Characteristics of Logistics Facilities Allocation, Size and Truck Generation by Tokyo Metropolitan Area Urban Freight Survey

Al Hanz Seiji Basa Lidasan, Shusaku Umeda, and Tetsuro Hyodo
Department of Logistics and Information Engineering, Tokyo University of Marine Science and Technology, 2-1-6 Etchujima, Koto-ku, Tokyo, 135-8533, Japan

Abstract: Logistics facilities should be considered in transportation plans and city plans because they generate/attract huge volumes of truck traffic. We utilized the 5th Tokyo Metropolitan Area Urban Freight Survey (TMAUFS) conducted in 2013 to analyze the relationship among land-use pattern, transportation demand, and transportation networks. This paper focuses on the questionnaire-survey to establishments and truck probe survey combined with the 1-km² mesh land-use data of the Japanese Statistics Bureau. Data from the TMAUFS are aggregated following a description of logistics facility allocation patterns. Transition patterns of logistics facilities including factories are analyzed through a comparison of the 4th (2003) and 5th (2013) TMAUFS. Specifically, we clarify the relationship between highway-construction and location of new logistics facilities. Truck-trip-generation models are examined by fitting conventional regression models and the Tobit model for a better fit. Finally, we summarize the required future policy measures forecasted by the developed models for transportation infrastructure systems.

Keyword — Logistics Facilities, Truck Trip Generation model, Urban Freight Survey, Sample Selection model, Tobit Regression

1. Introduction

Logistics and freight are seldom topics in transportation research. This highlights the importance of including logistics and freight systems as a subject matter in transportation research. One of the rationale is to account for the impacts of truck traffic produced by logistics facilities in transportation planning. While the relationship between land-use and transportation has been a well-researched subject, research that considers the dynamics of allocation of land-use and elements of logistics and freight networks are lacking. This paper aims to analyze the relationship between logistics facility allocation and truck trip generation by utilizing the 4th and 5th Tokyo Metropolitan Area Urban Freight Survey (TMAUFS) which were conducted in 2003 and 2013, respectively.

Specifically, we aim to relate land-use allocation and truck trip generation by formulating a Logistics Floor Area model and a Truck Trip Generation model using utility theory with land-use variables and other area characteristics as inputs to both models. Ultimately, the estimated models will be used to conduct sensitivity analysis on the effects of infrastructure and policy changes to the total logistics floor area and truck trip generation. Thus, our research objective is to analyze the effects of land-use policy changes and infrastructure improvements to logistics facility size and truck trip generation.

The structure of this paper is as follows: in Section 2, we briefly discuss TMAUFS data; In Section 3, we formally develop and estimate the Logistics Floor Area model and the Truck Trip Generation model using TMAUFS data and relate both models together to demonstrate their practical application. Finally, Section 4 concludes and summarizes this paper.

1.1 Review of Related Literature

Freight transport services are increasingly important for the regional competitiveness while freight traffic is a growing threat for urban sustainability (Lindholm & Behrends, 2012). Moreover, freight transport appears in the periphery of urban transport planners’ daily work, but they do neither know how nor have the capacity to tackle the issue (Lindholm & Behrends, 2012). There is also a lack of role models and inadequate monitoring, evaluation and dissemination of performed studies and projects, which makes it hard to follow good experiences as well as to avoid the bad examples (Lindholm & Behrends, 2012). Also, with the recent interest in freight planning, a concern for the phenomenon of logistics sprawl, i.e., the spatial de-concentration of logistics facilities and distribution centers in metropolitan areas, has been investigated (Woudsma, et al., 2015). However, the conduct of research on the

---

1 Corresponding author’s e-mail: hanzlidasan@gmail.com
dynamics of the logistics sector, especially in the context of urban transportation planning, is mired with the challenge of collecting reliable data (Sanchez-Diaz, 2016). On the other hand, the Tokyo Metropolitan Government has been conducting urban freight surveys roughly every 10 years since 1972 (Hyodo, 2017). The Tokyo Metropolitan Area has seen a de-centralization of logistics facilities or logistics sprawl since the 1970’s and is well documented by Tokyo Metropolitan Government’s urban freight survey and has informed policies for the urban freight and logistics sector in Tokyo (Hyodo, 2017).

The relationship between land-use and transportation is very-well established and a well-researched area (Geurs & van Wee, 2004; James et al. 1972; Wegener, 2004; Newman & Kenworthy, 1996) that it is already standard to consider land-use-transport (LUT) interactions in city/town planning and regional planning. However, we feel the lack of research that tackle the interaction between logistics land-use and transportation especially when utilizing models for policy analysis. Only a few has attempted to do so due to the complexities of the logistics sector (Hesse, 2002; Hesse, 2004; Wagner, 2010).

Previous research on logistics facilities distribution focus on decisions where to locate. For instance, evaluation criteria for the location selection of city logistics centers was formulated by combing economic, environmental, and social sustainability indicators through a fuzzy multi-attribute group decision making method (Rao et al. 2015). (Lindsey et al. 2014) used an econometric approach to evaluate longitudinal data of metropolitan markets wherein a methodology was developed to rank 20 metropolitan markets from 1997 to 2007 based on their potential for industrial space using macroeconomic, demographic, and freight flows as input variables. In relation to econometric modeling, (Woudsma et al. 2008) applied a spatial-temporal modeling approach to quantify the effects of transportation system performance on the patterns of logistics land-use. (Woudsma et al. 2015) investigated logistics sprawl and its relation to facility identification and location. (Sakai et al. 2016) presented the historical transition of logistics facilities in TMA from 1980 to 2003 which revealed that the asset pricing bubble in Japan during the period of 1986 to 1991 was a significant factor in the decentralization of logistics facilities into the suburbs. (Iwakata et al. 2015) highlighted the importance of accessibility to interchanges and expressways for mega distribution centers in TMA. (Hong, 2007) found that location of foreign logistics firms in Chinese cities depended on transport conditions in terms of roadway, railway and waterway, as well as market size, labor quality, agglomeration economies and government incentives.

Truck trip generation, on the other hand, has been modeled through various methods in the past. The most basic of which is through the use of trip generation rates (Kulpa, 2014; Sorratini & Smith 2000) which determines the number of truck trips generated per unit of independent variable (e.g., number of trips per number of employed persons). Multiple linear regression has also been used in numerous papers modeling truck trip generation either to develop generation rates or directly forecast truck trip generation (Tadi & Balbach, 1994; Holguín-Veras et al. 2002; El-maghraby, 2000; Sorratini & Smith, 2000; Kulpa, 2014). Truck trip generation models fall under the vehicle-based models as opposed to commodity based models in road freight transport trip generation modeling (Kulpa, 2014). While these methods have been the standard in urban-transport planning, the modeling of truck trip generation focused on certain types of land-use or facility one at a time (Tadi & Balbach, 1994; Holguín-Veras et al. 2002). This is especially flawed when considering mixed land-use patterns particularly in the regional level.

There is disconnect between research on logistics land-use and other land-use classifications and actual truck trip generation wherein past research consider one aspect independent of the others and vice-versa. Therefore, the contribution and the aim of this paper is to demonstrate that we not only link land-use and transport in the context of freight transport (vehicle-based) but simultaneously account for allocation and size of logistics facilities and truck trip generation considering all land-use classifications available in the data.

1.2 Transition of Logistics Facilities from the 4th to 5th Survey

In this section, we briefly discuss and show the general transition of the distribution of freight facilities in TMA from the 4th TMAUFS (2003) to the 5th TMAUFS (2013). However, only prefectures that are both in the 4th and 5th surveys are considered for consistency as shown in Figure 1 below.
We focus on logistics centers which is one of the facility types defined in the main part of the TMAUFS. The distribution of all facility types surveyed in TMA and the distribution of all logistics facility respondents in TMA are shown in Figure 2 below.

To understand the transition from the 4th to the 5th survey, we specifically focus on the statistics Total Number of Logistics Facilities and the Total Floor Area (m²) of Logistics Facilities in a geographical unit. Here we define a geographical unit as the secondary mesh unit (about 10-km²) as defined by Japanese Standards. Furthermore, we define Number of Logistics Facilities as the sum establishments per geographical unit transformed by an expansion factor. The Total Floor Area of Logistics Facilities is defined as the average in a geographical unit considering an expansion factor to reflect relative magnitudes.

Figure 3 below shows the increase and decrease in the Number of Logistics Facilities and Total Floor Area of Logistics Facilities. The relative sizes of the circles indicate the maximum absolute value of their respective percentage changes; black circles indicate an increase and red circles indicate a decrease from the 4th survey to the 5th survey.

Figure 1. TMA freight survey areas: (a) 4th TMA freight survey (left), and (b) 5th TMA freight survey (right)

Figure 2. Distribution of respondents from all facility types (left) and Logistics facilities (right)

Figure 3. Increase and decrease in logistics facilities and total floor area

□: New areas added for 5th survey
As seen in Figure 3, the Number of Logistics Facilities in Tokyo is observed to have decreased and by contrast, have increased in the suburbs (e.g., North Saitama, South Ibaraki). We suppose that this is due to the improved network of highways during the 5th TMAUFS relative to when the 4th TMAUFS was conducted (e.g., the completion of the Metropolitan Inter-City Expressway). Furthermore, Total Floor Area of Logistics Facilities is observed to have increased around Tokyo Bay and the Tohoku Expressway. We suppose that this is due to logistics facilities in Japan moving toward increasing in size due to consolidation of functions and services which, thus, leading to the decrease in small-scale logistics facilities and increase in large-scale logistics facilities.

2. DATA ABSTRACT AND METHODOLOGY

The Tokyo Metropolitan Area Urban Freight Survey (TMAUFS) is an urban freight survey conducted by the Tokyo Metropolitan Transportation Planning Commission of the Tokyo Metropolitan Government which started in 1972 and has since been conducted roughly every ten (10) years (Hyodo, 2017).

2.1 Questionnaire Survey to Logistics Establishments

The questionnaire survey to establishments of the TMAUFS’ primary purpose is to gather data on logistics facilities such as the location, scale, function, freight characteristics, inventory of goods, volume of goods inward/outwards, characteristics and information about trucks used by the facility, and OD information. The questionnaire survey was conducted by mailing-out the survey forms to over 140,000 establishments and among these, 44,000 forms were mailed back which results in a response rate of 31% (Hyodo, 2017).

2.2 Truck Probe Data

In this section, we discuss the truck probe data from the 5th TMAUFS. Among three (3) sources of truck probe data from the 5th TMAUFS we utilize the data that was collected for one (1) week, from October 6th (Monday) to October 12th (Sunday), by an OBU (On-Board Unit) manufacturer. As there was no specific sampling scheme applied, the data source with the most number of samples (22,995) collected for one (1) week was used to consider variations in behavior that may occur on different days of the week. The truck-probe data is categorized into four (4) levels: small, medium, large, and tractor; each level is categorized based on gross maximum weight in tons. The truck trip generation data is an aggregation of truck trips generated at the tertiary level mesh; that is, one data point of truck trip generation represents truck trips generated per 1-km2 area. However, as this research will specifically focus on large-sized trucks namely, large and tractor trucks, only the graphs of large and tractor trucks will be presented.
Figure 4 shows the common logarithm of one-week truck trip generation in Tokyo Metropolitan Area of (a) Large Trucks (left), and (b) Tractor Trucks (right).

Figure 4 shows the common logarithm of one-week truck trip generation of large trucks and tractor trucks. It can be observed that for large trucks (left) and tractor trucks (right), there are large concentrations of trip generation along Tokyo Bay. This is expected because the Port of Tokyo and the Port of Yokohama are located along Tokyo Bay as well as logistics facilities servicing these ports. These ports cater to international shipping and container ships and as such, their operations generate heavy volumes of truck traffic inbound and outbound of Tokyo Bay. Furthermore, a large concentration of tractor trucks trips generation (right) can be observed in the eastern region of TMA, specifically in the areas of Kashima City and Kamisu City. This is due to Kashima City and Kamisu City being part of the Kashima Rinkai Industrial Zone where about 1,500 factories of chemical, petrochemical, specialty chemical plants, steel, and oil refineries are located. Accompanying the Kashima Rinkai Industrial Zone is the Port of Kashima which further contributes to tractor trucks trip generation from the eastern region of TMA due to inbound and outbound international shipping containers.

Figure 5. (a) Aggregation of Large Trucks and Tractor Trucks’ one-week truck trip generation in Tokyo Metropolitan Area; (b) Scatterplot of truck trip generation against the number of logistics facilities and total logistics facility floor area.
Considering the size and nature of freight being transported by large trucks and tractor trucks, we further aggregate their one-week truck trip generation by combining their respective one-week trip generation. Figure 5 shows the combined one-week truck trip generation of large trucks and tractor trucks. We observe that large trucks and tractor trucks trip generation are concentrated around Tokyo Bay where the Ports of Tokyo and Yokohama are located as well as within proximity of ring roads (circumferential highways) and radial roads in the suburbs where numerous logistics facilities, warehouses, and factories are located. It is important to note that there are residential areas located in high-concentration trip generation areas of large trucks and tractor trucks especially along ring roads and radial roads; this exposes residents to safety risks. Furthermore, the high gross maximum weight of large trucks and tractor trucks exacerbate the deterioration and accelerate the wear-and-tear of roads which increase road maintenance costs. Based on the scatterplot of truck trip generation against logistics facility count and total floor area in Figure 5 (b) above, we also observe positive correlation between trip generation of tractor and large trucks with the number of logistics facilities and the total floor area of logistics facilities in an area. Given the observed positive correlation and the safety risks to residents and accelerating deterioration of roads caused by large trucks and tractor trucks, this paper will specifically focus on the trip generation of large trucks and tractor trucks in analyzing the dynamics of logistics facility allocation and size, land-use, and truck trip generation in Section 3.

2.3 Methodology

To achieve the research objective of analyzing how land-use policy changes and infrastructure improvements affect logistics facility size and truck trip generation, we propose the following framework for analysis shown in Figure 6. Model estimation is highlighted and contained in the polygon with dashed lines and shall be conducted before conducting the policy sensitivity analysis. The analysis starts with quantitatively understanding the allocation and size of logistics facilities in the Tokyo Metropolitan Area (TMA) by estimating a model for logistics facility allocation and size which is the Sample Selection Model (Tobit Type II) (Heckman, 1979). The Sample Selection Model will simultaneously determine the decision to locate a logistics facility in a geographical unit of not and the size of the logistics facility in square-meter (m²) units. This is followed by the estimation of the Truck Trip Generation Model which determines the total truck trips generated from a geographical unit through application and estimation of the Tobit Type I Model (Tobin, 1958). After obtaining the parameter estimates for both models, policy sensitivity analysis will be conducted in the form of changes in levels of service in a geographical unit or mesh. Thus, allowing for understanding how land-use policy and infrastructure improvements affect logistics facility size and truck trip generation.

![Figure 6. Framework for Analysis](image-url)
2.4 Tobit models

Tobit models were utilized due to the structural zeroes in the data, which is a consequence of the TMAUFS’s level of detail or data resolution, i.e., one data sample or data point corresponds to a 1-km² area, especially in the dependent variables being modeled: the decision to locate logistics facilities (binary), the total floor area of logistics facility (continuous), and total truck trips generated (continuous). As mentioned above, modeling the decision to locate logistics facilities and the size of the logistics facilities are simultaneously estimated while modeling the total truck trips generated are estimated in another model. The following sub-sections will introduce the Tobit models applied in this paper.

2.4.1 Sample selection model (Tobit type II)

The Sample Selection model (Heckman, 1979) also known as Tobit Type II model (Amemiya, 1984) is based primarily on two equations, namely the Selection equation (1) and the Outcome equation (3) and a set of conditions (2) and (4). The Tobit Type II model construction is discussed below within the context of the Logistics Facility Floor Area model as follows:

\[
y_{is}^s = \beta^s x_{is}^s + \epsilon_i^s \quad \text{ (Selection equation),} \tag{1}
\]

\[
y_{is}^s = \begin{cases} 0 & \text{if } y_{is}^s < 0 \\ 1 & \text{if } y_{is}^s \geq 0 \end{cases}, \tag{2}
\]

where the dependent variable \( y_{is}^s \) is the untransformed value of the total number of logistics facilities (Log.num) in the 1-km² mesh in the observed data;

\[
y_{is}^o = \beta^o x_{is}^o + \epsilon_i^o \quad \text{ (Outcome equation),} \tag{3}
\]

\[
y_{is}^o = \begin{cases} 0 & \text{if } y_{is}^o = 0 \\ \ln(\text{Log.areaE} + 1) & \text{if } y_{is}^o = 1 \end{cases} \tag{4}
\]

where \( y_{is}^o \) is the natural logarithm of the total logistics floor area scaled to size plus one,

\[
y_{is}^o = \ln \left( \text{Log.areaE} + 1 \right) \quad \text{and}
\]

\[
\left( \begin{array}{c}
\epsilon_i^s \\
\epsilon_i^o
\end{array} \right) \sim N \left( \begin{array}{c}
0 \\
0
\end{array} , \begin{bmatrix}
1 & \rho \\
\rho & \sigma^2
\end{bmatrix} \right), \tag{5}
\]

where equations (1) and (3) are the Selection and Outcome equations, respectively. We only observe the value of the latent outcome \( y_{is}^o \) in equation (3) only if the latent selection variable \( y_{is}^s \) is positive as described in the conditions in equations (2) and (4). Furthermore, it is assumed that the error terms follow a bivariate normal distribution as shown in equation (5) above. To estimate the Sample Selection model, the Maximum-Likelihood (ML) method is used. Equation (6) shows the likelihood function to be maximized as follows:

\[
L = \prod_i \Phi \left( -\beta^s x_i^s \right) \left( \frac{1}{\sqrt{1-\rho^2}} \phi \left( \frac{\beta^o x_i^o - \beta^s x_i^s}{\sqrt{1-\rho^2}} \right) \phi \left( y_{is}^o - \beta^o x_i^o \right) \right)^{y_{is}^s}, \tag{6}
\]

and the expected value of the outcome is shown in equation (7) as follows:

\[
E \left[ y_{is}^o | y_{is}^s \geq 0 \right] = \beta^o x_i^o + \rho \sigma \frac{\phi \left( \beta^o x_i^o \right)}{\Phi \left( \beta^s x_i^s \right)} \tag{7}
\]

2.4.2 Tobit regression model (Tobit type I)

The Tobit model (Tobin, 1958) is a linear regression model (8) with a latent dependent variable that is governed by a set of conditions (9). The Tobit regression formulation is presented below in the context of the Truck Trip Generation model.
where the latent variable \( y_i^* \) is the natural logarithm of the total truck trips generated plus one, \( \left[ y_i^* = \ln(y_i + 1) \right] \); and

\[
y_i = \begin{cases} 
  y_i^* & \text{if } y_i^* > 0 \\
  0 & \text{if } y_i^* \leq 1 
\end{cases},
\]

where the \((k = 9)\) input variables are the same as in Table 2 of the Sample Selection model above. Also, included are independent and normally distributed error terms, \( \varepsilon_i \), with mean 0 and standard deviation, \( \sigma \), in the latent formulation. Instead of observing \( y_i^* \) directly, which in this case, are truck trip generation (zero or non-zero), we observe \( y_i \) as in equation (9) above; we observe truck trip generation \( y_i^* \) if it is positive, and 0, otherwise. It is clearly seen here that the 1-km\(^2\) mesh data that have zero generation are properly considered in the Tobit model with the “potential” truck trip generated as a function of the input variables. To estimate coefficients \( \beta_i \), we maximize the likelihood function of equations (8) and (9) as shown in equation (10) below.

\[
L = \prod_i \frac{1}{\sigma} \phi \left( \frac{y_i - X_i \beta}{\sigma} \right) \left[ 1 - \Phi \left( \frac{X_i \beta}{\sigma} \right) \right]^{y_i^* - d_i}.
\]

### 3. LOGISTICS FACILITY FLOOR AREA AND TRIP GENERATION ANALYSIS

#### 3.1 Logistics facility floor area model

As observed in the scatterplot in Figure 5 (b), we focus on the relationship of truck trip generation and logistics facility count and total floor area. Given that total logistics facility floor area has a slightly higher correlation to truck trip generation than logistics facility counts, we deal with the former in developing the model, namely the Logistics Facility Floor Area (LFFA) model. In this section, we first develop a LFFA model to analyze factors that affect total floor area of logistics facilities in an area, specifically in a 1-km\(^2\) mesh, before proceeding to Truck Trip Generation model formulation which will be discussed in Section 3.2.

The unit of analysis for the LFFA model is the tertiary level mesh or geographical unit; that is, one data point is a 1-km\(^2\) area as defined by Japanese Standards which contains data collected from the 5th TMAUFS (2013).

A conventional multiple regression model was initially estimated for the LFFA model, where we set the dependent variable \((y^*)\) as the natural logarithm of the total logistics facility floor area scaled to size (Log.areaE) plus one, \([y^* = \ln(\text{Log.areaE} + 1)]\), with the results shown in Table 1. However, due to peculiarity of the data where numerous zero values were observed in the dependent variable, the estimated multiple regression model proved to be not a good fit for the data based on its low Adjusted R-squared (0.2744). Furthermore, using the estimation results in Table 1 for prediction of Total LFFA results in underestimated values, which are due to numerous zero values in the data. The structural zero in the data, especially in the dependent variable (Total Logistics Facility Floor Area), is a consequence of the level of detail or resolution of the unit of analysis. Given the scope of Tokyo Metropolitan Area and level of detail of the unit of analysis (i.e., 1-km\(^2\)), not all 1-km\(^2\) area will have logistics facilities located in them and correspondingly a measure of total floor area.

| Estimate | Std. Error | t-value | Pr(>|t|) |
|----------|------------|---------|----------|
| (Intercept) | -14.091 | 3.478 | -4.05 | 5.3E-05 |
| Population | -0.261 | 0.032 | -8.10 | 9.2E-16 |
| Working Population | 0.632 | 0.237 | 2.67 | 7.6E-03 |
| ACC.manuf | 0.052 | 0.013 | 3.94 | 8.6E-05 |
| ACC.cbd | 1.019 | 0.109 | 9.36 | <2e-16 |
| ICdistance | -0.030 | 0.020 | -1.50 | 1.3E-01 |
| TokyoPortDist | -0.018 | 0.004 | -4.17 | 3.1E-05 |
Because of the numerous zero values in the data, i.e., no logistics facilities located in an area, there is a need to effectively estimate a model without having to exclude data with no logistics facilities so that data in those areas are still considered. We do this by applying the Sample Selection model (Heckman, 1979) also known as Tobit Type II model (Amemiya, 1984) which was discussed in Section 2.4.1 above.

The input variables for the Sample Selection model are described in Table 2. We emphasize here the inclusion of land-use variables as well as accessibility to manufacturing, CBDs, distance to the closest expressway interchange, and distance to the Port of Tokyo.

We utilize the Sample Selection package in the R programming language (Toomet & Henningsen 2008) to estimate the model. Table 3 below shows the results of the Sample Selection model estimation for the LFFA model. Two other variations of the Sample Selection models were estimated, first of which include all variables in both the selection and outcome equations, and the second, which is like Model 1 only that it includes all variables in the outcome equation. However, the estimation results showed parameter estimates that do not satisfy the conditions for utility maximization; thus, only Model 1 and Model 2 shown in Table 3 are further discussed in this paper.

Table 2. Description of variables used in the Sample Selection model

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Population covered by the mesh</td>
</tr>
<tr>
<td>Working Population</td>
<td>Working population (daytime) covered by the mesh</td>
</tr>
<tr>
<td>ACC.manuf$^2$</td>
<td>Accessibility index to manufacturing sites</td>
</tr>
<tr>
<td>ACC.cbd$^3$</td>
<td>Accessibility index to CBDs</td>
</tr>
<tr>
<td>ICDistance.km</td>
<td>Distance of the mesh to the closest interchange</td>
</tr>
<tr>
<td>TokyoPortDis.km</td>
<td>Distance to the Port of Tokyo</td>
</tr>
<tr>
<td>landprice.yen</td>
<td>Average land price in the mesh</td>
</tr>
<tr>
<td>roadArea.m2</td>
<td>Total road area in the mesh</td>
</tr>
<tr>
<td>vacantArea.m2</td>
<td>Total vacant area in the mesh</td>
</tr>
<tr>
<td>residence.rate</td>
<td>Share of residence land-use</td>
</tr>
<tr>
<td>commercial.rate</td>
<td>Share of commercial land-use</td>
</tr>
<tr>
<td>quasiIndustrial.rate</td>
<td>Share of quasi-industrial land-use</td>
</tr>
<tr>
<td>industrial.rate</td>
<td>Share of industrial land-use</td>
</tr>
<tr>
<td>r.industrial.rate</td>
<td>Share of restricted industrial land-use</td>
</tr>
</tbody>
</table>

$^2$ \( ACC\text{\text{manuf}} = \sum_{j=1}^{J} M_j e^{-\ln(d_{ij})} \), where \( M_j \) is the total value of industrial products in area \( j \) and \( d_{ij} \) is the shortest distance between area \( i \) and area \( j \).

$^3$ \( ACC\text{\text{cbd}} = \sum_{j=1}^{J} B_j e^{-0.5\ln(d_{ij})} \), where \( B_j \) is the total working population in area \( j \) and \( d_{ij} \) is the shortest distance between area \( i \) and area \( j \).
We highlight from the estimation results in Table 3 that in the selection equation results, accessibility to manufacturing areas (ACC.manuf), accessibility to CBDs (ACC.cbd), and distance to the Port of Tokyo are statistically significant and are consistent with utility maximization based on their parameter signs. This indicates that accessibility to manufacturing areas and accessibility to CBDs of an area could affect the choice of location for a logistics facility. Population is also statistically significant and negative in parameter sign. This indicates that the
higher the population in an area, the less likely that a logistics facility will be located there. Taking a look at land-use variables in the outcome equation results, the share of residential land-use (residence.rate) is consistent in sign (negative) and is statistically significant. Although the other land-use variables seem not significant apart from quasi-industrial land-use (quasiIndustrial.rate), we left it as it is in the model to consider land-use in the model.

We apply the formula for the expected value of the outcome as in equation (7) above using the estimated parameters in Table 3 and the averages of the dependent variables to compute the calibrated average LFFA. Table 4 below shows the average calibrated values of LFFA including the expected values from the conventional multiple regression previously mentioned. Furthermore, we included sensitivity analysis in the form of policy changes and infrastructure improvements such as decreasing the distance to the closest expressway interchange and increasing the share of quasi-industrial, industrial, and restricted industrial land-use.

As shown in Table 4, as we decrease the accessibility distance of a 1-km2 mesh to the closest expressway interchange by 5-km, Model 1 and Model 2 result in an 8% and 7.8% increase in the Logistics Facility Floor Area, respectively. Furthermore, as the share of quasi-industrial, industrial, restricted industrial land-use are increased by 5%, Model 1 and Model 2 result in increases of 19.7% and 13.3%, respectively. This is invaluable information especially to city planners and road infrastructure managers because the impacts of such improvements or policy changes are quantified.

We can also observe from the rightmost column of Table 4 the expected values from the results of the conventional multiple regression which clearly underestimate values for the average LFFA (~3-m2). The difference between the Sample Selection model and the conventional multiple regression model is evident. This is because of the feature of Sample Selection Model that can better deal with zero values statistically.

### Table 4. Sensitivity analysis of the Expected Values of the Outcome

<table>
<thead>
<tr>
<th>Sensitivity Analysis</th>
<th>Tobit Type II</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Multiple Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m²</td>
<td>% increase</td>
<td>m²</td>
<td>% increase</td>
</tr>
<tr>
<td>Average Area of Logistics Facilities</td>
<td>4,922</td>
<td>5,035</td>
<td>2.89</td>
<td></td>
</tr>
<tr>
<td>Distance to closest IC decreased by 5 km</td>
<td>5,317</td>
<td>8.0%</td>
<td>5,426</td>
<td>7.8%</td>
</tr>
<tr>
<td>Share of quasi.Ind, Ind, res.Ind increased by 5%</td>
<td>5,891</td>
<td>19.7%</td>
<td>5,704</td>
<td>13.3%</td>
</tr>
</tbody>
</table>

Instead of the average increase in total floor area, we were also able to evaluate the increase in total floor area for each 1-km2 mesh, this time, considering the completion of all the planned and under-construction ring (loop) roads of the “Three Loop Roads of the National Capital Region” illustrated in Figure 7 and visualize the percentage increase in Total Floor Area as shown in Figure 8 below. These are reflected as Future Level of Service (LOS), that is, the changes in the variables of accessibility index for manufacturing areas and CBDs (ACC.manuf and ACC.cbd) as well as the distance to the closest expressway interchange (ICdistance.km) when all the ring roads are completed and operational.
At a glance, we can see that there will be significant increases in Total LFFA in the north and north-eastern region of TMA. Referring back to Figure 7, we were able to evaluate the impact of the completion of the remaining portions of Ken-O road (Metropolitan Inter-City Expressways/National Capital Region Central Loop Road) in the north to north-eastern region of TMA (dashed blue lines) which, as shown in Figure 8, will result in the increase in Total LFFA in surrounding areas.

Although, the percentage increase in total floor area of Model 1 and Model 2 in Figure 8 are almost similar, Model 2 can be considered relatively better given its higher Log-Likelihood of -2811.753 compared to the Log-likelihood of Model 1 of -2830.29. Furthermore, the Akaike Information Criterion (AIC), which determines the level of predictive error of the model, is lower for Model 2, hence the better model overall especially for out-of-sample predictions; this can be attributed to the inclusion of land-use variables in the Selection portion of...
Model 2 which contributes to a better estimated model.

3.2 Truck trip generation model

In this section, we formulate truck trip generation models, specifically for tractor trucks and large trucks. We emphasize in the model formulation the relationship of land-use composition and allocation in an area to its truck trip generation by including the corresponding shares of different land-use classifications in an area as inputs to the truck trip generation model.

Furthermore, we also take into consideration the relationship of the establishment of logistics facilities in an area to its truck trip generation by including as inputs to the model the total number of logistics facilities and the total floor area occupied by logistics facilities in the area.

A conventional multiple linear regression model was initially estimated for truck trip generation with the dependent variable ($y^*$) as the natural logarithm of the total truck trips generated ($y$) plus one in a 1-km² mesh, \[ y^*_i = \ln (y_i + 1). \] However, the estimation results proved to be a poor fit to the data due to their low adjusted R-squares shown in Table 5 below. Furthermore, even omitting data with zero truck trip generation does not result in a better estimated model (Excluding Zero Generation).

Table 5. Estimation results of conventional multiple regression analysis

<table>
<thead>
<tr>
<th></th>
<th>Excluding Zero Generation</th>
<th>Including Zero Generation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Estimate</td>
<td>Std. Error</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>2.489</td>
<td>0.315</td>
</tr>
<tr>
<td>Log.numE</td>
<td>0.058</td>
<td>0.041</td>
</tr>
<tr>
<td>Log.areaE</td>
<td>0.040</td>
<td>0.009</td>
</tr>
<tr>
<td>Population</td>
<td>-0.053</td>
<td>0.004</td>
</tr>
<tr>
<td>ICdistance.km</td>
<td>-0.060</td>
<td>0.014</td>
</tr>
<tr>
<td>portDistance.km</td>
<td>-0.320</td>
<td>0.030</td>
</tr>
<tr>
<td>Landprice.yen</td>
<td>0.087</td>
<td>0.019</td>
</tr>
<tr>
<td>residence.rate</td>
<td>-0.052</td>
<td>0.059</td>
</tr>
<tr>
<td>commercial.rate</td>
<td>-0.247</td>
<td>0.145</td>
</tr>
<tr>
<td>quasiIndustrial.rate</td>
<td>0.757</td>
<td>0.115</td>
</tr>
<tr>
<td>industrial.rate</td>
<td>1.053</td>
<td>0.164</td>
</tr>
<tr>
<td>r.industrial.rate</td>
<td>1.456</td>
<td>0.090</td>
</tr>
</tbody>
</table>

Moreover, the conventional multiple linear regression underestimates predictions of truck trip generation (almost 0 generation). Again, this is due to the numerous zero truck trip generation values in the 1-km² mesh data which is, like the LFFA model, a consequence of the level of detail or resolution of the unit of analysis. Thus, due to the scope of Tokyo Metropolitan Area and level of detail of the unit of analysis (i.e., 1-km²), not all 1-km² will have truck trip generation in them due to no logistics facility located in the area.

Because of the numerous zero values in the data, there is a need to effectively estimate a model without excluding data with zero generation so that such areas are still considered and appropriate model applied is the Tobit regression as discussed in Section 2.4.2.

Table 6 below shows the estimation results of the Tobit regression model for truck trip generation.
Based on the results of Table 6, we can see that all estimates are statistically significant and satisfy the conditions for utility maximization; in other words, the signs (positive or negative) of the estimates/parameters are consistent with the conditions for utility maximization. We specifically note that the estimate for total land area of logistics facilities (Log.areaE) is highly statistically significant. This is consistent with the findings that historically in TMA, the number of logistics facilities are generally decreasing and the average land area of logistics facilities are increasing due to consolidation of functions and diversification of business operations of logistics facilities. This means that as the total land area of logistics facilities in an area increases, the trip generation of tractor and large trucks also increases. The increase in truck trip generation, especially for tractor and large trucks, as the total land area of logistics facilities increase can possibly be attributed to the increase in allowable parking space for trucks of all types in the logistics centers. Hence allowing for a larger fleet of trucks and generally a larger scale of logistics operations. Furthermore, as the distance to the closest expressway interchange (ICdistance.km) and the distance to the Port of Tokyo (portDistance.km) decreases, the generation of tractor and large trucks increases; this is consistent with the findings in the LFFA model that the closer an area is to an expressway interchange and to the Port of Tokyo, the higher the likelihood that a logistics facility will locate in that area.

Considering the land-use input variables, all land-use classifications are statistically significant and positively affect trip generation of tractor and large trucks. It is not surprising that the three types of industrial land-use (i.e. quasi, industrial, and restricted) increase truck trip generation of tractor and large trucks because these are areas where logistics facilities are built and are being operated. On the other hand, although the estimate for commercial land-use is positive and statistically significant for tractor and large trucks, the magnitude is relatively low compared to the three (3) industrial land-uses. This is probably due to narrow roads and presence of many pedestrians and shoppers as well as private cars in commercial areas making it difficult for tractor and large trucks to maneuver especially during time-constrained operations. However, we note that the share of residence land-use is statistically significant and positively affect trip generation of Tractor and Large trucks. This is most likely due to the prevalence of logistics facilities, factories, and warehouses in the suburbs where a lot of residential areas are also located. Furthermore, the existence of mixed land-use patterns (e.g., quasi-industrial land-use) wherein residential and industrial structures are mixed together in an area might contribute to the significance of the residential share to trip generation of tractor and larger trucks.

Finally, to link the results of the LFFA model to the Truck Trip Generation model, we evaluate the average truck trip generation per day from the calibrated average LFFA in Section 3.1 and the estimated parameter for total logistics facility floor area (0.13) from the Truck Trip Generation model in Table 6. The proposed equation for calculating the average trip generation per day is based on the survey and data characteristics, such as the number of days in a week the data was collected (no. of days in a week) and the share of the type of trucks considered.
(expansion factor), and the results of the LFFA model namely, the estimated average floor area of logistics facilities:

\[
\text{trip gen per day} = \text{rate} \times \frac{\text{floor area}}{\text{no. of days in a week}} \times \text{expansion factor},
\]

(11)

where:

- **trip gen per day:** is the average Tractor and Large Truck trips generated per day per Logistics Facility with "floor area";
- **floor area:** is the average floor area (in 1,000's m\(^2\)) of Logistics Facilities evaluated from the Logistics Facilities Floor Area model in Section 4.1;
- **rate:** is the parameter estimate for the Logistics Floor Area (Log.areaE) variable in the Truck Trip Generation Model;
- **expansion factor:** \(1/(\text{share of Tractor and Large Trucks})\) in the data = 1/0.017;
- **no. of days in a week:** the total number of days the data was collected in a week (7 days)

Table 7. Sensitivity of Average Truck Trip Generation per Day

<table>
<thead>
<tr>
<th>Tobit Type II</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Multiple Regression</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>m(^2)</td>
<td>Average Truck Generation per day</td>
<td>% increase</td>
</tr>
<tr>
<td>Average Floor Area of Logistics Facilities</td>
<td>4,922</td>
<td>5.34</td>
<td></td>
</tr>
<tr>
<td>Distance to closest IC decreased by 5 km</td>
<td>5,317</td>
<td>5.76</td>
<td>8.0%</td>
</tr>
<tr>
<td>Share of quasi.Ind, Ind, res.Ind increased by 5%</td>
<td>5,891</td>
<td>6.39</td>
<td>19.7%</td>
</tr>
</tbody>
</table>

Table 7 shows the sensitivity analysis of average truck generation per day for a logistics facility with average floor area based on the estimation results of both the LFFA model and the Truck Trip Generation model. Again, we can see that the conventional multiple linear regression analysis gave underestimated results. In contrast, we observe realistic output from the combined LFFA model and Truck Trip Generation model.

4. CONCLUSION

We could show the transition and dynamics of logistics facility development from the 4th TMAUFS (2003) to the 5th TMAUFS (2013) focusing on total number of logistics facilities and total floor area of logistics facilities in an area. The results of the analysis showed that the total number of logistics facilities is decreasing especially in the central TMA. Furthermore, total number of logistics facilities decreased in quasi-industrial land-use areas along with little to no changes in restricted-industrial land-use areas. These observations indicate that there might be other suitable areas and factors being considered by logistics managers where logistics facilities would be developed. The decrease in number of logistics facilities might also be due to consolidation of functions and services of different logistics facilities which lead to logistics facilities with larger floor areas. We also presented the Truck Probe data portion of the TMAUFS and observed that there are areas where truck trip generation are concentrated depending on the category of trucks. For instance, small and medium trucks generation are mostly clustered around the center of TMA, where most CBDs are located. On the other hand, large and tractor trucks are mostly concentrated around Tokyo Bay, around ports as well as along the periphery of expressways in the suburbs. The general difference observed as to where truck trips are generated indicates that the location of logistics facilities might have an...
influence to where trucks are originating especially for large and tractor trucks. Hence, we formulated the LFFA model and the Truck Trip Generation model focusing on tractor and large trucks due to safety risks that they impose on the environment especially to people and road infrastructure.

We could show in the LFFA model that accessibility to CBDs and manufacturing areas increases the probability of logistics facilities locating in an area. Also, decreasing the distance to expressway interchanges and to Port of Tokyo through infrastructure improvements such as the completion of the Ken-O expressway at the western area of TMA, increases the probability that logistics facilities will be located and developed in an area. In terms of land-use allocation in an area, the results showed that the share of residential land-use in a 1-km² area is significant in decreasing the total floor area of logistics facilities in that area. The negative effect of the share of residential land-use makes sense especially for tractor and large trucks given their relative size to medium and small trucks. This is supported by the negative effect of population to the probability of logistics facilities being developed in an area. This is because the large volume and surface area of tractor and large trucks make it more challenging for truck drivers to maneuver in highly populated areas such as in residential areas. Moreover, we could estimate the average Total Logistics Floor area from the estimated model as well conduct policy sensitivity analysis for infrastructure improvements such as shortening the distance to expressway interchanges and increasing the share of industrial land-uses.

Regarding trip generation of tractor and large trucks, the results of the Truck Trip Generation model showed that total floor area of logistics facilities positively affect trip generation of tractors and large trucks in an area. Furthermore, the share for all land-use classifications are highly statistically significant and positively affect trip generation of tractors and large trucks especially on the share quasi-industrial, industrial, and restricted-industrial land-use. The results of the Truck Trip Generation model also showed that the distance of an area to the closest expressway interchange and to the Port of Tokyo negatively affect tractor and large trucks trip generation; meaning, as areas become further away from an expressway interchange and from Port of Tokyo, the fewer the trips for tractor and large trucks generated. These are important findings especially for city planners and road managers because the development of logistics facilities, land-use allocation as well as the development of expressways and expressways interchanges will have an impact on the trip generation of tractors and large Trucks.

We could link the LFFA model and Truck Trip Generation model by evaluating truck trips generated using the estimation results of the Truck Trip Generation model and average logistics facility floor area from the LFFA model as shown in Table 7 in Section 3. This paper contributes to the research of logistics and freight movements by developing a modeling framework that could be used to analyze the effects of land-use policy changes and infrastructure improvements to logistics facility size and truck trip generation by estimating separate models for total floor area of facilities and for truck trip generation and linking them together to forecast travel demand of trucks. Lastly, the modeling framework that we have demonstrated can be replicated in other cities. This is because we have simply applied statistical methods using data from an urban freight survey in modeling logistics land-use location choice and floor area in conjunction with truck trip generation. However, we stress the importance of a well-conducted urban freight survey such as the TMAUFS that include surveys to logistics firms and truck probe data as well as a cohesive database of land-use allocation. These are the keys to implementing statistical models that are not only for theoretical applications but also for practical purposes such as sensitivity analysis of certain policy changes.

Being able to consider both the allocation, size, truck trip generation, and the land-use, in freight planning, we could determine the impacts of certain infrastructure improvements and policy changes especially to logistics facility size and truck volumes. This implies that city planners and public authorities could influence the spatial allocation and travel demand of trucks by appropriately considering the land-use distribution shares and the location of future road infrastructure improvements such as expressway interchanges.

Finally, we recommend further analysis which considers the 2003 Tokyo Metropolitan Government policy of restriction code for diesel trucks inside the metropolitan area is recommended to assess the before and after effect on logistics facilities allocation, size and truck generation.

ACKNOWLEDGEMENTS

We would like to thank the Transport Planning Commission of the Tokyo Metropolitan Region for the provision of 2003 and 2013 TMAUFS data for this research. Also, we would like to thank Dr. Daisuke Fukuda (Tokyo Institute of Technology) for the discussion of the chief model ideas in this paper. Finally, we would like to thank the two reviewers that impart to us invaluable feedback and comments to improve our paper.
REFERENCES