SHORT-MEDIUM TERM PARAMETER STABILITY IN A NATIONAL FREIGHT DEMAND MODEL: AN EMPIRICAL INVESTIGATION

José Holguín-Veras[1], Iván Sarmiento[2], Carlos González[3], Ellen Thorson[4], Iván Sánchez[5]

Abstract

This paper conducts an empirical investigation on the temporal stability of parameters of freight demand models in the short-medium term. The analyses are based on seven national freight origin-destination samples conducted by the Colombian government during the years 1999 to 2005. The paper studies the stability of the parameters of freight generation, freight distribution, and empty trip models. Typical formulations of these models were calibrated using the cross-sectional data corresponding to each year. Then, to identify time-dependent effects models were estimated using a panel formulation with fixed effects. The results indicate the presence of statistically significant time-dependent effects on all freight generation models (production and attraction), as well as on the freight distribution model estimated with loaded vehicle trips. In contrast, the parameters of the freight distribution models based on commodity flows and the ones for the empty trips were found to be stable overtime. The reason may be related to the fact that the commodity flows reflect production-consumption patterns that are much slower to change overtime than vehicle trips that are the result of short term logistic decisions on the part of the carriers. The stability of the parameters of empty trips is also related to the stability of production-consumption patterns. This is because the percentage of empty trips—which is related to the parameters of the models—is directly determined by the degree of asymmetry of the commodity flow matrix, as the more asymmetric the matrix is the larger the percentage of empty trips. Since the parameters of the empty trip models are related to the percentage of empty trips, a stable percentage of empty trips lead to stable parameters.

---

[1] Professor Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute 110 8th Street, Room JEC 4030, Troy, NY 12180 USA

[2] Associate Professor Escuela Ingeniería Civil, Universidad Nacional de Colombia Sede Medellín Bloque M1-223.

[3] Assistant Professor Universidad de Antioquia. Medellin, Colombia. Ph.D. Student Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute 110 8th Street, Room JEC 4037, Troy, NY 12180 USA

[4] University Transportation Research Center, The City College of New York, Marshak Hall Suite 910A 138th Street & Convent Avenue, New York, NY 10031, USA.

[5] Ph.D. Student Department of Civil and Environmental Engineering, Rensselaer Polytechnic Institute 110 8th Street, Room JEC 4037, Troy, NY 12180 USA

I. Introduction

Analytical transportation planning relies on the use of demand and supply (network) models to predict future conditions, and the impact of different projects and programs on the transportation system. In most cases, the models used in these activities are calibrated using cross-sectional data from a calibration (base) year. Once the models have been successfully calibrated, their independent variables are forecasted for different scenarios, and used as an input to freight demand models to obtain estimates of future demand. A fundamental assumption is that the parameters of the different models used are stable overtime. This assumption is important because evidence of it provides an indication of the validity of a model (Gunn et al., 1985), as a model that is not stable over time is likely to produce inaccurate predictions (Ortúzar and Willumsen, 2001).

In the overwhelming number of cases, temporal parameter stability is implicitly invoked by assuming that the models' parameters do not change over time. The practical reasons are obvious as, more often than not, there are no data to study how the parameters evolve over time. This is not to say that there are no concerns about assuming parameter stability, as there are many indications that this assumption is problematic. First, since in most cases cross sectional data are used for model development and calibration, should the economic conditions be such that they change the structure of the freight flows captured in the data, the parameters of the models will be affected as well. Second, the ever changing nature of the world economy means that the economic linkages between different economic sectors are constantly being created, transformed, and sometimes eliminated. In this context, emerging economics sectors are likely to lead to freight flows not necessarily captured—and in some cases not even dreamed of—by the data collection efforts used for calibration purposes. For instance, the globalization of the world economy led to major increases in the volumes and the distances at which freight is transported. As a result, it is unlikely that freight models calibrated with data collected in the 1980s could have captured, and much less predicted, the freight flow patterns produced by globalization. Third, there is the role played by political and social drivers. Events such as the collapse of the Soviet Union altered the patterns of commerce in Europe. Societal trends, such as the increasing level of awareness of environmental concerns, are inducing companies to change their distribution and delivery patterns, out of a desire to be perceived as a responsible corporate citizen that cares for the environment. In many instances, companies have implemented changes in their distribution patterns that, although are more sustainable in the long term, are more costly. As in the previous case, the net effect is to change the geographical patterns of freight flows and their relation to transportation costs, which impact the parameters of freight demand models that capture the sensitivity to transportation costs.

The key implication of all of this is that the assumption of temporal parameter stability—which is at the heart of transportation demand modeling research and practice—is not likely to hold. Surprisingly enough, not much research has been reported in the literature concerning the subject of parameter stability, and even less in the area of freight demand. As a reflection of this, the literature review only found a handful of papers.

For the most part, the papers found have focused on transferability issues (McCarthy, 1982; Gunn et al., 1985; Tretvik and Widlert, 1998). However, during the 1980s a body of literature emerged with empirical evidence about the stability (or in most cases, lack of it) of parameters of disaggregate travel demand models, across space, cultures and time (Ortúzar and Willumsen, 2001). For example, McCarthy (1982) used a multinomial logit mode choice model to examine the validity of model specification, and the temporal stability of its parameters. The author found that the specific variables remain stable over the short run period, and that the parameters used are appropriate for immediate and short term forecasting. Other papers have made indirect mention of parameter stability issues (Robusté, 1994).

The main objective of this paper is to contribute to the study of parameter stability of freight demand models via a systematic study of seven national freight origin-destination (OD) matrices collected by the Colombia's Ministry of Transportation during the 1999 to 2005 time period. Since efforts were made to obtain older OD data but were not successful, the emphasis of the paper is on short-medium term changes as this is what is permitted by the data available. In the future, once additional data sets come in line, it will be entirely possible to re-examine the analyses made here with the perspective that provides the long term.

The main focus is on the various aspects concerning freight demand at the aggregate level. This includes generation, distribution, and empty trips. The OD data are used to estimate basic cross-sectional demand models for each of the various years for which data are available. A panel formulation with fixed effects was used to estimate the models and identify time-dependent effects.

This paper has five sections in addition to this introduction. Section 2 provides the reader with a brief description of the data used in the paper. Section 3, 4, and 5 discusses the results of parameter stability analyses of freight generation, distribution, and empty trip models respectively. Finally, Section 6 summarizes the key findings from the research.

II. Descriptions of the Data

The Freight Origin–Destination Survey (FODS) is a data collection program initiated by Colombia's Ministry of Transportation about 20 years ago. Initially, the FODS was conducted by the National Institute of Transport (Instituto del Transporte, INTRA) and, more recently, by the National Roads Institute (Instituto Nacional de Vías, INVIAS). The FODS is conducted at about 70 survey stations, where interviews are conducted to collect data about origins and destinations, commodity type, vehicles used, type of the container used, among others.

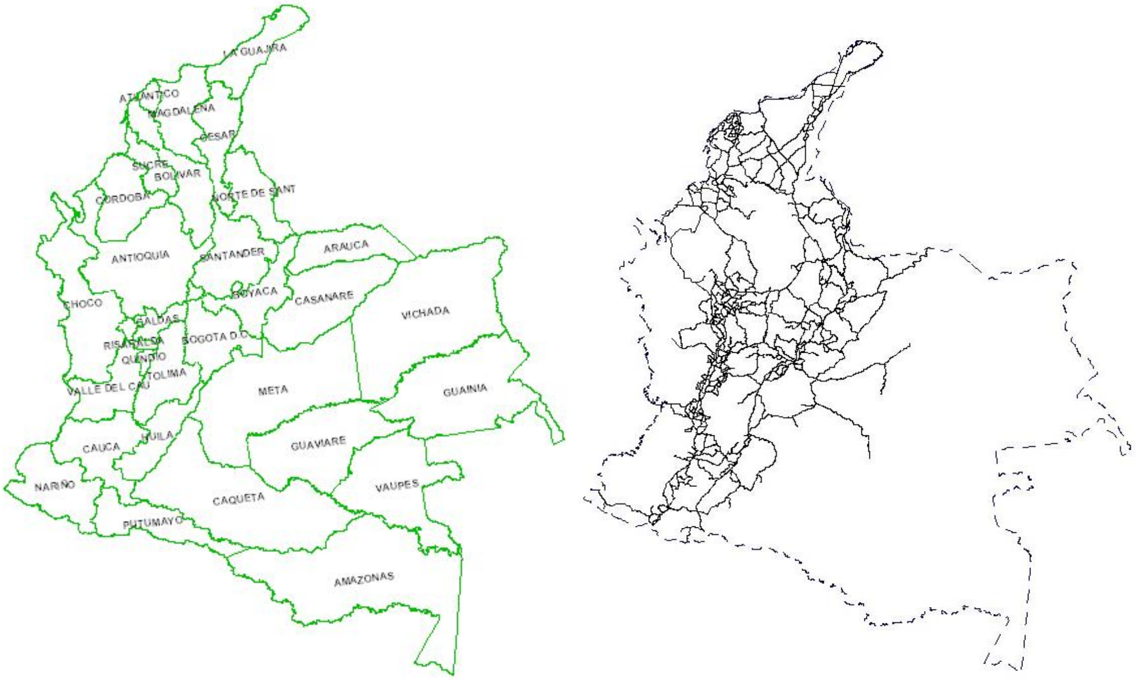
The FODS provides a comprehensive picture of national freight flows in Colombia. It provides the only available source of data for the highway modes that carry about 70% of the tonnage of freight transported, about the same mode split reported by the Commodity Flow Survey (CFS) in the United States. However, while the CFS is a shipper based survey, the FODS is a roadside survey that targets the carriers. The FODS is a sizable data collection program with more than 130,000 surveys every year. It provides useful data for freight transportation planning, though it has not been fully exploited for research purposes.

The FODS collect data for five consecutive days in a representative week of the year, between 6:00 AM and 6:00 PM. The target population includes all trucks with a capacity larger or equal to two metric tons, that pass the screenlines within the survey period. To avoid double counting a ticket is given to drivers who have been interviewed so that they do are not subsequently interviewed at another station along the trip.

At the finest level of geographic detail, the survey data are geocoded at the municipal level, which leads to OD matrices with 1,100x1,100 cells. The analyses in this paper are based on an aggregation to 36 transportation analysis zones (corresponding to political departments in Colombia) because in this format the coverage is more complete (electronic versions of the original files at a finer level of detail were only found for four years). The zoning system used includes 32 internal zones (departments) and 4 external zones (Ecuador, Venezuela, Peru and Panama). The zoning system and the corresponding highway network are shown in Figure 1

(Universidad Nacional de Colombia Sede Medellin et al., 2008). The total highway length used in the model is 27,469 Km.

Figure 1: Zoning System and Highway Network (Colombia)



III. Freight Generation

Freight generation is concerned with the estimation of the amount of freight produced and attracted by either individual establishments or zones. As in almost all other components of freight demand models, the amount of publications discussing freight generation is small. It suffices to say that only four out of the 1,500 pages of the ITE's Trip Generation Manual are dealing with freight (Institute of Transportation Engineers, 1991). The generation of freight is a complex topic. Part of its complexity stems from the fact that it is determined by the internal practices and operational procedures of the establishment that produces/attracts the freight. A restaurant with limited storage capacity, for instance relative to its size, is bound to attract more delivery trips than a restaurant with a larger storage space in equality of conditions, simply because of the more frequent deliveries required by the constrained storage space. Adding to the complexity, freight could be measured in multiple ways (e.g., weight, vehicle-trips, deliveries); and studied at the zonal or the establishment level. Regarding the level of aggregation, there seems to be consensus that establishment level models have a better chance of capturing the underlying dynamics of the process, though the data are costlier to collect. The paper focuses on zone level models, which is typical of national freight demand modeling projects.

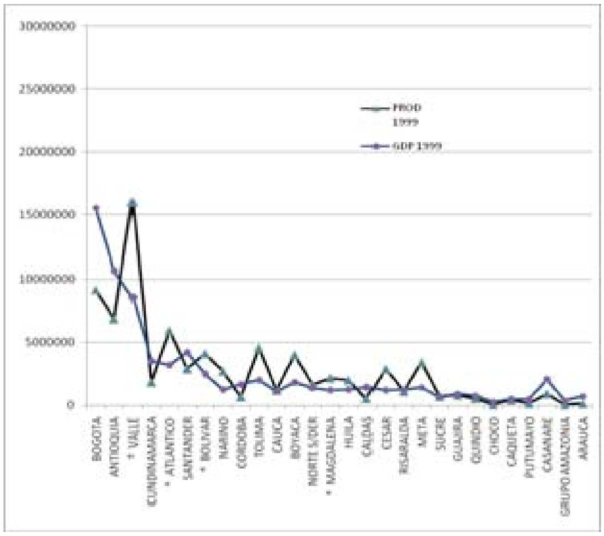
This chapter focuses on the temporal stability of parameters related to freight generation (i.e., production and attraction). The analyses use department level estimates of Gross Domestic Product (GDP) and population as the independent variables that explain production, and attraction of freight. Although other independent variables could be used, focusing on these two is important because they are frequently used to explain freight demand. The GDP and population were obtained from the Colombia's Department of National Statistics (DANE, 2009).

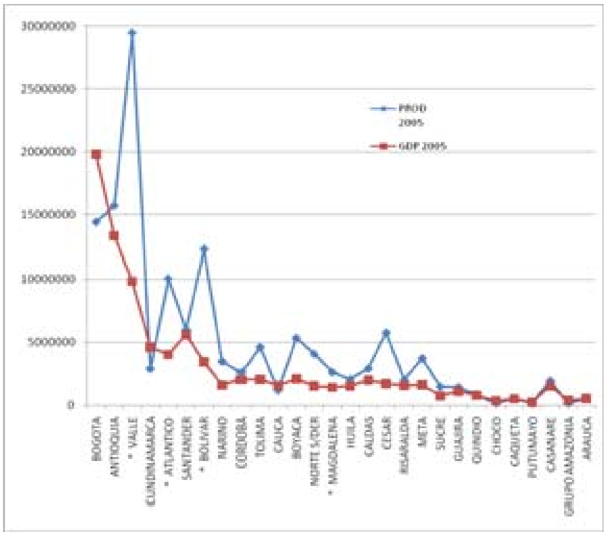
3.1 Freight Production

The data used to study the relationship between freight production and GDP spans over for seven years (1999-2005). There are good reasons to use GDP, or any other indicator of economic output, as the freight flows are nothing more than the physical representation of the trade patterns captured in these indicators. Two different sets of econometric models are estimated. The first one consists of models with GDP as the only independent variable, while the second set of models considers the effect of local ports.

To start, it is convenient to take a look at the relationship between the total tons produced by different zones and their GDP. Figure 2 shows the results for 1999 and 2005. The figure shows that in general, both GPD and freight production track each other fairly well and that, overtime more cargo is being transported for the same level of GDP. Figure 2 also shows a number of spikes in freight production that correspond to ports. Jurisdictions in which a port is located have been marked with an asterisk.

Figure 2: Tons Produced vs. GDP (1999 and 2005)





The first model estimated follows equation (1):

$$F_P = \beta_1 GDP \qquad (1)$$

Where: F_P, is the production function in tons; and *GDP* is the Gross Domestic Product of each department in millions of Colombian pesos (about US$1,600).

The second model is a power function (equation 2), which has the advantage that the exponent is the elasticity of production with respect to *GDP*. The model was originally estimated with an intercept, which was later removed because it was not significant. The results for the different cross sections are shown in Table 1.

$$F_P = \beta_0 GDP^{\beta_1} \qquad (2)$$

Table 1: Parameters of Freight Production Models under cross sections

	Linear Model				Power Model			
Year	Parameter β1	t	Adjusted R^2	F	Parameter β1	t	Adjusted R^2	F
1999	0.8763	4.18	0.73	17.46	0.9785	66.16	0.99	4,377.76
2000	0.8125	3.96	0.75	15.67	0.9854	98.42	1.00	9,686.92
2001	1.0764	3.91	0.76	15.28	1.0095	107.79	1.00	11,617.80
2002	0.8493	4.16	0.73	17.28	0.9927	93.07	1.00	8,662.40
2003	0.9087	3.23	0.65	10.42	1.0092	97.65	1.00	9,535.90
2004	1.0053	3.39	0.65	11.46	1.0185	105.43	1.00	11,114.90
2005	1.2574	3.61	0.70	13.03	1.0227	132.60	1.00	17,582.70

Table 1 shows that the parameter β_1 in the linear model varies between 0.81 and 1.26, indicating that on average one million of Colombian Pesos of GDP (about US$1,600) are needed to produce a ton of freight. The results also show that the parameter β_1 in the linear model increases overtime, and, in 2005 is 1.4 times the parameter in 1999. Although the parameters of the power function seem more stable than the ones in the linear model, the reader should keep in mind that a small change in the parameter value could lead to major changes in the estimates produced. As shown, the parameters were found to be statistically significant time-dependent.

In order to analyze the stability of the parameters a new model is estimated by regressing the freight production on the GDP with fixed time effects. The parameter β_i for the year binary variables representing each year can be thought as the effect on the freight production of year *i*.

Table 2: Time fixed effects for Freight Production Models

Variable	Linear Model		Power Model	
	Parameter βι	t-value	Parameter βι	t-value
GDP, ln(GDP)	0.85	-7.93	0.985	-68.01
2001	1,221,507.80	-2.44		
2003	1,079,577.10	-1.99		
2004	1,577,689.50	-2.55		
2005	2,189,346.00	-2.93	0.534	-2.23
Time dependent ?	Yes	Adjusted R2=0.711 F=39.01	Yes	Adjusted R2=0.996 F=9872.7

In the estimated models with time fixed effects in Table 2 the intercepts were not significant and were therefore not considered. In the linear model the time parameters were significant for years 2001, 2003, 2004 and 2005 at the 5% significance level. For the power model, the time parameter for the year 2005 was found significant at the 5% significance level.

An alternative model takes into account the additional freight produced by a port. This was accomplished with the introduction of a binary variable (δ_P) into the model. The zones with dedicated ports for oil or coal exports were not included in the analyses because they typically use either pipelines or rail to transport the cargo, and do not produce general cargo. The formulations are shown in equation (3) for the linear model, and in equations (4) and (5) for the power model, and the results in Table 3.

$$F_P = (\beta_1 + \beta_2\delta_P)(GDP) = \beta_1(GDP) + \beta_2\delta_P(GDP) \tag{3}$$

The power model is:

$$F_P = \beta_0 GDP^{\beta_1 + \beta_2 \delta_P} \tag{4}$$

$$\ln(F_P) = \beta_0 + (\beta_1 + \beta_2 \delta_P)\ln(GDP) \tag{5}$$

Table 3: Parameters of Freight Production Models Considering Ports under cross sections

Linear Model							Power Model					
Year	Parameter ß1	t1	Parameter ß2	t2	Adj. R^2	F	Parameter ß1	t1	Parameter ß2	t2	Adj. R^2	*F*
1999	0.6630	13.16	1.1793	21.79	0.93	4,398.20	0.9680	57.48	0.0709	4.20	0.99	251,675.70
2000	0.6270	7.41	1.0641	8.35	0.93	185.10	0.9780	86.01	0.0472	3.91	1.00	35,162.60
2001	0.8380	5.45	1.3363	8.49	0.92	2,103.50	1.0019	95.46	0.0487	4.50	1.00	79,687.90
2002	0.6543	8.74	1.1517	11.61	0.91	423.20	0.9840	82.45	0.0552	4.15	1.00	19,084.00
2003	0.6541	5.24	1.5402	11.25	0.89	768.30	1.0015	85.51	0.0487	3.94	1.00	39,262.90
2004	0.7240	6.41	1.7238	15.23	0.90	4,237.20	1.0104	93.83	0.0508	4.48	1.00	48,653.70
2005	0.9310	6.46	2.0384	11.98	0.94	557.90	1.0148	125.51	0.0501	4.67	1.00	19,238.90

It is important to note that in the linear model the parameter β_1 in 2005 is 1.4 times the parameter in 1999, while the combined parameter ($\beta_1 + \beta_2\ \delta_P$) for zones with a port, in 2005 is 1.6 times the corresponding to 1999. This clearly suggests that in the linear model the freight generation in port zones is growing more rapidly than in the rest of the country. Similarly, the parameter β_1 of the power model in 2005 is 1.045 times the parameter in 1999, while the parameter ($\beta_1 + \beta_2 \delta_P$) in 2005 is 1.025 times the parameter corresponding to 1999. The results show that the parameter of *GDP* in the linear model grew slower than the one for δ_P *(GDP)*. The same trend is observed in the power model.

As well as for the freight production model of Table 2 an analysis of the stability of the parameters was realized considering the effects of ports. The models in Table 4 are estimated considering fixed time effects and the effects of having a port in the zone.

Table 4: Time fixed effects for Freight Production Models Considering Ports

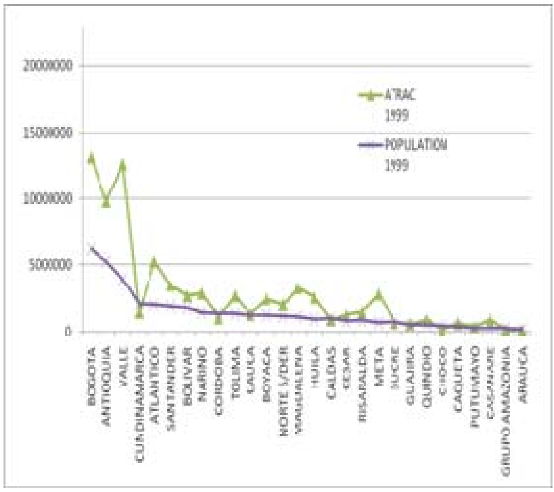
Variable	Linear Model		Power Model	
	Parameter $\beta 1$	t-value	Parameter $\beta 1$	t-value
GDP, ln(GDP)	0.64	-14.54	1.18	-15.67
Port*GDP, Port*ln(GDP)	1.43	-11.51	0.04	-7.56
2001	977,973.20	-3.66		
2003	827,928.70	-2.89		
2004	1,304,704.80	-4.00	0.58	-2.36
2005	1,914,645.20	-4.89	0.62	-2.78
Time dependent ?	Yes	Adjusted R2=0.92 F=126	Yes	Adjusted R2=0.736 F=83.86

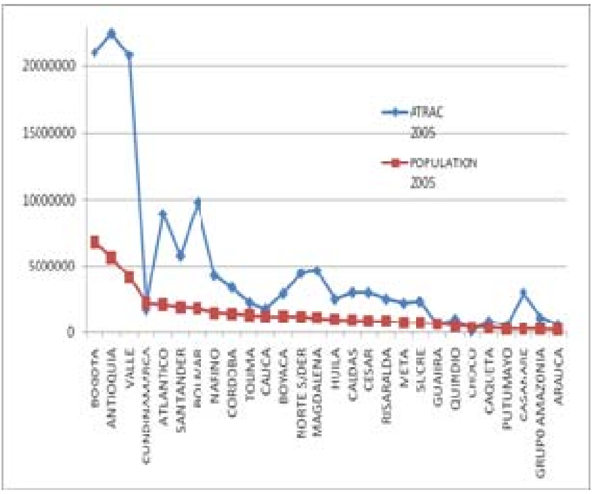
In the estimated models with time fixed effects in Table 4 the intercepts were not significant and were therefore not considered. In the linear model the time parameters were significant for years 2001, 2003, 2004 and 2005 at the 5% significance level. For the power model, the time parameter for the years 2004 and 2005 were found significant at the 5% significance level. In both linear models, whether considering or not the extra effect of having a port on freight production, the time parameter for years 2001, 2003, 2004 and 2005 is significant. In the case of power models the time parameter for year 2005 is significant, whether considering or not the effect of ports.

3.2 Attraction Models

In this subsection the relation between total tons attracted by the different zones and their corresponding populations (*P*) is studied. The data are shown in Figure 3 for 1999 and 2005. As in the case of freight production, the data show that the amount of freight attracted by the population centers has increased over time. For instance, while the city of Bogota attracted about 13 million tons in 1999, it attracted more than 20 million tons in 2005.

Figure 3: Tons attracted vs. Population (1999 and 2005)





As in the previous case, econometric modeling was undertaken using the cross-sections for each year. The linear model used is shown as in equation (6), and the corresponding power function in equation (7). The statistical results are summarized in Table 5.

$$F_A = \beta_1 P \tag{6}$$

Where F_A is the production function in tons, and P is the population of each zone:

$$F_A = \beta_0 P^{\beta_1} \tag{7}$$

Table 5: Parameters of Freight Attraction Models estimated under cross sections

	Linear Model				Power Model			
Year	Parameter β1	t	Adjusted R^2	F	Parameter β1	t	Adjusted R^2	F
1999	2.0948	11.78	0.92	138.80	1.0167	66.02	0.96	15,270.55
2000	1.9332	14.83	0.94	220.00	1.0253	141.11	0.96	18,499.24
2001	2.6450	14.56	0.94	212.10	1.0442	133.34	0.96	16,798.12
2002	2.0898	13.45	0.94	181.00	1.0331	101.70	0.96	9,064.24
2003	2.4910	12.04	0.90	145.10	1.0299	79.84	0.96	9,232.67
2004	2.9191	14.04	0.90	197.20	1.0499	114.10	0.96	12,148.76
2005	3.4782	11.36	0.92	129.00	1.0670	117.34	0.96	12,857.68

The parameters of the linear models vary between 1.93 and 3.48 with an average of about 2.50. This means that during this time period, one individual attracts about 2.5 tons of freight per year. This value is still small compared with developed countries like US, where the value is 38 tons per capita in all modes, and 26 tons in highways (69%) in 2002 (Bureau of Transportation

Statistics, 2009). Both models (linear and power) also show that the amount of cargo attracted by the population is increasing overtime.

As in previous cases, the stability of the parameters is estimated by regressing the freight attraction on the population with fixed time effects.

Table 6: Time fixed effects for Freight Attraction Models

Variable	Linear Model		Power Model	
	Parameter β1	t-value	Parameter β1	t-value
Intercept	-1,029,410.00	-3.63	-5.70	-3.27
Population, ln(Pop)	2.70	-19.36	1.43	-12.28
2004	1,121,962.30	-2.43		
2005	1,816,794.20	-3.26	0.68	-2.76
Time dependent ?	Yes	Adjusted R2=0.839 F=60.03	Yes	Adjusted R2=0.729 F=32.52

In the estimated models with time fixed effects in Table 6 the intercepts were significant and were therefore considered. In the linear model the time parameters were significant for years 2004 and 2005 at the 5% significance level. For the power model, the time parameter for the year 2005 was found significant at the 5% significance level.

As in the production case, an alternative model takes into account the additional freight produced by a port. Equation (8) shows a linear model with a binary variable to capture the port effect, while equation (9) shows the equivalent model with a power function. The statistical results are summarized in Table 5.

$$F_A = \beta_1 P + \beta_2 \delta_P P \tag{8}$$

$$F_A = \beta_0 P^{\beta_1 + \beta_2 \delta_P} \tag{9}$$

Table 7: Parameters of Freight Attraction Models Considering Ports under cross sections

Linear Model							Power Model					
Year	Parameter β1	t1	Parameter β2	t2	Adjusted R^2	F	Parameter β1	t1	Parameter β2	t2	Adjusted R^2	*F*
1999	1.8860	15.46	0.9950	3.07	0.95	165.60	1.0080	56.15	0.0545	2.62	0.99	6,753.80
2000	1.7860	17.39	0.7050	3.41	0.96	247.40	1.0200	123.66	0.0339	3.56	0.99	32,274.00
2001	2.3950	14.15	1.2060	6.42	0.97	#######	1.0364	122.68	0.0498	5.84	0.99	##########
2002	1.8860	14.48	0.9900	7.29	0.97	#######	1.0263	87.80	0.0431	3.55	0.99	55,215.40
2003	2.1720	14.91	1.5570	6.41	0.95	295.00	1.0180	72.29	0.0775	4.25	0.99	7,099.30
2004	2.5910	13.63	1.6070	4.93	0.95	218.50	1.0394	112.10	0.0666	4.11	0.99	9,754.10
2005	3.1324	10.68	1.7041	4.92	0.95	400.10	1.0601	103.47	0.0445	4.13	0.99	59,991.10

In the power model it is easily seen that the parameter β_2 is significant at the 5% level for the years 1999, 2000, 2001, 2002, 2003, 2004 and 2005. Therefore there is an additional effect on freight attraction when taking in account ports presence. Having a port produce an extra effect in average of 5% in the freight attraction. In the linear model the parameter β_1 in 2005 is 1.66 times the parameter in 1999, and the combined parameter ($\beta_1 + \beta_2 \ \delta_P$) for zones with a port, in 2005 is 1.68 times the corresponding to 1999. As far as the power model is concerned, the parameter β_1 in 2005 is 1.052 times the parameter in 1999, and the combined parameter ($\beta_1 + \beta_2 \ \delta_P$) for zones with a port, in 2005 is 1.04 times the corresponding to 1999. From the study of the interaction terms and their variation from 1999 to 2005, it is possible to conclude that having a Port in the zone of attraction does have an extra effect of its corresponding population on the amount of freight attracted.

A new model considering ports is estimated with time fixed effects.

Table 8: Time fixed effects for Freight Attraction Models Considering Ports

Variable	Linear Model		Power Model	
	Parameter β1	t-value	Parameter β1	t-value
Intercept	-1,075,483.90	-4.18	-4.857	-2.60
Population, ln(Pop)	2.45	-18.42	1.362	-10.85
Port*Pop,	1.27	-6.78	0.0345	-4.89
2001	663,721.10	-2.46		
2004	1,122,842.70	-3.14		
2005	1,817,665.10	-3.90	0.689	-2.82
Time dependent ?	Yes	Adjusted R2=0.92 F=126	Yes	Adjusted R2=0.736 F=83.86

In the estimated models with time fixed effects in Table 8 the intercepts were significant at the 5% significance level and were therefore considered. In the linear model the time parameters were significant for years 2001, 2004 and 2005 at the 5% significance level. For the power model, the time parameter for the year 2005 was found significant at the 5% significance level.

The analyses in this section clearly indicate that, regardless of the formulation, the parameters of freight attraction and production have increased overtime. The statistical models that express the parameter values as a function of the time index indicate that, in almost all cases, the parameters are time-dependent. As an illustration, the parameters of the linear model of freight production increased 40% between 1999 and 2005, while the parameter of the linear model of freight attraction increased 73% in the same time period. In essence, these results indicate that, overtime, more freight is being produced and attracted by a unit of GDP and population, respectively.

IV. Freight Distribution

The second group of models to be studied is the one that focuses on the estimation of freight distribution patterns. This is probably one of the processes in which freight demand modeling is the weakest, as it is the one where the mismatch between model assumptions and reality is the largest. Although the reasons are many and cannot be fully enumerated here, it is important to discuss the key ones.

The bulk of freight demand modeling applications relies on the use of distribution models originally designed for passenger demand modeling. The main focus of these models, e.g., gravity, is on modeling the flows between and origin *i* and a destination *j* as a function of the attributes of *i* and *j*, and the corresponding travel impedance. This approach is acceptable in passenger transportation as, in most cases, as the frequency of long trip chains is small and it could be argued that the assumption is appropriate. However, in freight transportation—where long tours are the norm and not the exception—this assumption could be problematic. In Denver, for instance, the number of stops per tour is 5.6 (Holguín-Veras and Patil, 2005). As result of the multiplicity of individual trips in a long tour, the physical origins and destinations of the individual trips tend not to match the production and consumption relations. When this happens, the assumption made by most distribution models, i.e., that the individual trips could be explained by the attributes of *i* and *j*, breaks down and the use of traditional distribution models is called into question. This is likely to be the case of urban freight. However, there are cases in which the mismatch between production-consumption (PC) and origin-destinations (OD) is less critical. One of such cases, is in intercity freight transportation in developing countries like Colombia. In these cases, the number of long tours is much less than in urban freight. This leads to a situation in which using trip-based distribution models is a reasonable decision.

In this section, a set of doubly-constrained gravity models are estimated for both commodity flows in tons, and loaded trips in vehicle units. These models account only for the loaded trips in the network, as the corresponding empty trips are analyzed in Section V. The mathematical form of the model is shown in Equation 10. The models were estimated for six years (2000-2005). Distance was used as the impedance variable because costs were no available for all years. Three different impedance functions were used: power, exponential, and gamma (which did not converge in all cases). The values of the parameters found are shown in Table 9.

$$T_{ij} = a_i * O_i * b_j * D_j * f(c_{ij}) \tag{10}$$

Where:

T_{ij} : Loaded trips vehicles / commodity flows from i to j

O_i= Loaded trips vehicles / commodity flows from origin i

D_j= Loaded trips vehicles / commodity flows to destination j

a_i= balancing factor for origin i

b_j= balancing factor for destination j

$f(c_{ij})$= impedance function

Table 9: Parameters of the Impedance Functions

Year	Loaded Trips Vehicles		Commodity Flows	
	Power (b)	Exponential (c)	Power (b)	Exponential (c)
2000	1.9121	0.0027523	1.2748	0.0016345
2001	1.6444	0.0023046	1.0550	0.0012418
2002	1.5916	0.0022399	0.9977	0.0011591
2003	1.7623	0.0024242	1.2032	0.0014859
2004	1.6317	0.0022692	1.0879	0.0013287
2005	1.5286	0.0019833	0.9535	0.0009933
Time dependent ?	Yes	Yes	No	No

Table 9 shows that, in general, the parameters of the models estimated using the commodity flows are lower than those from vehicle-trip models. As shown, the value of parameters of the power function varies between 1.53 (2005) and 1.91 (2000) in the case of the loaded vehicle trips model; and varies between 0.95 (in 2005) and 1.27 (2000) in the case of commodity flows model. In the case of the exponential function, the parameters varies between 0.0020 (2005) and 0.0028 (2000) in the loaded vehicle trips model, and between 0.0010 (2005) and 0.0016 (2000) in the case of commodity flows model.

The results show that the parameters of the impedance functions decline over time. It is worthy of note that since the analyses used distance as the impedance variable, the changes in the parameters are solely the product of changes in the demand (field observations indicate no major change in congestion during the period of analysis). These results have a direct impact in terms of prediction capability as a model calibrated with 2000 data will produce very different forecasts than a model calibrated in 2005 as the impedance effect did not remain constant.

However, the statistical tests of time dependence indicate a split situation. As shown, while the parameters of the models estimated using loaded vehicle trips were found to be time-dependent, the ones estimated with commodity flows were not. This obviously suggest a decoupling between vehicle-trips and commodity flows which makes perfect sense because the commodity flows reflect the production-consumption patterns, while the vehicle trips are a reflection of the logistical decisions made by the freight carriers. Since the production-consumption patterns are much slower to change than vehicle-trips, it should not be a surprised that the distribution patterns of the commodity flows are stable, while the one for vehicle trips are not. The decrease in the value of the parameters indicates that travel impedance is less important than before. To a certain extent, this result is not surprising as it is well known that economic globalization has led to increases in the both amount of freight transported, and the corresponding distance, which leads to a lowering of the value of the parameter of the distribution models.

V. Empty Trips

One of the most unique—and more frequently overlooked—aspects of freight transportation is the number of empty trips that it generates. The number of empty trips is so high that if air is considered a commodity, it would be the commodity most frequently transported. According to the Vehicle and Inventory Use Survey (U.S. Census Bureau, 2004), empty travel account for: 56% of the miles traveled by straight truck not pulling a trailer; 58% of the miles traveled by straight truck pulling a trailer, and 33% of the miles traveled by truck tractor (power-unit) pulling trailer(s). As a percentage of the number of trips made, empty travel typically account for about 20% of truck traffic in urban areas (Strauss-Wieder et al., 1989), and about 30-40% in intercity freight (Holguín-Veras and Thorson, 2003). These numbers clearly indicate their importance. More significant from the modeling point of view is that empty trip flows do not follow the pattern followed by the loaded trips as the empty trips tend to run counter to the commodity flows. As a result, trying to compensate for the empty trips by expanding the matrices of loaded trips, or not accounting for the empties, lead to major errors in the estimation of directional traffic (Holguín-Veras and Thorson, 2003).

The potential errors associated with not properly modeling empty trips could be made obvious with a simple example of a two zones (A and B) system, in which there is only a commodity flow of 100 tons from A to B and nothing from B to A. If the average payload is 20 tons, this would lead to a loaded truck traffic of 5 units/day from A to B. However, since after unloading the cargo the trucks have to return to the base in A, an additional flow of empty trucks of 5 units/day is generated, in this case from B to A. As the example indicate, not accounting for the empty traffic would underestimate total truck traffic (in this simplistic example by 100%) and, more importantly, it would incur in major errors in the estimation of the directional flows (which are the ones that determine capacity needs). It turns out that, as the reader could verify, it is impossible to obtain the correct answer of 5 units/day in both direction by simply playing with distribution model that generates the loaded trips (Holguín-Veras and Thorson, 2003).

The only way to properly account for the empty traffic is to use complementary empty trip models (Noortman and van Es, 1978; Hautzinger, 1984; Holguín-Veras and Thorson, 2003; 2003; Holguín-Veras et al., 2008). These models estimate the flows of empties from the commodity flow matrix with the use of simplifying assumptions of tour behavior. These models have been successfully incorporated in state of the art models in Sweden, Colombia, New York, among others. The general formulation for these models is (Holguín-Veras and Thorson, 2003) shown below. The reader should notice that the Noortman and Van Es' model is obtained by setting γ equal to zero in equation (11):

$$E\left(z_{ij}\right)=\frac{m_{ij}}{a_{ij}}+p\frac{m_{ji}}{a_{ji}}+\gamma\sum_{h\neq j}x_{hi}\left(P^{h}\left(j\right)P^{h}\left(E/j\right)\right) \tag{10}$$

Where:

a_{ij}= average payload (tons/trip) for loaded trips between origin i and destination j

p = probability of a zero order trip chain

γ, β = parameters to be determined empirically

d_{ij}= distance between origin i and destination j

x_{hi}= = m_{hi}/a = number of loaded trips from h to i

$P^{h}\left(j\right)$= probability that a vehicle that came from h to i chooses j as the next destination

$P^{h}\left(E/j\right)$ = probability that a vehicle following the tour h-i-j does not get cargo to j

$P^{h}\left(j\right)P^{h}\left(E/j\right)$ = probability that a vehicle traveling in h-i-j goes empty to j

$P^{h}(j)$ is a function of the attractiveness of zone j as a destination which can be assumed to be a function of the commodity flow from j to i, m_{ij} and the trip impedance. A number of different formulations could be obtained depending on the assumption made regarding $P^{h}(j)$. Equations (12) to (14) show the models used in the paper.

$$E(z_{ij})=\frac{m_{ij}}{a_{ij}}+px_{ji} \qquad \text{(NVE, Noortman and van Es')} \tag{11}$$

$$E(z_{ij})=\frac{m_{ij}}{a_{ij}}+p\frac{m_{ji}}{a_{ji}}+\gamma\sum_{h\neq j}x_{hi}\frac{m_{ij}}{\sum_{l}m_{il}}P(E/j) \qquad \text{(HVT1)} \tag{13}$$

$$E(z_{ij}) = \frac{m_{ij}}{a_{ij}} + p\frac{m_{ji}}{a_{ji}} + \gamma \sum_{h \neq j} x_{hi} \frac{m_{ij}\, e^{-\beta(d_{ij})}}{\sum_{l} m_{il}\, e^{-\beta(d_{il})}} P(E/j) \quad \text{(HVT2)} \qquad (12)$$

The analyses focus on the stability of parameters like the percentage of empty trips in the network, and the parameters of empty trip models. The results are shown in Table 10 and Table 11. As shown in Table 10, the percentages of empty trips are very stable over time. This is a consequence of the asymmetry of the commodity flow matrices that, as discussed before, is a very stable feature of the economic system.

Table 10: Percentages of empty trips

Year	Observed Empty Trips
2000	28.5%
2001	28.4%
2002	27.3%
2003	30.0%
2004	29.6%
2005	26.4%

The results in Table 11 show that the parameters of the empty trip models are stable. This is to be expected as previous research has already established that these parameters are related to the percentage of empty trips (Holguín-Veras and Thorson, 2003; Holguín-Veras et al., 2008). As a result, if the percent of empty trips is stable, the parameters of the models are likely to be stable too. The results also show the superiority of the empty trip models that use a first order trip chain representation. In all cases, the models HVT1 and HVT2 lead to lower estimation errors that are in between 11% to 27% lower than the ones produced by the Noortman and Van Es' model.

Table 11: Parameters of Empty Trip Models

Year	Model	p	γ	β	Error	% change in error	Observed Empty Trips
2000	NVE	0.4336	0.0000	0	16918076174	0.00%	28.5%
	HVT1	0.6064	-2.4035	0	14027243238	-17.09%	
	HVT2	0.5710	-1.9995	1.7309	13103661376	-22.55%	
2001	NVE	0.3551	0.0000	0	32076433692	0.00%	28.4%
	HVT1	0.4682	-1.0836	0	30170508445	-5.94%	
	HVT2	0.4739	-1.1368	1.3401	28985676061	-9.64%	
2002	NVE	0.3954	0.0000	0	17837070189	0.00%	27.3%
	HVT1	0.5975	-2.5666	0	14296395900	-19.85%	
	HVT2	0.5586	-2.0959	1.8563	12937632459	-27.47%	
2003	NVE	0.4442	0.0000	0	29564066991	0.00%	30.0%
	HVT1	0.6190	-2.2630	0	24068897094	-18.59%	
	HVT2	0.5841	-1.8209	1.3761	23416435079	-20.79%	
2004	NVE	0.4211	0.0000	0	36106960195	0.00%	29.6%
	HVT1	0.5858	-1.9384	0	29644261702	-17.90%	
	HVT2	0.5397	-1.4084	-0.2393	30501650930	-15.52%	
2005	NVE	0.3694	0.0000	0	43625622505	0.00%	26.4%
	HVT1	0.5374	-1.8970	0	38675516143	-11.35%	
	HVT2	0.5029	-1.6319	2.7040	32235418335	-26.11%	
Time dependent ?	NVE	No					
	HVT1	No	No				
	HVT2	No	No	No			

Moreover, the statistical analyses of the parameters of the empty trip models considered (i.e., NVE, HVT1 and HVT2) found that they are not time dependent, i.e., they are stable over time. This should not be a surprise as previous research has already concluded that the percentage of empty trips is related to the symmetry of the commodity flow matrices. In this context, the more symmetric the commodity flow matrix, the easier for the carriers to find a backhaul and the lower the percentage of empty trips (Holguín-Veras, 2004). Since the parameters of the empty trip models are related to the percentage of empty trips (Holguín-Veras and Thorson, 2003), commodity flow matrices that are stable overtime, lead to percentage of empty trips and model parameters that are also stable.

VI. Conclusions

This paper conducts an empirical investigation on the temporal stability of parameters of freight demand models in the short-medium term. The analyses are based on seven national freight origin-destination samples conducted by the Colombian government during the years 1999 to 2005. The paper studies the stability of the parameters of freight generation, freight distribution, and empty trip models. Typical formulations of these models were calibrated using the cross-sectional data for to each year. Then, to identify time-dependent effects the resulting parameters were regressed as a function of a time index. The results indicate the presence of statistically significant time-dependent effects on all freight generation models (production and attraction), as well as on the freight distribution model estimated with loaded vehicle trips. In contrast, the parameters of the freight distribution models based on commodity flows and the ones for the empty trips were found to be stable overtime. The reason may be related to the fact that the commodity flows reflect production-consumption patterns that are much slower to change overtime than vehicle trips, that are the result of short term logistic decisions on the part of the carriers. The stability of the parameters of empty trips is also related to the stability of production-consumption patterns. This is because the percentage of empty trips—which is related to the parameters of the models—is directly determined by the degree of asymmetry of the commodity flow matrix, as the more asymmetric the matrix is the larger the percentage of empty trips. Since the parameters of the empty trip models are related to the percentage of empty trips, a stable percentage of empty trips lead to stable parameters.

The results indicate that the amount of cargo produced by a unit of GPD, and the amount cargo attracted by a unit of population have increased overtime. This seems to indicate a lowering of the unit value of the cargo transported in Colombia, and an increase in the amount of goods consumed by the citizenry. The decrease in the value of the parameters of the vehicle-trip distribution models indicates that travel impedance is less important than before. In contrast, the commodity flow distribution models were found to have stable parameters. This suggests that the difference is because of a change in the logistical patterns of the freight industry, as opposed to a change in the underlying demand. Faced with increasing customer demands, freight carriers are frequently pushed to provide faster service with lower payloads at longer distances. Such changes could explain the finding concerning the vehicle-trip distribution models.

Taken together, this research has provided a mixed bag of evidence concerning the validity of the parameter stability assumption. As discussed in the paper, the assumption was rejected in half the cases (freight generation and vehicle-trip distribution models), and found to hold in the other half (commodity flow distribution and empty trip models). Such result is befitting for such a complex and important topic as it leads to the obvious conclusion that, still, more research is needed to reach solid conclusions.

VII. References

DANE (2009). Population and GDP data by regions for Colombia. www.dane.gov.co.

Gunn, H., M. Ben-Akiva and M. Bradley (1985). "Test of the Scaling Approach to Transfering Disaggregate Travel Demand Models." Transportation Research Record(1037): 21-30.

Hautzinger, H. (1984). The Prediction of Interregional Goods Vehicle Flows: Some Modeling Concepts. Ninth International Symposium on Transportation and Traffic Flow Theory, VNU Science Press.

Holguín-Veras, J. (2004). On the Estimation of the Maximum Efficiency of the Trucking Industry: Implications for City Logistics. City Logistics III. E. Taniguchi and R. Thomson, Elsevier, Amsterdam**:** 123-134,.

Holguín-Veras, J. and G. Patil (2005). "Observed Trip Chain Behavior of Commercial Vehicles." Transportation Research Record **1906**: 74-80.

Holguín-Veras, J. and E. Thorson (2003). "Modeling Commercial Vehicle Empty Trips with a First Order Trip Chain Model." Transportation Research Part B: Methodological **37**(2): 129-148.

Holguín-Veras, J. and E. Thorson (2003). "Practical implications of Modeling Commercial Vehicle Empty Trips." Transportation Research Record **1833**: 87-94.

Holguín-Veras, J., E. Thorson and J. C. Zorrilla (2008). "Commercial Vehicle Empty Trip Models with Variable Zero Order Empty Trip Probabilities." Networks and Spatial Economics. http://dx.doi.org/10.1007/s11067-008-9066-7

Institute of Transportation Engineers (1991). Trip Generation Washington, D.C., Institute of Transportation Engineers.

McCarthy, P. (1982). "Further evidence on the temporal stability of disaggregate travel demand models." Transport Research Part B: Methodological **16B**(4): 263-278.

Noortman, H. J. and J. van Es (1978). Traffic Model. Manuscript for the Dutch Freight Transport Model.

Ortúzar, J. D. and L. G. Willumsen (2001). Modelling Transport. New York, John Wiley and Sons.

Robusté, F. (1994). "Network simulation, tolls and some conjectures on regional development: the case of Catalonia." Flux(17): 33-40.

Strauss-Wieder, A., K. Kang and M. Yokei (1989). The Truck Commodity Survey in the New York-New Jersey Metropolitan Area. Good Transportation in Urban Areas, Proceedings of the 5th Conference, Santa Barbara, CA.

Tretvik, T. and S. Widlert (1998). The stability of parameter estimates in household-based structured logit models for travel-to-work decisions. Travel behavior research: Updating the state of play**:** 457-472.

U.S. Census Bureau (2004). Vehicle Inventory and Use Survey. http://www.census.gov/prod/ec02/ec02tv-us.pdf.

Universidad Nacional de Colombia Sede Medellin, Universidad del Norte and Universidad Pedagogica Tecnologia de Colombia (2008). Desarrollar y poner en funcionamiento modelos de demanda y oferta de transporte, que permitan proponer opciones en materia de infraestructura, para aumentar la competitividad de los productos Colombianos. Ministry of Transport.