MODELLING THE SPATIAL DISTRIBUTION OF COMMERCIAL VEHICLE OWNERSHIP: AN APPLICATION TO WINDSOR, ONTARIO

Aya Hagag and Hanna Maoh
Department of Civil and Environmental Engineering
University of Windsor

Introduction

Commercial vehicles play an important role in the movement of goods in urban areas. Although they account 10 – 15 percent of the total daily urban traffic (Hunt and Stefan, 2007), they have a higher impact on the performance of the transportation system. More specifically commercial vehicles on average have significantly larger gross vehicle weight (GVW), and as such contribute to the deterioration of pavement and air quality.

The spatial distribution of commercial vehicles, as in the case of private vehicles, is a key determinant of the level of traffic volume that can be observed on the road network. To date, analysis to analyze the spatial location of commercial vehicles is limited. Most of the existing efforts have been focused on private vehicles instead. Identifying the factors that explain the prevalence of specific types of urban commercial vehicles at a given location in the city could help transportation modelers to devise effective travel demand models.

In this paper the determinants which lead to the prevalence of a specific type of commercial vehicles (as derived from the GVW classes) in a given zone are investigated. A multinomial logit (MNL) model is formulated to test the influence of various zonal and vehicle characteristics on the prevalence of commercial vehicles in traffic analysis zones (TAZs). When considering a set of alternative TAZs, the MNL model predicts the probability of finding commercial vehicle type t in TAZ i. To our knowledge, no previous study has developed such model.

The remainder of this paper is organized as follows. The next section highlights the data used in this analysis. It also discusses the specification of the MNL model. This will be followed by a section to present and discuss the statistical results. Finally, a conclusion to our study will be provided in the last section.

Data and Methods

The data used for this analysis was acquired from two main sources: (1) R. L. Polk and Co for the Windsor-Essex region, and (2) 2011 Canadian Census. The former is a dataset composed of all registered commercial vehicles at the census tract level in the region, where commercial vehicles are classified into eight different classes according to their Gross Vehicle Weight (GVW) and geo-referenced to the census tract level. The latter is a data source that contains demographic information such as the population and the employment number by industry type for each census tract.

Discrete choice modeling techniques namely, the Multinomial logit (MNL) model, were used to capture the effect of different zonal and vehicle characteristics on the prevalence of commercial vehicles in different zones. Even though the Polk dataset includes the entire population of 13,983 registered commercial vehicles in Windsor for the year 2013, only a ten percent random sample was used for the analysis to reduce the computation time required to perform the model estimation. This resulted in a total of 1400 commercial vehicles that were checked and validated to be a good representative sample of the entire population. The econometric software NLOGIT 5 was used to estimate the MNL model.

Given the population of commercial vehicles, the MNL model predicts the probability of finding a given commercial vehicle v in one of the 73 zones comprising the study area. Although the choice probability could be modeled across all 73 zones, we opted for generating a smaller choice set of ten alternative zones. The later was formed by randomly selecting nine zones and adding them to the actual zone where vehicle type t is located. Each zone i is associated with a utility function, U_i^t, that can be expressed as follows:

$$U_i^t = V_i^t + \varepsilon_i^t$$

where V_i^t is a linear-in-parameter deterministic function characterizing the nature of alternative zone i and the attributes of commercial vehicle t. On the other hand, ε_i^t is a random term that is assumed to be independently and identically distributed (iid) according to a Gumbel p.d.f. The choice probability can be represented according to the following MNL model:

$$\Pr(i) = \frac{\exp(V_i^t)}{\sum_{j=1}^{10} V_j^t}$$

The starting point for specifying the deterministic utility function is to consider jobs in the various zones. These are the locations where commercial vehicles are housed by their respective firms. Our rationale is based on the premises that locations where jobs are found give rise to the existence of commercial vehicles since firms housing these jobs will need to own vehicles for their business transportation activities (i.e. delivering goods and/or providing services). Given the diversity in the clustering pattern of industries over space (Maoh and Kanaroglou 2007), we expect to observe variability in the spatial distribution of commercial vehicles in the various zones across the city.

Intuitively, not all industries will be dependent on the same types of vehicles for their business transportation activities. For instance, the basic industry is more likely to own heavy duty trucks while firms from the services sector are more prone to own and use small cars and light commercial trucks. The inclusion of a list of variables which represent the number of jobs (by zone) from a particular industry would allow us to capture these differences especially in the presence of interaction terms characterizing the class of commercial vehicles. For the latter, the GVW classification is utilized. More specifically, the eight GVW classes are grouped based on the Federal Highway Administration (FHWA) classifications to: (1) Cars (GVW unknown); (2) Light Duty Trucks (GVW 1-2); (3) Medium Duty Trucks (GVW 3-6); and (4) Heavy Duty (GVW 7-8). Figure 1 presents the breakdown of the 13,983 commercial vehicles that were registered in Windsor in the year 2013, while Figure 2a – 2d highlights the spatial distribution of these vehicles.

Zonal jobs are classified into 5 major industrial groups following the categorization used in Hunt and Stefan (2007). That is, basic industry, wholesale, retail, transportation and services, as shown in Figures 3 and 4. The natural logarithm is applied to the generated zonal jobs to capture the non-linear effect between the size of jobs in a given zone and the probability of finding a particular type of commercial vehicle in that zone. This transformation produced more stable results in the model.

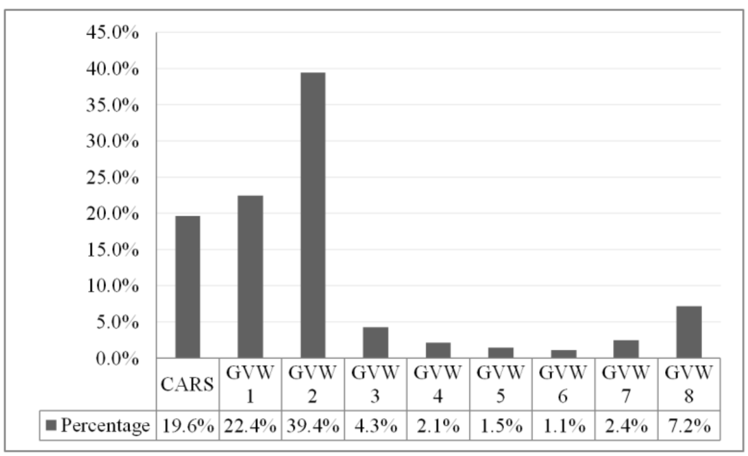

Figure 1: Distribution of commercial vehicle classes

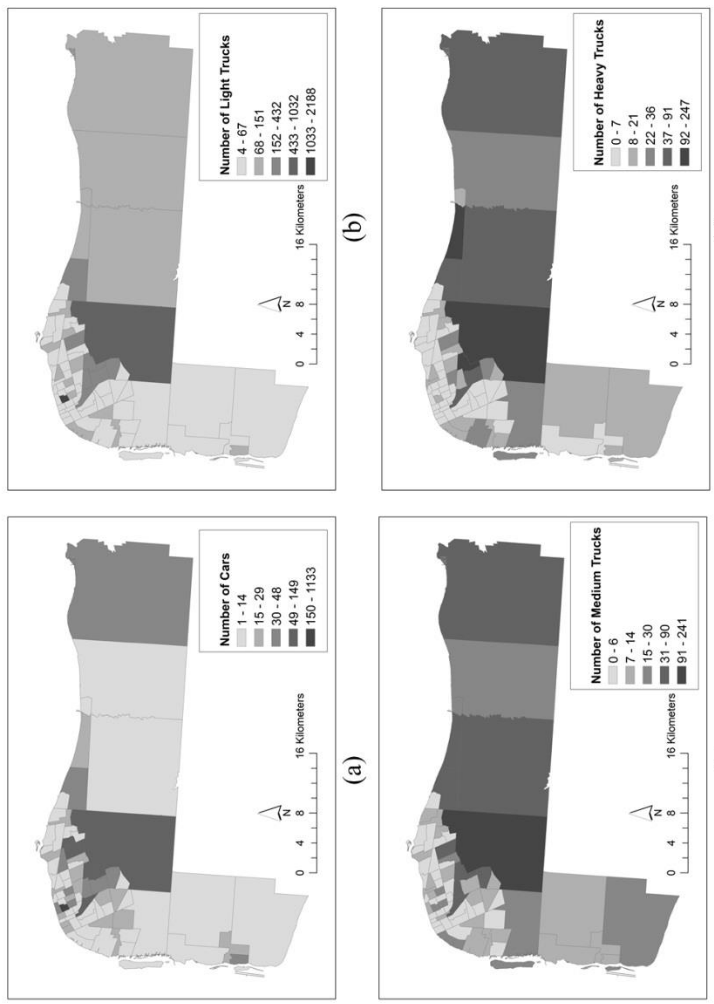

Figure 2: Spatial distribution of (a) commercial cars; (b) light commercial trucks (c) medium commercial trucks; and (d) heavy commercial trucks by place of registration in Windsor in 2013

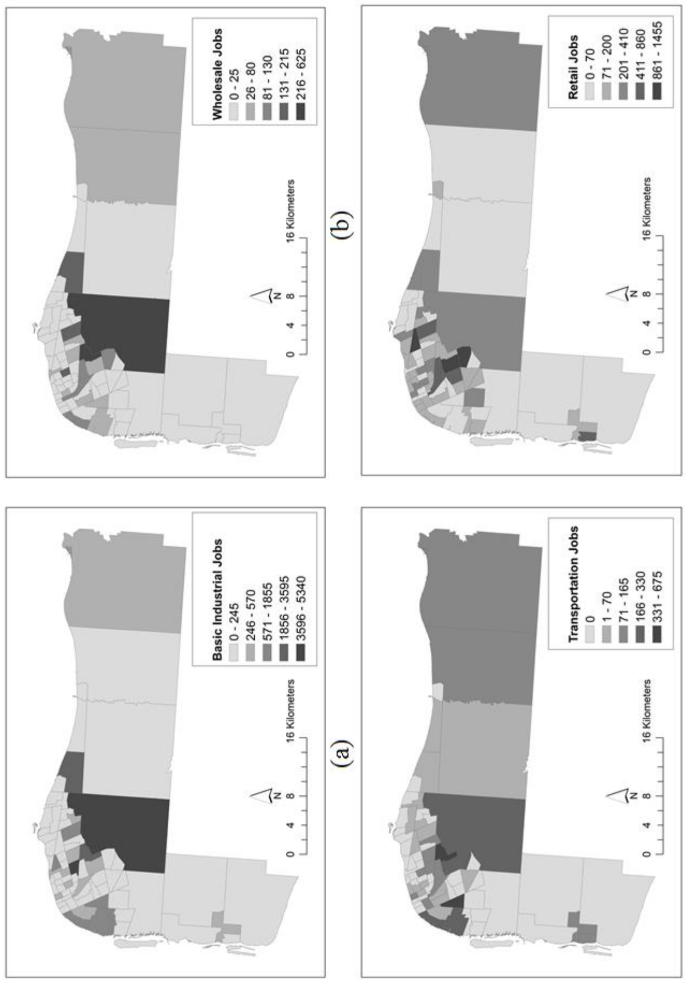

Figure 3: Spatial distribution of (a) basic sector jobs; (b) wholesale trade jobs; (c) transportation jobs; and (d) retail trade jobs, in Windsor in 2011.

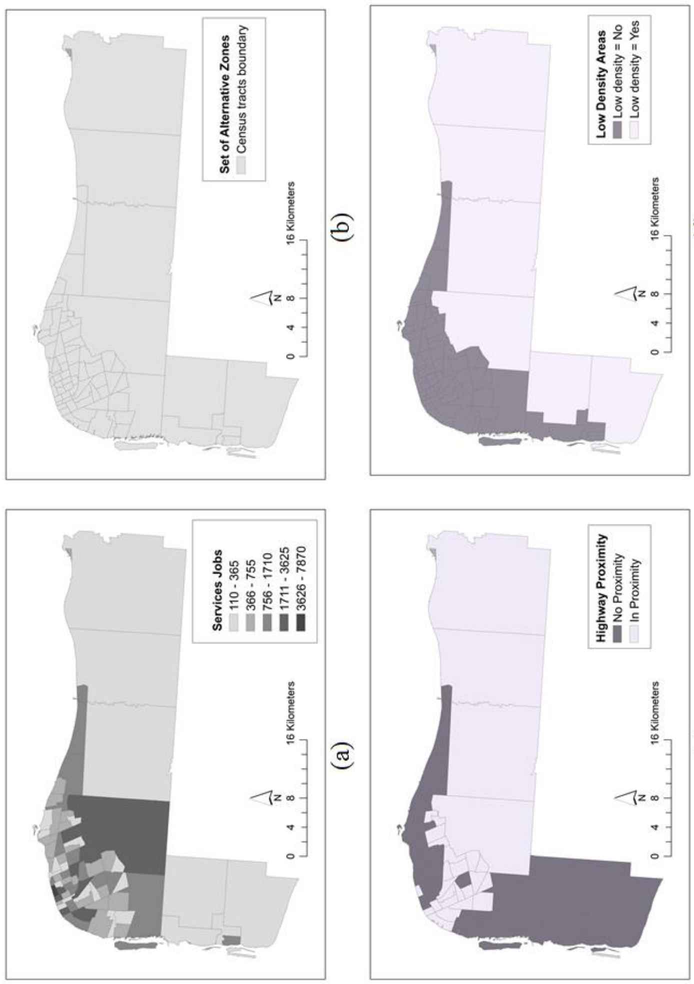

Figure 4: (a) Spatial Distribution of Services Sector jobs in Windsor in 2011; (b) Set of alternative zones used to form choice sets; (c) zones in proximity to highways; and (d) low density suburban zones.

Table 1 lists the variables used in the specification of the utility equation with their description. Our first a priori expectation is that zones with basic industry firms will be more prone to housing heavy and medium vehicles, other things being equal. On the other hand, these zones would be less likely to house cars and light trucks since their use to deliver goods is not common. When it comes to the wholesale industry, zones with firms pertaining to this industry are more likely to house various types of commercial vehicles. However, it is expected that firms from this industry will be more affiliated with medium and heavy trucks. Likewise, we expect to see a strong affiliation between zones housing the transportation industry and the prevalence of heavy trucks. At the same time, this industry is expected to show preference to other types of commercial vehicles albeit we do not expect the same influence as in the case of heavy trucks. Zones housing the retail industry are more likely to be affiliated with the prevalence of cars and light trucks. In contrast, we do not expect retail firms to own medium and heavy trucks in general. The same could be said about firms from the services sector.

Table 1: Description of Explanatory Variables

Variable	Description
$\ln(IND_i)$	The natural log of the total number of Basic Industrial jobs in each zone *i*
$\ln(WHL_i)$	The natural log of the total number of Wholesale jobs in each zone *i*
$\ln(TRA_i)$	The natural log of the total number of Transportation jobs in each zone *i*
$\ln(RET_i)$	The natural log of the total number Retail jobs for each zone *i*
$\ln(SER_i)$	The natural log of the total number of Service jobs in each zone *i*
$\ln(Area_i)$	The natural log of the area in kilometers squared for each zone *i*
$HWYPRO_i$	1 if the zone *i* is in close proximity with highways, 0 otherwise
$LDENS_i$	1 if the zone *i* has less than 100 jobs per km^2 and less than 100 residents per km^2, 0 otherwise

$\ln(IND_i) \times$ M	Interaction term between the natural log of basic industrial jobs in each zone *i* and Medium Duty Trucks
$\ln(IND_i) \times$ H	Interaction term between the natural log of industrial jobs in each zone *i* and Heavy Duty Trucks
$\ln(WHL_i) \times$ M	Interaction term between the natural log of Wholesale jobs in each zone *i* and Medium Duty Trucks
$\ln(TRA_i) \times$ L	Interaction term between the natural log of Transportation jobs in each zone *i* and Light Trucks
$\ln(TRA_i) \times$ H	Interaction term between the natural log of Transportation jobs in each zone *i* and Heavy Duty Trucks
$\ln(RET_i) \times$ C	Interaction term between the natural log of Retail jobs in each zone *i* and Cars
$\ln(RET_i) \times$ L	Interaction term between the natural log of Retail jobs in each zone *i* and Light Trucks
$\ln(SER_i) \times$ C	Interaction term between the natural log of Service jobs in each zone *i* and Cars
$\ln(SER_i) \times$ L	Interaction term between the natural log of Service jobs in each zone *i* and Light Trucks
$HWYPRO_i \times$ L	Interaction terms between Light Trucks and zones in close proximity with highways
$HWYPRO_i \times$ M	Interaction terms between Medium Duty Trucks and zones in close proximity with highways
$HWYPRO_i \times$ H	Interaction terms between Heavy Duty Trucks and zones in close proximity with highways

C = Cars; L = Light Trucks; M = Medium Trucks and H = Heavy Trucks

Next, we hypothesize that zones with larger land areas will house more commercial vehicles, other things being equal. The rationale here is that smaller zones are typically associated with high population densities and not as many business establishments especially in a sprawled city like Windsor. Similarly, low density zones (i.e. mainly suburban zones, as shown in Figure 4, in which density is less than hundred jobs per km^2 and less than hundred residents per km^2) are expected to be more specialized in terms of their firm population (e.g. zones housing warehouses) and as such are more likely to house commercial vehicles, other things being equal. Furthermore, zones in close proximity to highways and interchanges are also expected to attract more commercial vehicles since locations in proximity to transportation infrastructure are normally considered prime sites for business establishments to locate. Figure 4c highlights the zones in proximity to highways and interchanges.

Results

The estimation results, as shown in Table 2, suggest a well behaved model with an acceptable McFadden pseudo $\boldsymbol{\rho}^2$ of 0.237. Based on the estimated coefficients, zones with higher number of basic industry jobs tend to be strongly affiliated with medium and heavy duty trucks, as discerned from the parameters of the two interaction terms *Inter1* and *Inter2*. As expected, those zones are also less prone to give rise to the prevalence of cars and light trucks. These results are expected given the nature of these industries (e.g. manufacturing and mining) which relies heavily on larger trucks for their business operation.

In a similar vein, the results from the interaction terms *Inter3* and *Inter4* suggest that zones housing wholesale firms are strongly affiliated with medium and heavy duty trucks. As we anticipated, the wholesale industry is likely to own cars and light trucks (as discerned by the positive sign of the $\ln(WHL_i)$ parameter when compared to the parameters of the two interaction terms *Inter3* and *Inter4*) but their dependency on larger trucks is evident in the model. In the case of the transportation industry, zones housing firms of this type tend to rely on all types of vehicles although their dependence on light commercial vehicles is more pronounced, as can be discerned from the interaction term *Inter5* when compared to the parameter of the $\ln(TRA_i)$.

Table 2: Estimation Results of MNL Model

Parameter	Beta	t-stats
$\ln(IND_i)$	-0.109	-4.96
$\ln(WHL_i)$	0.062	2.82
$\ln(TRA_i)$	0.108	2.89
$\ln(RET_i)$	-0.152	-3.29
$\ln(SER_i)$	0.436	4.62
$\ln(Area_i)$	0.096	2.00
$HWYPRO_i$	-1.104	-6.95
$LDENS_i$	0.950	5.16
$\ln(IND_i) \times \mathrm{M}$	0.244	2.94
$\ln(IND_i) \times \mathrm{H}$	0.277	4.40
$\ln(WHL_i) \times \mathrm{M}$	0.153	1.98
$\ln(TRA_i) \times \mathrm{L}$	0.094	2.17
$\ln(TRA_i) \times \mathrm{H}$	0.120	1.58
$\ln(RET_i) \times \mathrm{C}$	0.388	5.11
$\ln(RET_i) \times \mathrm{L}$	0.288	5.32
$\ln(SER_i) \times \mathrm{C}$	0.697	5.55
$\ln(SER_i) \times \mathrm{L}$	0.348	3.36
$HWYPRO_i \times \mathrm{L}$	0.402	2.33
$HWYPRO_i \times \mathrm{M}$	1.269	4.27
$HWYPRO_i \times \mathrm{H}$	1.157	4.17
$\boldsymbol{L(0)} = -3223.619$, $\boldsymbol{L(\beta)} = -2460.748$		
$\rho^2 = 0.237$		

Zones housing retail trade firms tend to be highly affiliated with cars and light trucks although the prevalence of cars is more pronounced. By comparison, these zones repels medium and heavy duty trucks as can be deduced from the sign of the $\ln(RET_i)$ parameter which is negative and significant relative to the positive and significant parameters of the two interaction terms ($\ln(RET_i) \times$ C) and ($\ln(RET_i) \times$ L). Zones housing services sectors firms are more prone to give rise to the presence of cars and light trucks, as in the case of the retail trade sector. However, the presence of cars is more pronounced for these services sectors. Also, unlike the retail trade case, services firms tend to rely on medium and heavy duty trucks for their business transportation activities as can be discerned from the positive and significant parameter of the $\ln(SER_i)$ variable.

As expected, larger zones tend to have a positive influence on the presence of commercial vehicles. Likewise, zones with low land use densities tend to have a positive influence on the presence of commercial vehicles, other things being equal. It is found that zones in proximity to highways tend to be more associated with the presence of heavier vehicles, mainly medium duty trucks. On the other hand, those zones are less prone to having cars. These results are sensible especially that locations in proximity to highways and interchanges (i.e. highway ramps) are attractive to manufacturing and heavy industry firms that rely heavily on accessibility (Maoh and Kanaroglou, 2009). These firms are more likely to own medium and heavy trucks for their goods movement activities.

Conclusions and Future Research

In this paper we investigated the determinants which lead to the prevalence of a specific type of commercial vehicle in a given zone. The total number of jobs per zone, zone’s land area and whether the zone is in close proximity to a highway were found to have positive effect on the prevalence of commercial vehicles in specific zones. Moreover, we were able to explain the variability in the spatial distribution of the key types of commercial vehicles based on the presence and dominance of certain industries in these zones. The analysis conducted here is unique in that no previous study has attempted to explain the determinants that give rise to the presence of the commercial vehicles owned and registered by firms at small area geography like a Traffic Analysis Zone (TAZ).

From a transport policy perspective, modeling and understanding car ownership is essential to both urban transport and land-use planning, since vehicles influence trip generation, mode choices and also location choices (Whelan, 2007; Potoglou and Kanaroglou 2008). Moreover, vehicles are responsible for the number of kilometer travelled and energy consumed in an urban area. To date, all of the existing efforts have been solely focused on private car ownership modeling. Therefore, the research presented in this paper provides the basis for more work on the topic to develop predictive transport demand models and subsequently evaluate the implication of enforcing or changing certain policies (Jong et al., 2004).

It should be noted that one of the main reasons for the underdevelopment of commercial vehicle ownership models in the literature is due to the lack of detailed commercial vehicle travel data. Obtaining such information via surveys is mainly dependent on the cooperation of private firms, who most of the time vacillate to share information related to their business freight and/or transportation activities. As a results most of the existing efforts to collect detailed commercial travel data resulted in a low response rate and very few observations. In short, collection of a meaningful sample for statistical analysis is costly and very timely exercise. Hence, modeling efforts that make use of commercial vehicle registration data like the one used in this paper can help overcome some of these limitations. Future research emanating from the current work will focus improving the quality of the utilized Polk vehicle registration data by introducing fuel efficiency information from other resources such as fuel efficiency data that is published by the US Department of Energy. It is hoped that the enrichment of the Polk information could improve the capabilities of the model developed in this paper.

Acknowledgement

The authors are thankful to the Natural Science and Engineering Research Council of Canada (NSERC) for supporting this research through a Discovery Research Grant.

Bibliography

Hunt, J., and K. Stefan (2007). Tour-Based Microsimulation of Urban Commercial Movements. *Transportation Research Part B:* 41(9): 981-1013.

Maoh, H. and P. Kanaroglou (2007). Geographic Clustering of Firms and Urban Form: A Multivariate Analysis. *Journal of Geographical Systems*, 9: 29-52.

Maoh, H. and P. Kanaroglou (2009). Intrametroplitan Location of Business Establishments: Microanalytical Model for Hamilton, Ontario, Canada. *Transportation Research Record: Journal of Transportation Research Board*, 2133: 33-45.

Jong, G., Fox, J., Daly, A., Pieters, M. and Smit, R. (2004). Comparison of car ownership models. *Transport Reviews*, 24(4), pp.379-408.

Potoglou, D. and Kanaroglou, P. (2008). Modelling car ownership in urban areas: a case study of Hamilton, Canada. *Journal of Transport Geography*, 16(1), pp.42-54.

Whelan, G. (2007). Modelling car ownership in Great Britain. *Transportation Research Part A: Policy and Practice*, 41(3), pp.205-219.