

## AN EXPLORATORY ANALYSIS OF SPATIAL EFFECTS ON FREIGHT TRIP ATTRACTION

**Iván Sánchez-Díaz, Ph.D.**

Corresponding Author

Postdoctoral Research Associate. Department of Technology Management and Economics

Chalmers University of Technology. Vera Sandbergs Allé 8, Room 3332

SE - 412 96 Göteborg, Sweden

E-mail: [ivan.sanchez@chalmers.se](mailto:ivan.sanchez@chalmers.se)

**José Holguín-Veras, Ph.D.**

William H. Hart Professor. Director of the Center for Infrastructure, Transportation, and the

Environment (CITE). Department of Civil and Environmental Engineering, Rensselaer

Polytechnic Institute. 110 Eighth Street, Troy, NY, USA. Phone: 518-276-6221.

E-mail: [jhv@rpi.edu](mailto:jhv@rpi.edu)

**Xiaokun (Cara) Wang, Ph.D.**

Assistant Professor. Department of Civil and Environmental Engineering, Rensselaer

Polytechnic Institute. 110 Eighth Street, Troy, NY, USA. Phone: 518-276-2098.

E-mail: [wangx18@rpi.edu](mailto:wangx18@rpi.edu)

### ABSTRACT

This paper conducts an exploratory analysis of freight trip attraction and its relationship with key features of the urban environment. Using establishment level data, the authors explore the role of business attributes, as well as network and land use descriptors. The research uses data from 343 establishments from five different industry sectors in New York City. These establishments are geo-located, and spatial association indicators are estimated to assess the presence of spatial effects. Spatial econometric techniques are used to assess the role of spatial effects among establishments and the urban environment. The empirical evidence

suggests that establishments' location, such as land-value and front street width, play an important role on Freight Trip Attraction (FTA), and that retail industries located in high employment zones tend to produce higher FTA per employee. Another key finding is that FTA is better modeled using non-linear models for all industry sectors. Specifically, the freight trip attraction of business establishments is concave with employment, flattening as employment increases. This is confirmed by the modeling results for which the range of coefficients estimated for employment reveals that, although larger establishments have higher FTA than small establishments, FTA increases at a diminishing marginal rate. These exploratory findings shed light on the use of locational variables, and nonlinear spatial effects specifications to enhance FTA models.

## **KEYWORDS**

Urban transportation systems, trip generation, freight, land use, spatial autocorrelation.

### *Bios:*

*Dr. Sanchez-Diaz received his Ph.D. from Rensselaer Polytechnic Institute. His research focuses on freight demand modeling, behavioral models and model development using large datasets.*

*Dr. Holguin-Veras is the Director of the VREF Center of Excellence for Sustainable Urban Freight Systems. His expertise is on freight transportation sustainable systems, policy development and transportation economics.*

*Dr. Wang received her Ph.D. from The University of Texas at Austin. Her research mainly focuses on data-driven models for analyzing the interactions between land use, transport, energy and environment, as well as the spatial dependence of transportation behaviors.*

## **1. INTRODUCTION**

As a result of the industrialization processes that have shaped the global economy, metropolitan areas have become centers of manufacturing and commerce, magnets of development that tend to attract new residents. As part of this trend, an increasingly large portion of the world's

population currently lives in urban areas, which, in most cases, are not fully prepared to accommodate the new residents. The implications of this population growth are significant, as the number of urban dwellers is expected to grow from 3.5 billion to 5 billion individuals by 2030. This increase in urbanization is bound to create major challenges for already strained urban infrastructure systems, and will likely have substantial environmental impacts. As a palliative to this situation, urban planners have proposed ideas to foster sustainable growth patterns and enhance urban livability, such as increasing walkable zones, adopting mixed land-uses with high population density and promoting the use of transit systems (Smart Growth Network and ICMA, 2002).

One of the factors that affects urban life the most, and that often draws the attention of urban residents, is the traffic congestion produced by urban activity. Due to its visibility, freight transportation is often associated with heavy traffic and negative environmental impacts. However, while freight transportation does produce negative impacts, it is also necessary for economic development. As stated by Ogden (1992), "freight transport is absolutely essential to modern urban civilization. The very concept of urbanization requires a freight system to sustain it, since urbanization means that large numbers of people are accumulated in areas remote from their sources of food, sources of raw materials for industry, markets for industrial products, and places to dispose of their waste." Despite its importance, the connection between freight and land-use is often disregarded. For instance, while some studies have analyzed the role of built-environment factors and land use decisions, such as the ones proposed by Smart Growth, on passengers vehicle miles traveled (Hong et al., 2013), this is not the case for freight where very few studies have analyzed how these factors and decisions can impact the efficiency of freight movements (Bassok et al., 2013). To avoid unexpected negative outcomes, it is important to consider the freight system when designing urban transportation policy, and seek strategies that reduce negative externalities without hampering economic progress.

Traffic congestion affects both citizens and businesses. For businesses, congestion impacts affect primarily accessibility, accidents, and operations (e.g., productivity, deliveries reliability). Moreover, Hartgen et al. (2014) show that businesses tend to select their location based on the convenience of the location, the accessibility, the visibility and also based on traffic congestion. In particular, businesses located in downtown areas and in wide streets tend to see traffic congestion as a problem. According to the study, while few businesses would consider relocation, a large portion of them would react to rising congestion by changing their delivery and shipment decisions, e.g., use 3<sup>rd</sup> party delivery services, consolidate shipments (Hartgen et al., 2014).

A key step in planning transportation systems and designing transportation policies is to estimate the amount of trips generated by the various industry sectors in a city. As in the traditional transportation demand literature (Ortúzar and Willumsen, 2011), the term “generation” is used here to refer to both attraction and production of trips. In the case of freight, it is also crucial to establish the difference between freight generation (FG) and freight trip generation (FTG). FG and FTG are two different concepts: FG refers to production and consumption (demand) of actual cargo while FTG refers to freight traffic required to transport the FG (Holguín-Veras et al., 2011). As the traffic is the one producing the negative externalities, it is essential to understand and measure FTG for the different economic activities taking place in the urban environment.

This paper seeks to enhance FTG modeling techniques by exploring freight trip attraction’s locational and spatial determinants. This is important because spatial considerations are central to business location decisions. In addition to the variables traditionally incorporated in FTG models, such as employment and industry sector, a complementary set of variables is used to relate the establishment FTG to the urban geography. The paper builds on preliminary results that Sanchez-Diaz et al. (2013) presented. Although FTG is composed of freight trip attraction

(FTA) and freight trip production (FTP), the main focus of this research is on FTA, as FTP data were not available.

## 2. LITERATURE REVIEW

The amount of existing research on FTG is very small. The Institute of Transportation Engineers Trip Generation Manual (Institute of Transportation Engineers, 2008) uses procedures that estimate the number of truck-trips as a percentage of passengers trips, which implicitly assumes that passenger and freight follow similar patterns. This simplification leads to significant errors as shown in Lawson et al. (2012) and Holguin-Veras et al. (2013). In search of better models, researchers have studied the relationship between FTG and economic variables. Brogan (1980) and Lawson et al. (2012) study FTG using employment and land-use as a categorical factor. Iding et al. (2002) and Barlett and Newton (1982) consider other variables in addition to employment. Bastida and Holguín-Veras (2009) and Holguín-Veras et al. (2011) show that economic variables work well to produce FTG estimates.

FTG is often estimated for facilities that have large impacts on local traffic. Not surprisingly, distribution facilities and warehouses have been the focus of specific studies to identify and quantify the traffic impacts produced by logistic-related land-uses (DeVries and Dermisi, 2008, Wagner, 2010). Similarly, Holguin-Veras et al. (2002) and Al-Deek et al. (2000) conducted FTG studies at marine container terminals to analyze their traffic impacts. As shown in this review, variables such as type of facility, employment, and land-use types have been used as independent variables of FTG models. However, no publications discuss the effect of the locational variables on FTG.

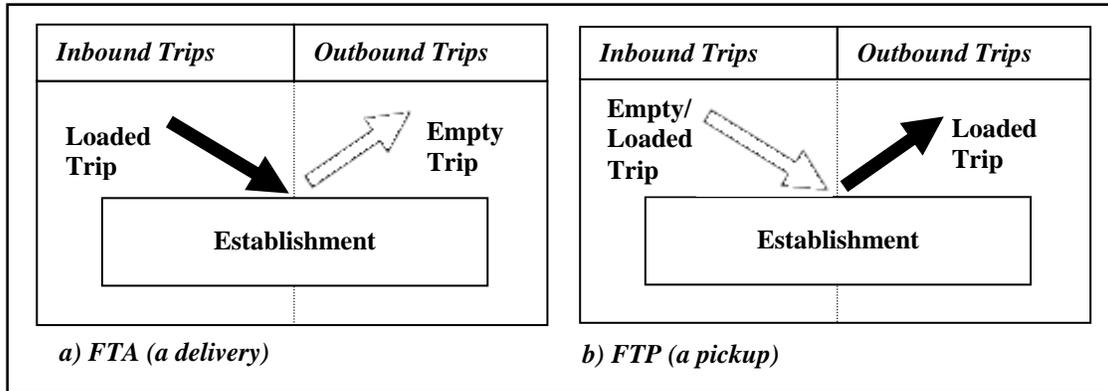
In terms of modeling approaches, most models are based on regression analysis or trip rates (Al-Deek et al., 2000, Wagner, 2010, Campbell et al., 2011, Holguín-Veras et al., 2011, Holguín-Veras et al., 2012). Starkie (1967) proposes a nonlinear specification to model FTG, though he only estimated it for manufacturing and engineering establishments. More elaborate

models use multiple classification analysis (Bastida and Holguín-Veras, 2009, Campbell et al., 2011, Holguín-Veras et al., 2012), and artificial neural networks (Al-Deek, 2001). Garrido and Mahmassani (2000) study the effects of geographical location and economic activity of firms on shipments' pattern for a single motor-carrier. Novak et al. (2011) evaluated the use of spatial regression variables, and estimated spatial regression models for freight generation at the national level. Kawamura and Miodonski (2011) estimated retail goods consumption using socioeconomic and land-use variables as explanatory variables while controlling for spatial interactions. In essence, the literature has highlighted the importance of geographical location and spatial effects to improve FG modeling. However, there is a lack of research on how these characteristics affect FTG. This research attempts to help fill these knowledge gaps by exploring the locational and spatial effects on FTA using econometric techniques.

### **3. FREIGHT TRIP ATTRACTION MODELING**

FTG estimates the total number of freight-related trips generated by an establishment or a transportation analysis zone (TAZ). Some basic relations hold (see Figure 1). First, at the establishment level, the number of inbound and outbound truck-trips on any given day are likely to be the same, as most trucks do not spend the night on the establishment they deliver to. Second, there are basically two types of activities that can produce a freight-related trip: a delivery (FTA) or a pickup request (FTP); each of these activities generates one inbound trip and one outbound trip in the absence of trips where a pickup and a delivery are made in the same stop. It is important to make a distinction between FTA and FTP as they are driven by different factors.

#### **Figure 1: FTG at the establishment level**

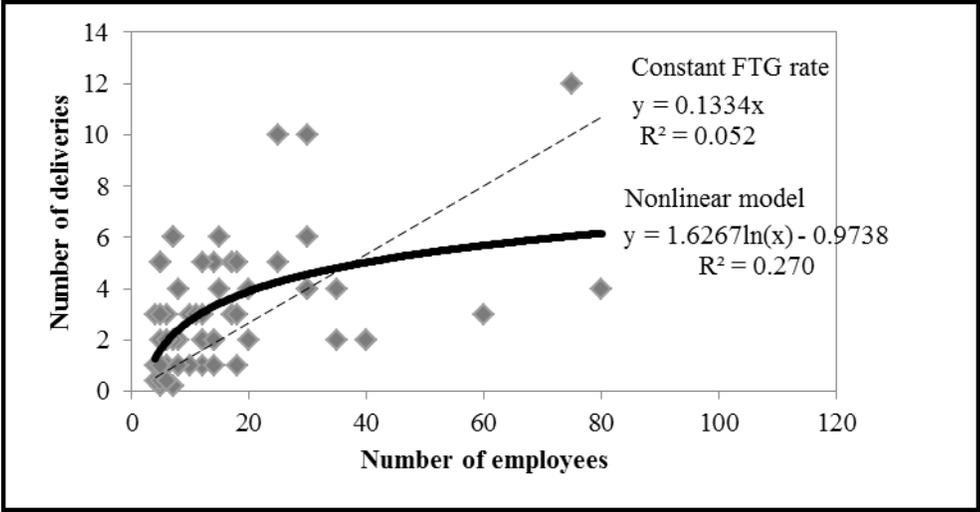


As shown in Figure 1, for trips with one stop, the outbound trips resulting from FTA and inbound trips resulting from FTP are empty trips. However, most carriers distribute goods using multiple stops tours to maximize efficiency, so that FTA and FTP from multiple receivers/ shippers are served in one single tour. Using this principle, tour-based models use carriers data to derive FTG by considering that each trip-end in the tour is the result of an FTA, or an FTP, see Sánchez-Díaz et al. (2014).

### 3.1. FTA Models

An important aspect for improvement of traditional FTG models is the assumption of proportionality between FTA and business size. As noted by Holguín-Veras et al. (2011), while the amount of cargo needed at the establishment—referred to in this paper as freight attraction (FA)—tend to be a proportional to business size, this is not necessarily true for FTA. The reason is that an establishment’s FTA is determined by the frequency of shipments, which is decided by the firm on the basis of logistical considerations. In essence, a firm will decide the optimal shipment size and delivery frequency based on the FA, the operational constraints, and the costs (shipping, handling and inventory costs). Thus, an increase in the FA can be handled through an increase in the shipment size and, if needed, a change in the size of the vehicle, or an increase in FTA. Small firms produce proportionally large amounts of FTA as the transport of a small amount of FA requires a truck trip (FTA). To minimize costs, large firms can increase the shipment size and change the truck size without affecting FTA; so although they have larger FA,

the firm will still attract a single trip. Large firms will, thus, have smaller FTA per ton than small firms. As FA is proportional to the business size, the relationship between FTA and business size is better explained with a nonlinear function. These concepts are illustrated in Figure 2. As shown, for a given industry sector, a marginal increase in business size tend to generate a larger increase in FTA for small establishments than for large ones.



**Figure 2: Non-linear Effects on FTG**

Holguín-Veras et al. (2012) explored different modeling techniques and the effects of using different industry/land-use classification systems. They considered different specifications: FTA can be constant per establishment, can depend on the business size, or can have both a constant term and a business size dependent term. The results for FTA are presented in Table 1.

**Table 1: Freight Attraction Models**

NAICS	Obs.	Const. / Empl.		RMSE	R <sup>2</sup> adjusted	Akaike Information Criterion (AIC)
		c	b			
Construction (NAICS 23)	25	10.800		6.81	0.00	155.4
Manufacturing (NAICS 31, 32, 33)	51	14.157		15.58	0.00	426.4
Wholesale Trade (NAICS 42)	117	11.360	0.345	18.68	0.07	998.7
Accommodation and Food (NAICS 72)	56	6.535	0.403	10.60	0.22	398.0
Retail Trade (NAICS 44, 45)	98	15.348	0.313	22.50	0.12	874.3

Source: Adapted from Holguín-Veras et al. (2012) to estimate weekly deliveries.

Note: AIC estimated based on (Akaike, 1987)

As shown by Holguin-Veras et al. (2013), these models outperform most models in the literature in terms of their prediction power; though they only explain a small part of FTA variance. One of the reasons for the low prediction power is the small set of variables used: business size (e.g., number of employees, area), industry sector and land-use. These variables are often used because they are readily available and have proven to be effective to estimate FTA; though there are other unexplored variables that can enhance current modeling techniques. This is the case of locational effects that influence FG, the transportation costs, the inventory costs and some operational constraints. For instance, an establishment that is located in a major arterial, which can use its front space to promote its business and is more accessible to customers, is likely to generate more business than another one that is hidden in an alley; the higher FA required may lead to larger FTA even if both establishments have the same number of employees. Similarly, establishments that pay for premium space to be located in a high income location will have different FTA patterns than establishments located in isolated suburban areas, not only because of their higher FA but also because of limited space availability and higher inventory costs that will influence their logistic decisions. As a result, variables that capture, even in a proxy manner, the importance of business location could play a vital explanatory role. This paper uses New York City (NYC) as a case study to investigate whether the inclusion of locational variables, spatial interactions, and nonlinearity can enhance FTA models.

#### **4. DESCRIPTION OF THE DATA**

As part of previous research projects, the authors collected establishment-level FTG data from 343 receivers in NYC. The data cover the entire borough of Manhattan and the western portion of Brooklyn. The data were collected using two phases: in the first phase, letters describing the objective of the data collection and a questionnaire were sent to the sampled establishments; in the second phase the questionnaires were completed using a computer aided telephone

interviews (CATI). While the interviews for Manhattan were conducted on April 2005, the ones for Brooklyn were conducted on March 2006. The sample was randomly selected to ensure the representativeness of the sample, though more weight was given to freight-intensive industries over service and only receivers with more than five employees were selected. For more details on the sampling and data collection process, see (Holguín-Veras, 2006).

The data contain information about employment, sales, type of commodities, land-use and industry sector. A breakdown by industry sector reveals that most of the observations are in the wholesale trade sector (34%), retail (28%), and accommodation and food (16%). In terms of business size, most of the establishments (90%) in the sample are small to medium-sized, with 5 to 50 employees. This is consistent with the pattern in Manhattan and Brooklyn, where more than 90% of establishments are within this employment range (U.S. Census Bureau, 2011).

The data used for the external validation were collected from receivers in NYC in 2012 using mail-out letters and questionnaires and complemented with CATI. The data cover the FTA patterns for 100 retailers, 43 establishments in the manufacturing sector, 26 wholesalers, and 7 establishments in the food and accommodation sectors. The ranges for FTA are similar to the ones in the calibration datasets: 1 to 100 weekly deliveries. Unfortunately, as opposed to the calibration dataset, the validation dataset does not contain data for construction establishments. Similar to the calibration dataset, freight-intensive industry sectors were the focus of the study. The distribution by employment was selected to resemble the distribution in NYC. For a detailed description of the survey conducted and the descriptive analysis, see (Holguín Veras et al., 2014).

#### **4.1. Explanatory Variables**

This research examines the explanatory value of variables related to land-use, economic attributes, and network characteristics to model FTA. In doing this, the establishments were geo-located using the addresses provided in the survey. Additional data were gathered from

different sources, such as the NYC Department of City Planning website (NYC Department of City Planning, 2011), the NYC Property website (The City of New York, 2011) and satellite images. Table 2 presents a brief description of the variables considered. These variables cover different aspects that can affect FTA and that are relatively easy to observe. Land-use variables include land-value, zoning regulation, and geographic location. While land-value is a proxy of the customers' income and the inventory costs, the zoning regulation and geographic location are binary variables that group establishments according to shared characteristics (e.g., floor area ratio, distance to distribution centers) that can affect FTA patterns. Economic attributes include the industry sector, the role of the establishment in the firm (e.g., headquarter), the main type of commodities received, number of vendors and employment. These variables are crucial to measure the business activity in the establishment, and to make a distinction between the economic activities taking place at the premises. Network characteristics include distance to truck route, distance to large traffic generators (i.e., buildings or facilities receiving high number of deliveries), the distance to the primary network, and the width of the street. These variables are used to explore the role of location on FTA. Distance to truck routes and distance to primary network measure the effect of being located on main streets, and in their proximities. To complement the study of location, width is used as a proxy of the degree of importance of the street where the establishment is located.

To ensure a meaningful connection between the models and the economic activity performed, the observations were grouped by industry sector; the 2-digits North American Classification System (NAICS) codes were selected to classify the data. For a detailed description of the industry sectors, see U.S. Census Bureau (2010b). As not all the establishments have complete data, the number of observations per sample differ slightly, depending on which variables were included.

**Table 2: Description of Variables Considered**

Type of Variable	Variable	Description
<b>Land-Use Variables</b>	Land-Market Value	A proxy variable for the price of land, estimated using the NYC Property website. The website provides the building size in square feet (BS), the number of stories (NS), the number of buildings (NB), and the estimated market value (MV). The unit value is estimated as: $Value = \frac{MV}{BS * NS * NB * 100}$
	NYCZR	Binary variable that represents the location of the establishment in one of the districts: 8 different commercial districts/ 3 types of manufacturing districts/ 1 residential district. Districts are also classified by floor to area ratio allowances
	Geographic Location	Binary variable that represents establishments' location: Manhattan (MH), Downtown Manhattan (DM), Midtown Manhattan (MM), or Brooklyn (BK)
<b>Economic Attributes</b>	NAICS	Binary variable representing the industry sectors: Construction, Manufacturing, Wholesale, Retail, Accommodation and Food
	Type of Establishment	Binary variable indicating the function of the establishment in the organization ( <i>branch, headquarters, or single</i> )
	Commodity Type	Binary variable indicating the type of commodities received (e.g., textile and clothing)
	Number of Vendors	This variable indicates the number of suppliers of the establishment
	Employment	Number of employees working in the establishment on a typical day
<b>Network Characteristics</b>	Distance to Truck Route	Euclidian distance (in miles) to the closest truck route as defined by the NYC Department of Transportation
	Distance to the Primary Network	Euclidian distance (in miles) to the closest street belonging to the primary network
	Minimum Distance to an LTG	Network distance (in miles) to the closest large traffic generator (LTG). An LTG is a building or facility that receives a fairly high number of deliveries
	Mean Distance to LTG	Average network distance (in miles) to the LTGs in New York City
	Width	Width of the street (in feet) in front of the establishment

Table 3 summarizes the descriptive statistics of the sample. As shown, the FTA varies from 1 to 125 weekly deliveries. On average, construction establishments attract the smallest number of trips (11 weekly deliveries), while retail establishments attract the largest number of trips (21 weekly deliveries) most likely due to the vastly different shipment sizes. The number of

employees varies from 1 to 80, with retail establishments having the lowest average employment (15 employees), and accommodation and food services establishments having the highest (21 employees). In terms of locational variables, the data show that wholesale trade establishments have the widest front streets on average (41 feet), while construction has the narrowest (37 feet). The statistics for the land market-value proxy show that accommodation and food and retail establishments are located in the most expensive areas (\$253 and \$211 per square feet respectively), while wholesale, construction and manufacturing establishments are located in less expensive ones (\$155, \$100 and \$96 per square feet respectively). These results make sense on account of the consumer orientation of food and retail establishments, and the space needs of manufacturing.

**Table 3: Descriptive Statistics of NYC Sample**

NAICS 23: Construction							NAICS 31-33: Manufacturing				
Variable	Unit	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
<i>Deliveries</i>	deliveries/week	24	11	7	1	25	51	14	16	1	100
<i>Employment</i>	employees	24	19	13	3	62	51	22	17	3	75
<i>Width</i>	feet	24	37	9	25	57	51	41	12	20	71
<i>Dist_TR</i>	miles	24	0	0	0	1	51	0	0	0	8
<i>Min_LTG</i>	miles	22	6	1	2	11	49	28	2	0	65
<i>Value</i>	US\$ / sq. ft	22	100	75	19	276	48	96	87	7	390

NAICS 42: Wholesale Trade							NAICS 44-45: Retail Trade				
Variable	Unit	Obs.	Mean	Std. Dev.	Min	Max	Obs.	Mean	Std. Dev.	Min	Max
<i>Deliveries</i>	deliveries/week	115	17	19	1	125	96	21	24	2	125
<i>Employment</i>	employees	115	17	16	2	80	96	16	15	1	75
<i>Width</i>	feet	112	41	11	20	77	96	44	14	20	86
<i>Dist_TR</i>	miles	115	0	0	0	1	96	0	0	0	1
<i>Min_LTG</i>	miles	112	3	2	0	13	94	3	2	0	14
<i>Value</i>	US\$ / sq. ft	110	155	109	8	648	92	211	167	33	1017

NAICS 72: Accommodation and Food Services						
Variable	Unit	Obs.	Mean	Std. Dev.	Min	Max
<i>Deliveries</i>	deliveries/week	54	14	10	3	50
<i>Employment</i>	employees	54	21	20	3	80
<i>Width</i>	feet	54	40	14	11	71
<i>Dist_TR</i>	miles	54	0	0	0	1
<i>Min_LTG</i>	miles	53	2	1	0	7
<i>Value</i>	US\$ / sq. ft	51	253	180	27	796

To get an idea about each sector's contribution to FG and FTA in the New York City Metropolitan area, the authors used Wood and Leighton (1969)'s FG estimates (although outdated, these are the latest FG data available), and applied Holguín-Veras et al. (2012)'s FTA models. The models in Holguín-Veras et al. (2012) consider only freight-intensive industry sectors, such as retail, wholesale, accommodation and food, construction, and manufacturing, for the other sectors the authors assumed FTA as one daily trip.

**Table 4: FG and FTA estimates for New York City Metropolitan Area**

Industry Classification	New York City Metropolitan Area					
	Establishments	Percent of Establishments	Daily tons	Percent of tonnage	Daily trips	Percent of trips
Retail trade	49,944	14.1%	73,242	9.6%	185,628	28.9%
Wholesale trade	26,758	7.6%	94,606	12.4%	82,390	12.8%
Accommodation and food	26,611	7.5%	n.a.	n.a.	67,520	10.5%
Construction	30,098	8.5%	139,620	18.3%	65,012	10.1%
Manufacturing	11,349	3.2%	125,887	16.5%	33,197	5.2%
For-hire carriage	8,050	2.3%	189,212	24.8%	8,050	1.3%
Utilities and communications	248	0.1%	55,695	7.3%	536	0.1%
Others	200,596	56.7%	84,688	11.1%	200,869	31.2%
Total	353,654	100.0%	762,950	100.0%	643,201	100.0%

Note: The FTA models were applied using the NAICS and data from the county business patterns (U.S. Census Bureau, 2010a) from the following counties: Bronx, Kings, Nassau, New York, Putnam, Queens, Richmond, Rockland, Suffolk and Westchester. For the geographic area covered by the FG, refer Wood and Leighton (1969). Wood and Leighton (1969) numbers should be used cautiously as the estimates are very old.

The results revealed that the retail sector represents 14.1% of the establishments in the city, attracts 9.6% of the freight within the city and 28.9% of FTA; the 7.6% of establishments in the wholesale sector attracts 12.4% of the freight and 12.8% of FTA, while the 7.5% of establishments in the accommodation and food sector attracts 10.5% of FTA (no estimates where reported for freight attracted). The construction sector represents 8.5% of the establishments and is responsible for 18.3% of the FG and 10.1% of the total FTA of the city. The manufacturing sector represents only 3.2% of the establishments in the city and is

responsible for 16.5% of the freight and 5.2% of FTA. According to Wood and Leighton (1969), the for-hire carriage sector accounts for 24.8% of freight movements; utilities and communications accounts for another 7.3%, and other industry sectors for the remaining 11.1%. These sectors attract about 31.2% of the FTA.

The variables in this section describe the economic characteristics of the establishments and relate the establishments to their environment. Another aspect that plays an important role in FTA, but which cannot be captured using these variables, is the spatial interaction between establishments. The next section sheds light on this important aspect.

#### **4.2. Spatial Effects Indicators**

Spatial indicators are measures used to quantify relative spatial location; their objective is to convey the underlying spatial processes. They include a wide range of techniques such as centro-graphic studies, point pattern analysis and spatial autocorrelation. The paper focuses on the spatial autocorrelation because its presence affects the estimation of FTA model parameters, resulting in biased and inefficient parameters (Anselin, 1988). The most used specification test for spatial autocorrelation is the Moran's  $I$ , which is defined as:

$$I = \frac{N \sum_i \sum_j \mathbf{w}_{i,j} (X_i - \bar{X})(X_j - \bar{X})}{(\sum_i \sum_j \mathbf{w}_{i,j}) \sum_i (X_i - \bar{X})^2} \quad (1)$$

Where  $N$  is the number of observations,  $\bar{X}$  is the mean value of  $X$ 's,  $X_i$  is the value at location  $i$ , and  $W_{i,j}$  is the weight indexing location of  $i$  relative to  $j$ .  $W_{ij}$  is defined as the inverse of the distance between establishments.

Moran's  $I$ —the two-dimensional equivalent to the serial correlation metric—is a global measure that studies the spatial interaction between observations in the study region. Moran's  $I$  coefficient varies between -1 and 1. An absolute value close to +- 1 reflects a high spatial autocorrelation; while a value of zero means that there is no spatial autocorrelation implying a

random pattern. The expected value of Moran's  $I$  under the null hypothesis of no spatial autocorrelation is used to assess the sign and degree of the autocorrelation (Moran, 1950), it is defined as follows:

$$E(I) = \frac{-1}{N-1} \quad (2)$$

A Moran's  $I$  value greater than  $E(I)$  indicates a tendency toward clustering, and a value less than  $E(I)$  indicates a tendency toward dispersion or uniformity (Moran, 1950). However, autocorrelation may exist in some zones of the region, or have different signs in different zones. For these cases, the area of study is divided using Thiessen polygons combined with Local Indicators of Spatial Association (LISA) to calculate the interaction between neighboring polygons (Anselin, 1988). Since each index has an associated test statistic, it is possible to map which of the polygons have a statistically significant relationship with its neighbors.

A descriptive statistical analysis of spatial patterns in the FTA data was conducted to assess spatial autocorrelation in each industry sector.

Table 5 presents a summary of the spatial effects indicators for each industry sector, while Figure 3 provides a visual representation of clusters with similar FTA identified through a local analysis.

**Table 5: Summary of Spatial Effects Indicators**

NAICS	Description	Moran's I		LISA analysis		Remarks
		I	E(I)	Clusters	Outliers	
NAICS 23	Construction	0.03	-0.05	YES / LOW	YES	Moran's I indicates a global tendency towards clustering in FTA
NAICS 31-33	Manufacturing	-0.04	-0.02	YES / HIGH	YES	LISA reveals some positive associations (i.e., clusters)
NAICS 42	Wholesale Trade	-0.02	-0.01	YES / LOW	YES	LISA reveals several positive associations
NAICS 44-45	Retail Trade	-0.11	-0.01	YES / LOW	YES	LISA reveals the presence of some clusters and several outliers
NAICS 72	Accommodation & Food	-0.02	-0.02	NO	YES	Moran's I shows no significant indication of spatial autocorrelation. LISA reveals two outliers

As shown in,Table 5, the Moran's *I* indicators are close to zero. However, most of the sectors have *I* values different from the expected value, which indicate some degree of spatial autocorrelation. In the case of construction, *I* is greater than the expected value, which is an indication of tendency towards clustering. This pattern is corroborated in the LISA cluster map of Figure 3, where the cluster indicates a positive association between a polygon and its neighbors. In this case a polygon with a low level of deliveries is surrounded by polygons with low levels of deliveries. In contrast, manufacturing, wholesale trade and retail trade exhibit *I* values lower than the expected value. Although the global tendency is towards dispersion in these cases, the local analyses shown in Figure 3 reveal the presence of several low-level deliveries clusters. For the manufacturing sector, the positive association takes place between polygons with high levels of deliveries. Although at this point it is not possible to make conjectures about the reasons for these patterns, these results suggest that neighborhood

effects and shared locational effects, such as land-value, land-use regulation, employment of neighbor establishments, and width of street, likely play a role in FTA.

In general, the LISA indicators suggest the presence of spatial autocorrelation in FTA in most of the sectors. The next step is to study the systematic spatial autocorrelation between establishments and the relationship between the variables listed in Table 2 and FTA using econometric models, and explore how these relationships can be exploited to enhance FTA models.

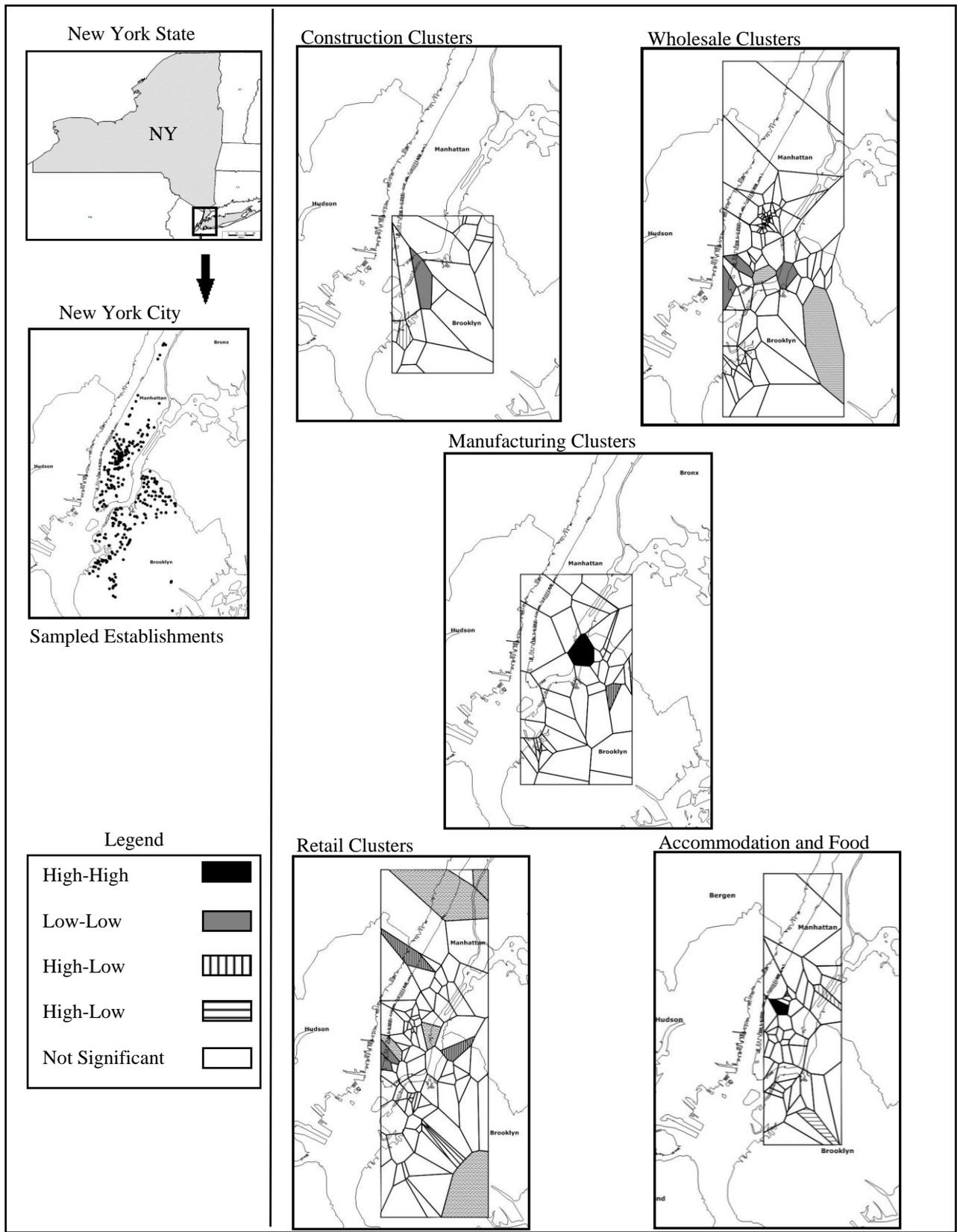


Figure 3: Local Indicator of Spatial Autocorrelation Cluster Maps

## 5. ECONOMETRIC MODELS

### 5.1. Methodology

The authors use two different econometric approaches. The first one is ordinary least squares (OLS), which is used to determine the empirical associations between geographical features (e.g., floor area ratio allowance, network characteristics), economic attributes of establishments, and FTA. This approach, though very useful, ignores the spatial interaction between establishments described in the previous section. The second approach uses spatial econometric models to assess spatial autocorrelation effects. Spatial econometric methods specify, estimate and test for the presence of spatial interactions (Anselin, 1988). In essence, they explicitly consider spatial weights and lags in regression models. The concept of neighborhood is explained by a spatial weight matrix  $\mathbf{W}$  built with the distances between every pair of observations. The spatial lag for variable  $\mathbf{X}$  at  $i$  can therefore be expressed as:

$$\mathbf{W} \bullet \mathbf{X} = \sum_{j=1 \dots N} (w_{ij} * X_j) \quad (3)$$

Where  $\mathbf{X}$  is an  $N$  by 1 vector of observations on the random variable,  $j \in N$  are the neighborhood observations,  $\mathbf{W}$  is the spatial weight matrix,  $X_j$  is the lag variable and  $w_{ij}$  is the Euclidian distance between  $i$  and  $j$ . *Since the NYC area and the sample are bisected by the East River, the use of network distance would be a more appropriate choice. However, for practical reasons the exploratory analysis of this paper considers the Euclidian distance.*

These spatial effects can be incorporated into the standard linear regression model in two different ways: as an additional regressor in a Spatial Lag Model (SLM) or as a specific component of the error structure ( $E[\varepsilon_i \varepsilon_j] \neq 0$ ) in a Spatial Error Model (SEM). In the former case a spatial lag variable ( $\mathbf{W} \bullet \mathbf{X}$ ) is used as an additional regressor. The SLM is appropriate when the focus of interest is the assessment of the existence and strength of spatial interaction. The SEM is appropriate when the concern is with correcting for the potentially biasing influence of the spatial autocorrelation, due to the use of spatial data. Anselin (1988) proposes the spatial

two stage least squares and the method of moments estimators to estimate these models. The spatial econometric models in this paper are estimated using maximum likelihood in the software developed by (Anselin et al., 2006).

Three criteria were used to assess if a spatial econometric model is a better alternative than the corresponding OLS version: the conceptual validity of the spatial terms, the statistical significance of these terms, and the Akaike information criterion (AIC) (Akaike, 1987). The latter deals with the tradeoff between bias and variance in model construction, and is a function of the number of parameters and the log likelihood of the model. Whenever the AIC is lower for the spatial model than for the OLS model, the phenomenon is better modeled using a spatial econometric model.

## **5.2. Results**

The estimated OLS and spatial econometric models are summarized in Table 6. Multiple linear and nonlinear specifications were assessed for the industry sectors. A Ramsey Regression Equation Specification Error Test (RESET) (Ramsey, 1969) confirmed the need for a nonlinear specification for the wholesale trade, where the null hypothesis ( $H_0$ : all the nonlinear coefficients are zero) was rejected at the 5% level of confidence; for the other sectors the RESET was not conclusive. Table 6 presents a summary of the econometric models selected for each industry sector. SEMs and SLMs were estimated for all the industry segments. However, as spatial terms were statistically significant only for retail, the other results are not included in the table.

As shown in Table 6, the models selected are all nonlinear functions of employment for all the sectors. In the cases of construction, and accommodation and food sectors, locational variables (i.e., width of the front street and the value of land) also play a significant role; while for retail establishments, the location with respect to other establishments plays a significant role. As revealed by the AIC, the models in Table 6 are in general a better alternative than the benchmark models in Table 1.

**Table 6: Econometric Models Summary**

Dependent variable: $\ln(\text{Weekly deliveries} / \text{establishment})$									
	Construction		Manufacturing	Wholesale	Accommodation and food		Retail		
Regressor	NLM-1	NLM-2	NLM	NLM	NLM-1	NLM-2	NLM	SLM-1	SLM-2
<i>Constant</i>	1.11* (1.93)		1.30 (3.33)	0.97 (3.47)	1.44 (5.09)		1.73 (6.59)	0.03 (0.22)	0.08 (1.01)
$\ln(\text{Employment})$	0.41 (1.99)	0.45 (2.42)	0.36 (2.67)	0.57 (5.37)	0.35 (3.47)	0.40 (4.29)	0.37 (3.56)	0.65 (7.23)	0.69 (7.37)
$W \bullet \ln(\text{Employment})$									0.23 (2.59)
$\ln(\text{value of land})$						0.24 (5.01)			
<i>Width of front street (in feet)</i>		0.03* (1.86)							
<i>Lambda</i>									
$W \bullet \ln(\text{Deliveries})$								0.36 (3.30)	
n	23	23	51	115	54	51	96	94	94
Adj. R <sup>2</sup>	0.12	0.90	0.79	0.20	0.17	0.94	0.11	0.76	0.77
F-stat	3.9	108.8	7.1	28.9	12.0	368.9	11.0	n.a.	n.a.
Akaike Information Criterion (AIC)	52.1	52.4	123.1	308.4	105.7	99.5	224.2	216.7	185.4

Notes: t-statistics are shown in parenthesis; all variables are significant at the 5% level except (x) which is significant at the 10% level; NLM are non-linear ordinary least squares models; SLM are spatial lag models.

As shown, FTA has a nonlinear relationship with employment and the  $\ln$ -coefficient ranges from 0.35 (accommodation and food) to 0.69 (retail). This range of coefficients reveals that larger establishments have higher FTA than small establishments; though FTA increases at a diminishing marginal rate. For instance, while for a small manufacturing establishment of 3 employees, an additional employee is likely to increase FTA from 6.0 to 5.5 (i.e., an increase by 0.5 trips) per week according to equation (4); for a large establishment of 75 employees, an additional employee does not increase significantly its FTA (i.e., 17.4 trips for 75 or 76 employees).

$$FTA = 3.67 * E^{0.36} \tag{4}$$

To analyze the implications of the nonlinearity in more detail, the authors computed the partial derivative of equation (4) with respect to employment for the manufacturing sector. See equation (5). The negative sign of employment indicates that the marginal increase in FTA becomes smaller as employment increases. The same relationship holds for the other sector where the coefficient for employment is always less than one.

$$\frac{\partial FTA}{\partial E} = 1.32 * E^{-0.64} \quad (5)$$

As shown in Table 6, the variables relating the establishment to its environment play a key role in improving the explanatory power of FTA models for some sectors. In terms of land-value, the proxy variable used to assess the effect of the property value was found to be statistically significant for accommodation and food establishments, but has no significant effect on the other industry sectors. For this sector, FTA can be estimated as follows:

$$FTA = E^{0.40} LV^{0.24} \quad (6)$$

Where  $E$  is the employment and  $LV$  the land value. The positive coefficient for  $LV$  reveals that an increase in the property value is associated with a higher FTA. However, as the coefficient is less than 1 and the relationship is nonlinear, FTA increases at a diminishing marginal rate of land-value. This coefficient seems to capture the FTA pattern of restaurants and hotels located in premium properties, where storage space is small and inventory costs are high, thus generating a higher FTA.

In the case of construction, the variable measuring the width of the establishment's front street has a statistically significant effect on FTA. This relationship can be stated as follows:

$$FTA = E^{0.45} e^{0.03*WS} \quad (7)$$

Where  $WS$  is the width of the front street.

The authors' conjecture is that the street width is acting like a proxy of the degree of importance of the road where the establishment is located. Establishments with wider front streets attract significantly more trips than the ones with small front streets. For instance, a construction establishment located in a 25 feet street attracts half the number of deliveries than one located in a 48 feet street, even though they have the same number of employees. One may argue that this model suffers from simultaneous causality bias because construction firms with large FTA select establishments with wide front streets. To shed light on this issue, the presence of a reverse causality is assessed by, (1) computing the correlation between the width of the front street and the error term in the regression, and (2) estimating a model where the width of the front street is the dependent variable and FTA and employment are the independent variables. In the first case, the results reveal a low negative correlation (-0.17) between the width of the street and the regression error term. For the second approach, none of the independent variables, i.e., FTA,  $\ln(\text{FTA})$ , employment,  $\ln(\text{employment})$ , are statistically significant to explain the street width. Although there is no strong evidence of a simultaneous causality bias for this particular case, further research on this reverse causality can provide insights on how FTA affects the firms' location decisions.

In the case of retail establishments, two statistically significant and conceptually valid models were found, in addition to the OLS non-linear model. As shown in Table 6, the SLM-1 has a significant positive effect of the lag term (*with deliveries*) that reveals the presence of cluster effects: establishments that are closer to peers with high FTA tend to have larger FTA. This model is useful to detect the sign of the spatial autocorrelation but has limited applicability because one needs to know the establishment neighbors' FTA beforehand. SLM-2 found that the coefficient of the employment lag term (*with employment*) is positive and statistically significant. Therefore, a retail establishment that is located in a zone with high retail employment and/or high retail establishment density tends to generate more FTA per employee than one in a zone with low retail employment, as shown below:

$$FTA_i = 1.03 * E_i^{0.69} * \prod_{j \in J} E_j^{0.23w_{ij}} \quad (8)$$

Where,  $i$  is the establishment under study,  $j \in J$  are retail establishments in the area of influence (the threshold is set to 0.023 in Euclidean distance to guarantee at least one neighbor per observation), and the weights for the sample range between 0.004 and 0.023 in Euclidean distance. The weights can be converted to other units using Geographic Information System (GIS) software. As both SLM-1 and SLM-2 have smaller AIC than the NL model, they represent a better alternative in statistical terms (with SLM-2 being preferred over SLM-1).

### **5.3. External validation and applicability**

The models presented in the previous section and the benchmark models in Table 1 were applied to the validation datasets; the Mean Absolute Percentage Error (MAPE) was estimated to assess their external validity and compare their performances. In the cases where a locational variable is statistically significant, both the nonlinear model without the locational variable (NL-1) and including it (NL-2) were assessed. The results are shown in Table 7.

As shown in Table 7, the MAPEs for the validation datasets are generally slightly larger than the ones for the calibration datasets, which is expected due to the exploratory nature of the models and the need to gain a deeper understanding of the determinants of FTA. The exceptions are the retail models.

**Table 7: External Validation Results**

Industry Sector (NAICS)	Model		MAPE	
			Calibration	Validation
Construction (23)	Benchmark	10.8	97%	n.a.
	NL-1	$3.03 * E^{0.41}$	72%	n.a.
	NL-2	$E^{0.45} + \exp(0.03 * WS)$	65%	n.a.
Manufacturing (31-33)	Benchmark	14.16	134%	146%
	NL	$3.67 * E^{0.36}$	79%	114%
Wholesale (42)	Benchmark	$11.36 + 0.35 * E$	176%	198%
	NL	$2.64 * E^{0.57}$	107%	148%
Accommodation and food (72)	Benchmark	$6.54 + 0.40 * E$	79%	114%
	NL-1	$4.22 * E^{0.35}$	57%	76%
	NL-2	$E^{0.40} * LV^{0.24}$	53%	80%
Retail (44-45)	Benchmark	$15.35 + 0.31 * E$	129%	299%
	NL	$5.64 * E^{0.37}$	78%	198%
	SLM-2	$1.03 * E^{0.69} * \prod_{i \in I} (E^{0.23 w_{ij}})$	87%	89%

Notes: *E*: employment; *WS*: width of front street; *LV*: land value

The results also show that the incorporation of the locational variable decreases the MAPE for the calibration dataset by about 10% for the construction sector. For the accommodation and food sector, the findings are more nuanced; while for the calibration dataset the incorporation of land value decreases the MAPE by 7%, for the validation dataset the MAPE increases by 5%. These findings suggest the need for further research on the transferability of models that include land-value. In the manufacturing, the wholesale, and the accommodation and food industries, the MAPE of the nonlinear models displays lower values for both the calibration and the validation datasets than for the benchmark linear models. In the case of retail, the nonlinear model produces a smaller MAPE than the benchmark model, and the MAPE for the validation dataset is significantly higher than the one for the calibration dataset. However, the application of the SLM-2 results in similar MAPEs for both datasets, suggesting that, as opposed to the Benchmark and the NL models, SLM-2 is externally valid. To confirm the internal validity of these models, the authors estimated a SLM using the validation dataset and found a model with the same specification of SLM-2. The model is shown below:

$$FTA_i = 1.10 * E_i^{0.18} * \prod_{j \in J} E_j^{0.92w_{ij}} \quad (9)$$

Despite the weak Moran's  $I$  found in Section 4.2, these findings confirm the presence of spatial autocorrelation in retail establishments. Although the coefficients reveal that the parameters of the models are sensitive to the calibration dataset, the validation analysis shows that incorporating a spatial employment lag term improves significantly the performance of the model and eliminates the possibility of biased coefficients. In essence, retail establishments FTA is affected by location: an establishment neighboring large employment peers, or in a zone with high density of retail establishments tend to attract more trips per employee than isolated establishments. An interesting analysis for further research is to assess these spatial effects in less dense cities, with less expensive land use and different urban geographies.

The application of these models requires disaggregated employment information, land-value information and geographical information. In the US, information on employment per industry sector is publicly available (U.S. Census Bureau, 2010a), though a disaggregation process is necessary; land-value information is also available (The City of New York, 2011); geographical variables can be measured from free satellite images, and the current availability of GIS software makes possible to account for spatial effects. In essence, collecting the data for the variables proposed can be done at low costs and, as shown, through their incorporation enhance the quality of FTA estimates.

## 6. CONCLUSIONS

This paper provides a background on freight trip generation modeling, proposes a new set of variables that relate the establishment to its location, and assesses the performance of novel explanatory variables. Using New York City (NYC) as a case study, the paper explores the relationship between these variables and freight trip attraction (FTA). The analysis shows that in spite of having the lowest employment, retail establishments tend to have the largest FTA;

wholesale establishment tend to be located in wider streets; and accommodation and food establishments tend to be located in the most expensive properties.

The results from the econometric analysis and the specification tests show that FTA is better modeled as a nonlinear function of employment and other locational variables. Moreover, the range of the coefficients for employment, between zero and one, reveals that larger establishments have higher FTA than small establishments; with FTA increasing at a diminishing marginal rate. These results imply that large establishments tend to generate proportionally (i.e., per employee) less FTA than small ones. As predicted by the theory, this is the result of logistical decisions and how higher amounts of freight attraction can be handled by a smaller FTA through an increase in shipment size and a change in the truck size. This trend is more accentuated for the manufacturing, and the accommodation and food sectors. In essence, the nonlinear relation between FTA and employment is one of the main findings from this research and its implementation will improve FTA estimates.

For the accommodation and food sectors, the establishment location also plays a key role, as the establishments located in premium properties tend to have high FTA. For the construction industry, establishments located in important streets tend to attract a higher FTA. For retail establishments, there is a notable spatial autocorrelation. The results show that retail establishments located in zones with high retail employment tend to have more FTA per employee than the ones located in zones with low retail employment.

The application of the models to external data shows that the models including locational variables, spatial effects and nonlinear specifications perform better than the benchmark models (Holguín-Veras et al., 2012); and that they are in general externally valid. The case of retail is noteworthy, because models that do not account for spatial autocorrelation have very low performance when applied to external data, suggesting their lack of transferability and the need for models that account for spatial autocorrelation.

This research sheds light on some key factors driving FTA. The findings show that the FTA of an establishment is better estimated using spatial effects, locational variables, and nonlinear specifications. In spite of the complexity required to deal with the spatial dimension, the availability of Geographic Information System (GIS) software and free software with satellite images like Google Earth opens the possibility for researchers to continue the exploration of innovative FTA models that incorporate locational and spatial effects. Moreover, this research open the possibility for transportation agencies and practitioners to use models that go beyond generic trip generation rates by taking into account some of the operational needs of each industry sector and capturing the effects of urban geography on their logistics decisions. The understanding and incorporation of locational and spatial factors on freight demand models will enhance urban freight planning, improve the accuracy of traffic impact studies and contribute towards more efficient urban freight policies.

## REFERENCES

- Akaike H (1987) Factor analysis and AIC. *Psychometrika* **52**, pp. 317-32
- Al-Deek HM (2001) Comparison Between Neural Networks and Multiple Regression Approaches for Developing Freight Planning Models with Specific Applications to Seaports. In *80th Annual Meeting of the Transportation Research Board*. Vol. 1763, Washington, D.C.
- Al-Deek HM, Johnson G, Mohamd A, and El-Maghraby A (2000) Truck Trip Generation Models for Seaports with Container and Trailer Operation. *Transportation Research Record* **1719**, pp. 1-9
- Anselin L (1988) In *Spatial Econometrics: Methods and Models*. Kluwer Academic Press, Dordrecht.
- Anselin L, Syabri I, and Kho Y (2006) GeoDa: an introduction to spatial data analysis. *Geographical analysis* **38**, pp. 5-22
- Bartlett RS and Newton WH (1982) Goods Vehicle Trip Generation and Attraction by Industrial and Commercial Premises. Transport and Road Research Laboratory (No. LR 1059 Monograph), pp.1-65. Crowthorne, Berkshire, United Kingdom. f

Bassok A, Johnson C, Kitchen M, Maskin R, Overby K, Carlson D et al. (2013) NCFRP Report 24: Smart Growth and Urban Goods Movement. In *National Cooperative Highway Research Program / National Cooperative Freight Research Program*, pp. 96, Transportation Research Board of the National Academies, Washington D.C.

Bastida C and Holguín-Veras J (2009) Freight Generation Models: Comparative Analysis of Regression Models and Multiple Classification Analysis. *Transportation Research Record* **2097**, pp. 51-61.

Brogan JD (1980) Improving Truck Trip Generation Techniques Through Trip End Stratification. *Transportation Research Record* **771**, pp. 1-6.

Campbell S, Jaller M, Sanchez-Diaz I, Holguin-Veras J, and Lawson C (2011) Comparison Between Industrial Classification Systems in Freight Trip Generation (FTG) Modeling. In *91st Annual Meeting of the Transportation Research Board*. Washington, D.C

DeVries JB and Dermisi SV (2008) Regional Warehouse trip production analysis: Chicago Metro Analysis. Illinois Center for Transportation, Chicago, Illinois.

Garrido RA and Mahmassani HS (2000) Forecasting freight transportation demand with the space-time multinomial probit model. *Transportation Research Part B* **34**, pp. 403-18.

Hartgen DT, Fields G, Layzell A, and San Jose E (2014) How Employers View Traffic Congestion: Results of a National Survey. *Transportation Research Record* 2319, pp. 56-66. 2012.

Holguín-Veras J (2006) Potential for Off-Peak Freight Deliveries to Congested Urban Areas, pp. 294, Rensselaer Polytechnic Institute, Albany, NY.

Holguín-Veras J, Jaller M, Destro L, Ban X, Lawson C, and Levinson H (2011) Freight Generation, Freight Trip Generation, and the Perils of Using Constant Trip Rates. *Transportation Research Record: Journal of the Transportation Research Board* **2224**, pp. 68-81.

Holguín-Veras J, Jaller M, Sánchez-Díaz I, Wojtowicz J, Campbell S, Levinson H et al. (2012) NCHRP Report 739 / NCFRP Report 19: Freight Trip Generation and Land Use. In *National Cooperative Highway Research Program / National Cooperative Freight Research Program*, pp. 165, Transportation Research Board of the National Academies, Washington D.C.

Holguín-Veras J, López-Genao Y, and Salam A (2002) Truck-Trip Generation at Container Terminals: Results from a Nationwide Survey. *Transportation Research Record* **1790**, pp. 89-96.

Holguin-Veras J, Sánchez-Díaz I, Lawson C, Jaller M, Campbell S, Levinson HS et al. (2013) Transferability of Freight Trip Generation Models. *Transport Research Record* **2379**, pp. 1-8.

Holguín Veras J, Wang C, Sánchez-Díaz I, Campbell S, Hodge SD, Jaller M et al. (2014) Fostering Unassisted Off-Hour Deliveries: The Role of Incentives. *Transportation Research Part A: Policy and Practice* (in press),

Hong J, Shen Q, and Zhang L (2013) How do built-environment factors affect travel behavior? A spatial analysis at different geographic scales. *Transportation* **41**, pp. 419-40.

Iding MHE, Meester WJ, and Tavasszy LA (2002) Freight trip generation by firms. In *42nd European Congress of the European Regional Science Association*. Dortmund, Germany.

Institute of Transportation Engineers (2008) In *Trip Generation 7th* (ed.), Institute of Transportation Engineers, Washington, D.C.

Kawamura K and Miodonski (2011) Examination of the Relationship between Land Use and Retail Freight Delivery. In *91st Annual Meeting of the Transportation Research Board*, pp. 13, Chicago.

Lawson C, Holguín-Veras J, Sánchez-Díaz I, Jaller M, Campbell S, and Powers E (2012) Estimated Generation of Freight Trips Based on Land Use. *Transportation Research Record: Journal of the Transportation Research Board* **2269**, pp. 65-72.

Moran PA (1950) Notes on continuous stochastic phenomena. *Biometrika* **37**, pp. 17-23.

Novak DC, Hodgdon C, Guo F, and Aultman-Hall L (2011) Nationwide Freight Generation Models: A Spatial Regression Approach. *Network and Spatial Economics* **11**

NYC Department of City Planning (2011) Zoning Districts. New York City Department of City Planning. <http://www.nyc.gov/html/dcp/html/zone/zonetext.shtml>. Accessed December 3 2011

Ogden KW (1992) In *Urban Goods Movement: A Guide to Policy and Planning*. Ashgate Publishing Company, Brookfield, VT

Ortúzar JD and Willumsen LG (2011) In *Modelling Transport*. 4th (ed.), pp. 608, John Wiley and Sons, New York

Ramsey JB (1969) Tests for specification errors in classical linear least-squares regression analysis. *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 350-71

Sánchez-Díaz I, Holguín-Veras J, and Ban X (2014) A Time-Dependent Freight Tour Synthesis Model. In *Transportation Research Board 93rd Annual Meeting*. Washington, D.C.

Smart Growth Network and ICMA (2002) Getting to smart growth: 100 policies for implementation. pp. 104, U.S. Environmental Protection Agency

Starkie DNM (1967) Intensity of Commercial Traffic Generation by Industry. *Traffic Engineering and Control* **7**, pp. 558-60

The City of New York (2011) New York Property Information. <http://nycprop.nyc.gov/nycproperty/nynav/jsp/selectbbl.jsp>. Accessed March 11 2011

U.S. Census Bureau (2010a) County Business Patterns. <http://www.census.gov/econ/cbp/>. Accessed April 2 2010

U.S. Census Bureau (2010b) North American Industry Classification System. <http://www.census.gov/eos/www/naics>. Accessed February 13 2014

U.S. Census Bureau (2011) Zip Code Business Patterns. [http://www.census.gov/epcd/www/zbp\\_base.html](http://www.census.gov/epcd/www/zbp_base.html). Accessed February 17 2014

Wagner T (2010) Regional traffic impacts of logistics-related land use. *Transport Policy* **17** (4), pp. 224-229.

Wood RT and Leighton R (1969) Truck Freight in the Tri-State Region. *Traffic Quarterly* **23**, pp. 323-40