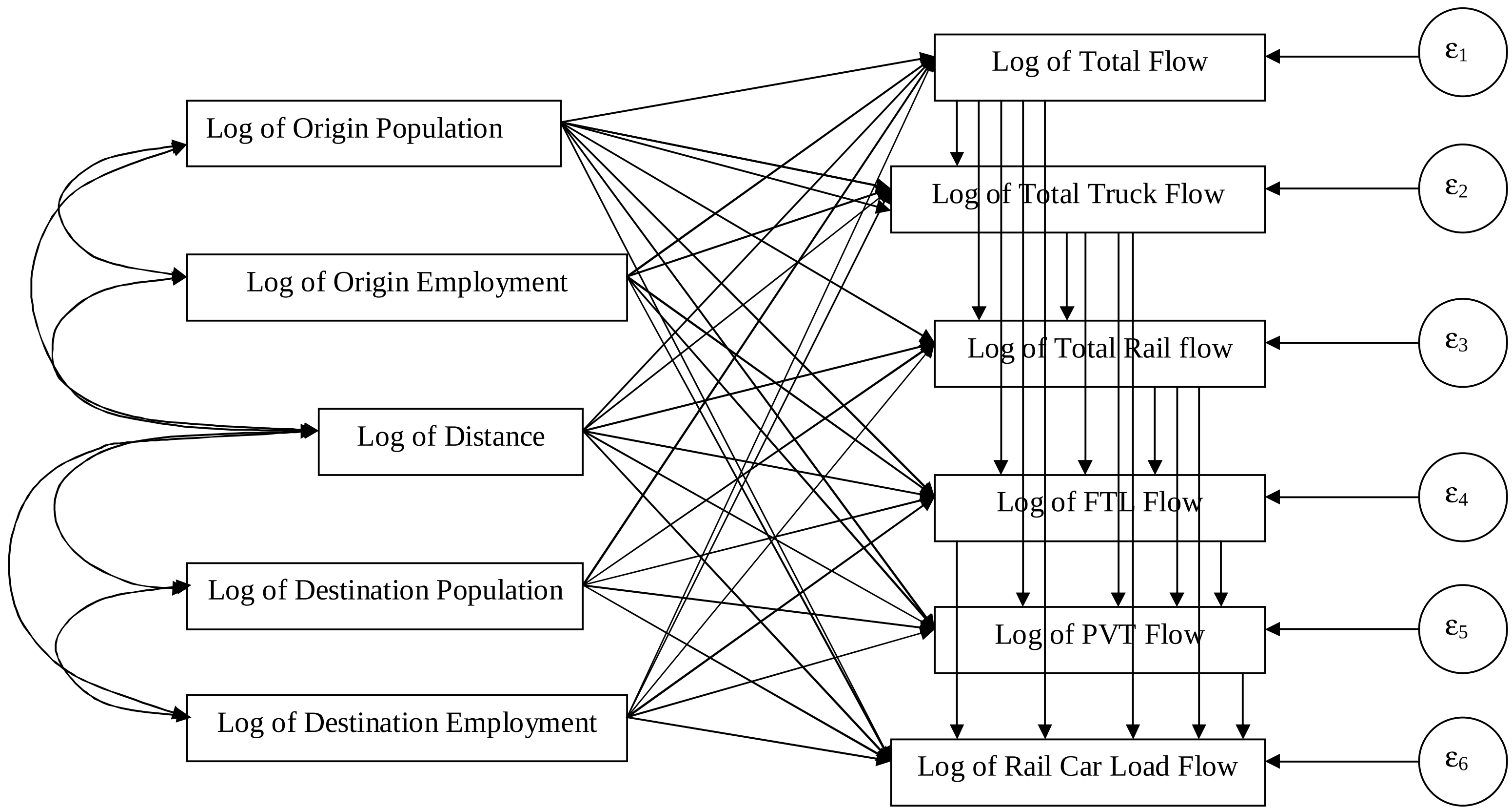

Figure 2. Path Diagram for the All Commodity Groups Structural Equations Model

Thus origin employment affects the total flow by rail through the intermediate variables, total flow and total truck flow. In some cases, a variable may have both a direct and an indirect effect on another variable. Then the total effect is the sum of the direct and indirect effects.

Table 2 shows that employment, both at the origin and destination end, has a positive impact on all flows by various modes. As expected, total flow affects total truck and rail flows with coefficients less than one. These coefficients represent how the total flow contributes to the different types of modal flows between an origin-destination pair. Also, as expected, the total flow by truck has a negative effect on the total flow by rail signifying the competitive nature of modal flows. It can also be seen that distance has a negative impact on all flows by various modes. While there are certainly strategic level decisions regarding facility location and customer clustering that tend to make distance a secondary variable in influencing freight flows, one can not ignore the possibility that distance is correlated with the quantity of freight flow between an origin-destination pair. For the state of Florida, recent Commodity Flow Surveys have indicated that about 60 percent of freight movements by value and 80 percent of freight movements by weight occur within the state. Clearly, distance is playing a major role in shaping the distribution and quantity of freight flows in Florida. In fact, about 70 percent (by weight) of all commodity flows originating in Florida travel less than 100 miles. The distance variable in the models simply reflects this tendency in the freight flow database and is found to offer statistically significant and intuitively plausible coefficients.

A rather surprising finding is that the origin and destination population variables are found to have a negative impact on freight flows in both the models. It was originally expected that population variables would have a positive impact on the quantity of flow. However, estimation results show that population variables are associated with negative coefficients. On the other hand, the employment variables have positive coefficients. Thus, it appears that employment is the key driver of freight flow activity while resident population is not a key driver of statewide freight flow activity. Business establishments, manufacturing and production operations, and other industrial land uses contribute to heavier volumes of freight flow. The presence of a residential population does not necessarily contribute positively to freight flows between origin-destination pairs at a statewide level.

Within an urban area context, when one is concerned with movement of goods and services within an urban area, then one may conjecture that both business establishments and residential population contribute positively to truck trip generation. However, within the context of a statewide freight flow analysis where the freight flows mostly raw industrial goods, residential population is not likely to attract freight trips. Indeed, many industrial sites are located in zip codes with minimal residential population and attract large amounts of freight flow. Thus, it appears that this finding may have some merit in the statewide modeling context. This finding also lends credence to the approach taken by many states and urban areas that try to attract "jobs" to their area to promote economic activity. The notion is that people will then come to where the "jobs" are located. Previous research using the Transearch database has reported similar findings (e.g., Brogan *et al.,* 2001).

6. CONCLUSIONS

Freight transportation model development is now a critical component of the overall transportation planning process as agencies consider mobility strategies for enhancing the safety and efficiency of freight transportation. This paper was aimed at developing a simple but practical modeling framework for directly estimating freight transportation flows by commodity between origin-destination pairs. The model development is based on a comprehensive integrated database that includes population, employment, and freight flow data for Florida. The only missing component in the database is modal level of service attributes; the process of merging modal level of service attributes is currently ongoing and will result in further enhancement of the model presented in this paper.

The modeling methodology consists of a structural equations model that can accommodate multiple dependent variables simultaneously. This structural equations model can be applied to all origin-destination zip code pairs in a region. In this model system, explanatory variables representing origin and destination population and employment characteristics and impedance (distance between the origin-destination pair) are included. The model is found to offer statistically valid indications and plausible interpretations suggesting that these models may be suitable for application in freight flow forecasting.

The model can be further enhanced by including additional explanatory and dependent variables. Inclusion of detailed modal level-of-service variables, policy and regulatory variables, and commodity characteristics would enhance both the sensitivity and the explanatory power of the model. Also, the model presented in this paper includes only shipments by truck and rail due to the very limited commodity flows by water and air in Florida. Air and water are important modes of freight transportation that should be considered in future model development efforts.

Undoubtedly, further research is important in the area of freight transportation demand modeling. The development of freight databases and collection of freight movement data continues to be a challenge for model development and estimation. There is always a degree of uncertainty regarding the coverage of the database with respect to geography, commodity groups, and modes and regarding the accuracy of the data as one goes to greater levels of spatial detail. The development of freight transportation models is making great strides, but there is some question as to how transferable these models are between geographic contexts and between geographic scales within the same context. How applicable is it to use a model system estimated at the zip code level at another level of aggregation such as census tract or traffic analysis zone? Also, Florida is a state dominated by the citrus industry, tourism industry, and a housing and office building construction boom fueled by strong population and employment growth. How transferable is a model that is estimated based on Florida freight flow data. Research into these issues will greatly enhance the ability to develop freight transportation models and estimate freight flows accurately while analyzing the effects of alternative freight mobility strategies and policies.

ACKNOWLEDGEMENTS

Funding and freight databases provided by the Florida Department of Transportation for undertaking this research are gratefully acknowledged. The paper benefited greatly from constructive comments provided by two anonymous reviewers. The author is responsible for all errors and omissions.

REFERENCES

Allen, B.W. (1977) The demand for freight transportation: a micro approach. *Transportation Research*, Vol. 11, No. 1, 9-14.

Arbuckle, J.L. (2000) *AMOS Users Guide Version 4.01*. SmallWaters Corporation, Chicago.

Baumol, W.J. and Vinod, H.D. (1970) An inventory theoretic model of freight transport demand. *Management Science*, Vol. 16, No. 7, 413-421.

Bollen, K.A. (1989) *Structural Equations with Latent Variables*. Wiley, New York.

Brogan, J., Stephen, C.B., and Demetsky, M.J. (2001) Application of a statewide intermodal freight planning methodology. Research Report 02-R5, Virginia Transportation Research Council, Virginia.

Bronzini, M.S. (1980) Evolution of a multimodal freight transportation network model. *Proceedings of the Transportation Research Forum*, Vol. 21, No. 1, pp. 475-485.

Browne, M.W. (1984) Asymptotically distribution free methods for the analysis of covariance structures. *British Journal of Mathematical and Statistical Psychology*, Vol. 37, 62-83.

U.S. Census Bureau (2005) Website http://www.census.gov.

Cambridge Systematics, Inc. (1997) A guidebook for forecasting freight transportation demand. NCHRP Report 388, National Cooperative Highway Research Program, Transportation Research Board, National Research Council, Washington, D.C.

Cambridge Systematics, Inc., COMSIS Corporation, and University of Wisconsin-Milwaukee (1997) Quick response freight manual. Final Report DOT-T-97-10, U.S. Department of Transportation, Washington, D.C.

Crainic, T. (1987) Operations research models of intercity freight transportation: the current state and future research issues. *Logistics and Transportation Review*, Vol. 23, No. 2, 189-206.

Friesz, T.L. (2000) Strategic Freight Network Planning Models. In D.A. Hensher and K.J. Button (eds.) *Handbook of Transport Modeling*, Elsevier Science Publishers, B.V., The Netherlands, pp. 527-537.

Garrido, R.A. and Mahmassani, H.S. (2000) Forecasting freight transportation demand with the space-time multinomial probit model. *Transportation Research*, Vol. 34B, No. 5, 403-418.

Harker, P.T. (1985) The state of the art in the predictive analysis of freight transportation systems. *Transport Reviews*, Vol. 5, No. 3, 44-62.

Harker, P.T. and Friesz, T.L. (1986) Prediction of intercity freight flows. Part I: Theory and Part II: Mathematical Formulations. *Transportation Research*, Vol. 20B, No. 2, 139-153, 155-174.

Holguin-Veras, J. and Thorson, E. (2000) Trip length distributions in commodity-based and trip-based freight demand modeling. *Transportation Research Record, Journal of the Transportation Research Board*, No. 1707, TRB, National Research Council, Washington, D.C., 37-48.

InfoUSA.com, Inc. (2005) Website http://www.infousa.com.

List, G.F. and Turnquist, M.A. (1994) Estimating truck travel patterns in urban areas. *Transportation Research Record*, No. 1430, TRB, National Research Council, Washington, D.C., 87-93.

Ogden, K.W. (1991) Truck movement and access in urban areas. *ASCE Journal of Transportation Engineering*, Vol. 117, No. 1, 71-90.

Pendyala, R.M. (2004) Modeling Origin-Destination Freight Flows in Florida by Mode and Commodity. Research Report, Department of Civil and Environmental Engineering, University of South Florida, Tampa.

Pendyala, R.M., Shankar, V.N., and McCullough, R.G. (2000) Freight travel demand modeling: synthesis of approaches and development of a framework. In *Transportation Research Record, Journal of the Transportation Research Board*, No. 1725, TRB, National Research Council, Washington, D.C., 9-16.

Reebie Associates (2005) Website http://www.reebie.com/images/transearch.asp.

Regan, A.C. and Garrido, R.A. (2001) Modeling Freight Demand and Shipper Behavior: State of the Art, Future Directions. In D.A. Hensher (ed.) *Travel Behaviour Research: The Leading Edge*, Elsevier Science Ltd., Oxford, UK, pp. 185-216.

Shankar, V.N. and Pendyala, R.M. (2001) Freight Travel Demand Modelling: Econometric Issues in Multi-level Approaches. In D.A. Hensher (ed.) *The Leading Edge of Travel Behaviour Research*, Elsevier Science Publishers, B.V., The Netherlands, pp. 629-644.

Slavin, H.L. (1976) Demand for urban goods vehicle trips. *Transportation Research Record* No. 591, TRB, National Research Council, Washington, D.C., 32-38.

Souleyrette, R., Maze, T.H., Strauss, T., Preissig, D., and Smadi, A.G. (1998) Freight Planning Typology. *Transportation Research Record*, No. 1613, TRB, National Research Council, Washington, D.C., 12-19.

Southworth, F. and Peterson, B.E. (2000) International and intermodal freight network modeling. *Transportation Research*, Vol. 8C, No. 1, 147-166.

Winston, C. (1983) The demand for freight transportation: models and applications. *Transportation Research*, Vol. 17A, No. 6, 419-427.