

# **Factors Affecting the Capacity Utilization of Road Freight Transport in Turkey**

**Murat ÖZEN<sup>1</sup>**

**Muhammad FAYYAZ<sup>2</sup>**

**Hediye TÜYDEŞ YAMAN<sup>3</sup>**

## **ABSTRACT**

Road freight accounts for 89% of the inland freight transportation in Turkey. This study examined the factors affecting capacity utilization in Turkey with a specific focus on load factors and empty running. These two decisions, the load factor and the choice of running empty, are modeled jointly in a Heckman model. The model results showed that rigid and articulated trucks represent relatively different transport markets. Their age and the distance profile of the shipments affect the load factor and the probability of running empty. These effects are further examined for different commodity types and by testing different econometric specifications.

**Keywords:** Capacity utilization, load factor, empty running, heckman model, commodity types.

## **1. INTRODUCTION**

Road dominates the inland freight transportation in Turkey with 89% market share [1]. The situation is quite similar in many other countries [2]. For instance, in the European Union-28 region, 75% of the inland freight is transported by trucks [3]. Although there have been several policies to reduce share of road freight, the dominance of trucking is expected to continue due to its flexibility and door-to-door delivery [4]. This problem is more critical in developing countries, such as Turkey, in which other alternatives are not competitive due to different reasons (i.e. shorter hauling distances, lack of intermodal freight infrastructure and legislation, etc.). Therefore, it becomes more important to evaluate efficiency of road freight

---

Note:

- This paper has been received on April 27, 2018 and accepted for publication by the Editorial Board on February 04, 2019.
- Discussions on this paper will be accepted by May 31, 2020.

• <https://dx.doi.org/10.18400/tekderg.413807>

1 Mersin University, Department of Civil Engineering, Mersin, Turkey - [ozen.murat@mersin.edu.tr](mailto:ozen.murat@mersin.edu.tr) - <https://orcid.org/0000-0002-1745-7483>

2 Institute of Transport and Logistics Studies, The University of Sydney, Sydney, Australia - [muhammad.fayyaz@sydney.edu.au](mailto:muhammad.fayyaz@sydney.edu.au) - <https://orcid.org/0000-0003-1075-5688>

3 Middle East Technical University, Department of Civil Engineering, Ankara, Turkey - [huydes@metu.edu.tr](mailto:huydes@metu.edu.tr) - <https://orcid.org/0000-0003-2053-992X>

movements to reach more efficient and environmentally friendly freight transportation system.

Capacity utilization (load factor) is a key factor in assessing efficiency of freight transportation. Load factor can be defined as ratio of actual weight of goods to the maximum weight that can be transported on a fully loaded trip [7]. It is a weight based measure and it may underestimate actual utilization of vehicle in some sectors, such as automobile, food and parcels where utilization of vehicles are limited by volume rather than weight [8]. Abate and Kveiborg [5] stated that, the capacity utilization of the trucks depends on several factors including characteristics of a trip (length, geographical area that it takes place, etc.), type of truck (operator type, size, etc.) as well as type of goods carried. The decision on how much to load or whether the truck should run empty represents an economic trade-off for the truck operator. Empty running occurs due to the unavailability of a load at the time of the return trip. However, it may also occur, if carriers have to deliver to several locations during a trip chain. Sometimes it is unavoidable, since freight trips are generally performed from point of production centers to consumption centers [5-6]. Despite an increase in revenue, a return load may also represent additional costs in terms of detouring and waiting costs, which may outweigh the additional revenue [9]. As a result, the frequency of empty runs is not random across the regions. For example, for importing regions it is harder to find matching return trips.

It is essential to represent this decision problem in a joint framework in which the joint decision of running empty and choosing a specific load-factor was modeled as a sample selection problem [10-12]. The approach was applied to Danish heavy truck trips, for which it was clearly shown that there are correlations between the two stochastic processes. The objective of this study is to examine the factors affecting capacity utilization in road freight transportation for Turkey by investigating the joint decision problem of running empty (a binary choice) and having a certain load factor. Using disaggregate data collected by the Turkish General Directorate of Highways (TGDH) from 2007 to 2011, this study has applied a joint estimation methodology for empty running and the load factor decision.

The main contribution of the study is the use of generally accepted models within a developing country context such that dynamics of empty runs in Turkey are determined. In a broader perspective, the paper contributes to the understanding of truck capacity utilization trade-offs in developing countries where there is a high share of road transport, an old fleet of trucks, lower GDP, and a different degree of urbanization compared to industrialized countries. Furthermore, the findings of this study have relevance for other developing countries, where there are few alternatives to road transport due to lack of rail and shipping infrastructure. A limitation of the study is the slightly short term of data availability, which was still long enough to derive statistically significant patterns, which can be used to develop more efficient trucking sector policies towards a low carbon development structure for Turkey.

The next section of this paper describes the data used in this study; which includes data characteristics and limitations. The section thereafter describes the econometric framework which consists of factors affecting load factor and empty running, and model specification based on the Heckman model. The following section presents the overall results for the various models. The concluding section provides some policy implication and future research directions.

## 2. DATA

This study used the roadside axle surveys conducted on intercity highways. These surveys have been regularly conducted to determine characteristics of heavy vehicle and freight movements in Turkey evaluated by TGDH. These surveys included weighing of trucks, randomly stopped according to a predetermined sampling ratio, in addition to interviews collecting information about the truck type (rigid or articulated), age of truck, commodity type, empty weight, load carrying capacity and total weight, as well as origin and destination provinces and country of the trip. Using the information of the load carrying capacity, empty weight and total weight, it was possible to determine the load factor (capacity utilization) of trucks. Here, it should be noted that as these surveys were performed on inter-city roads, they were capable of capturing mostly inter-city truck movements, and did not provide insight on the intra-city movements; thus, they are excluded from the scope of this study.

While a more elaborate discussion of the road side axle surveys was provided by Ozen [13], reports evaluating truck freight transportation by TGDH covering the periods of 2007-2009 and 2010-2014 have been published [14,15]. However, aforementioned specific data requirements of the model used in this study enabled the use of data only from 2007 to 2011, which is one of the reasons between small differences in selected measures also determined in this study. While this caused slight differences in measures, another reason is the data cleanup process followed in this study to eliminate surveys with missing data.

*Table 1 - Descriptive statistics for model variables*

Variable	Variable description	Average
LF	Load factor for loaded trips (Load factor for all trips)	78% (58%)
L	1 if a truck is loaded and 0 otherwise	69.1%
Distance	Average distance of the trip (in km)	435
Size	Average gross vehicle weight (tons) (Empty weight + load carrying capacity)	29.0
MC	Maximum allowable limit (tons)	17.9
CW	Average commodity weight (tons)	10.4
Age	Average age of vehicle (years)	6.7
Rigid	Rigid trucks	68%
Articulated	Articulated trucks	32%
Net-importer*	1 for trips made towards a net importing region; 0 otherwise	11%
Voluminous**	1 for voluminous cargo; 0 otherwise	39%

Source: Turkish General Directorate of Highways (TGDH): Road Side Axle Surveys from 2007 to 2011.

\*Net Importer: regions in Turkey whose imports (in terms of the goods) is higher than exports.

\*\*Voluminous: Goods which need an extra amount of space (e.g., timber, grains etc.).

Descriptive statistics about the data from 53,327 intercity trucks used in this study are summarized in Table 1. While 69.1% of the surveyed trucks were loaded, load factor on

average was determined as 78% with an average commodity weight of 10.4 tons. While average trip distance was calculated as 435 km for the study data, it was determined by assigning the trip on the time-based shortest path between the stated origin and destination, as established by Ozen [13]. It should be noted here that this may be a partial reason causing trip length deviating slightly from the statistics published by the TGDH.

Leaving empty trucks aside constituting approximately one-third of the sample, further analysis by load type of the laden trips showed that while some load types were overrepresented others were not represented enough either due to their smaller market shares or sampling process during the axle load surveys which were performed in short sampling periods (such as 2-3 days at a study location). Thus, the current 20-commodity types by TGDH following to “Standard Goods Classification for Transport Statistics (NST)-2007” was further simplified into 6 categories as shown in Table 2, similar to the NST categories employed by TGDH prior to 2007 [16]. In this categorization Commodity Type 5 (including coke, refined petroleum, etc.) ranked first with a share of 18.3% of the sample, followed by Commodity Type 3 (food products, beverages and tobacco) with a share of 12.7 % (see Table 2). Commodity Type 1 (products of agriculture, hunting and forestry) with 10.8% share was followed by 7.9% share of Commodity Type 2. This simpler categorization enabled generation of dummy variables for load types to be used in the models. Note: Grouped goods or unidentifiable goods by NST-2007 (NST Group 18 to 20) were omitted as well as commodities that had really small shares which could not be aggregated under the selected categories in Table 2, such as, secondary raw materials, municipal wastes and other wastes (NST Group 14); mail, parcels (NST Group 15) and goods moved in the course of household and office removals (NST Group 17).

*Table 2 - Classification of commodity groups*

Commodity	Commodity Type	Frequency	Percent
0	Empty	16388	30.8
1	Products of agriculture, hunting, and forestry	5735	10.8
2	Coal and lignite; peat; crude petroleum Metal ores and other mining products	5328	10.0
3	Food products, beverages and tobacco	6763	12.7
4	Textiles and textile products Wood and products of wood and cork Furniture, other manufactured goods	4268	8.0
5	Coke, refined petroleum products Chemicals, chemical products Other non-metallic mineral products	9753	18.3
6	Basic metals; fabricated metal products Machinery and equipment n.e.c. Transport equipment Equip. and materials utilized in the transport of goods	4992	9.4

Source: Turkish General Directorate of Highways (TGDH): Road Side Axle Surveys from 2007 to 2011. Number of observations: all trips (53,327); loaded trips (36,838) and empty trips (16,489).

Figure 1 and 2 show the market share of loaded (not empty, L=1) and empty (L=0) trips as a function of distance and size. Approximately 71% of the articulated trucks were loaded when surveyed, and almost half of those were traveling more than the average hauling distance of 435 km. The empty running by the articulated trucks followed the same trend, although on a smaller scale. When the rigid trucks were analyzed, a significant portion of them (loaded or not) were observed on rather short hauls of less than 400 km. More importantly, the number of trucks running empty seemed to fade out quite quickly. It is also worth noting that, for trips longer than 1000 km, distance seems to matter more for rigid trucks than for articulated trucks.

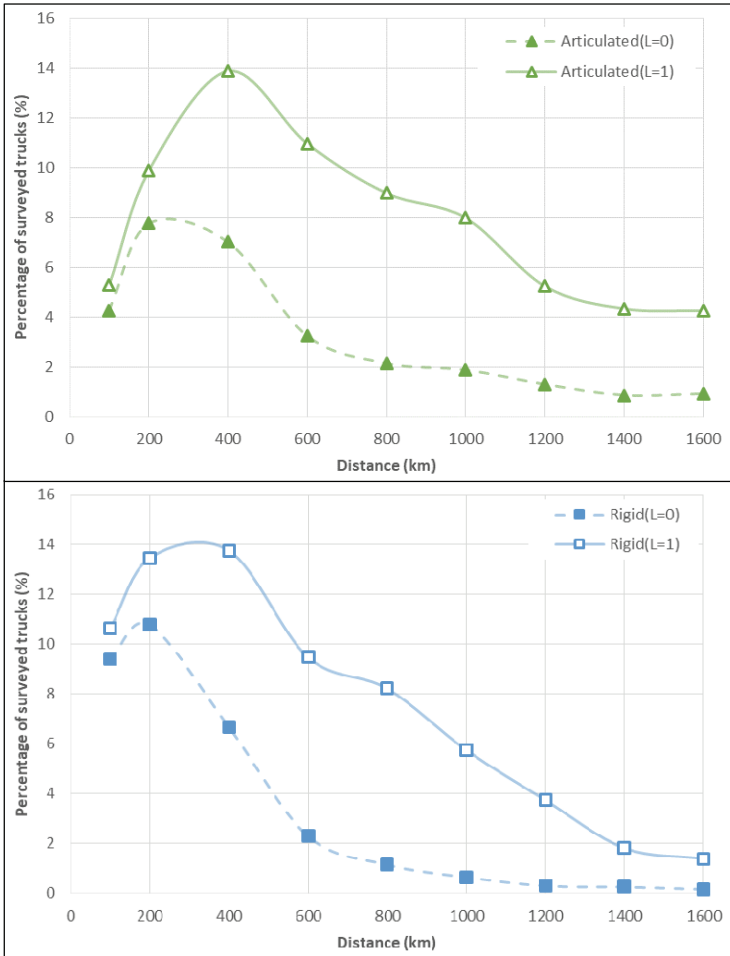


Figure 1 - Market share of loaded and empty trips as a function of distance

The distribution of road freight movements surveyed during 2007 to 2011 by truck size (see Figure 2) supports the fact that articulated and rigid trucks have different characteristics. In this study, truck size is defined in Gross Vehicle Weight (GVW), which is the sum of

maximum load carrying capacity and empty weight of the vehicle. As the size of the articulated trucks increases, the percentage of loaded trips within the size categories increases, hence the percentage of empty runs decrease. For rigid trucks, the percentage of loaded trips within each truck size category fluctuates without suggesting a clear trend as is the case for articulated trucks.

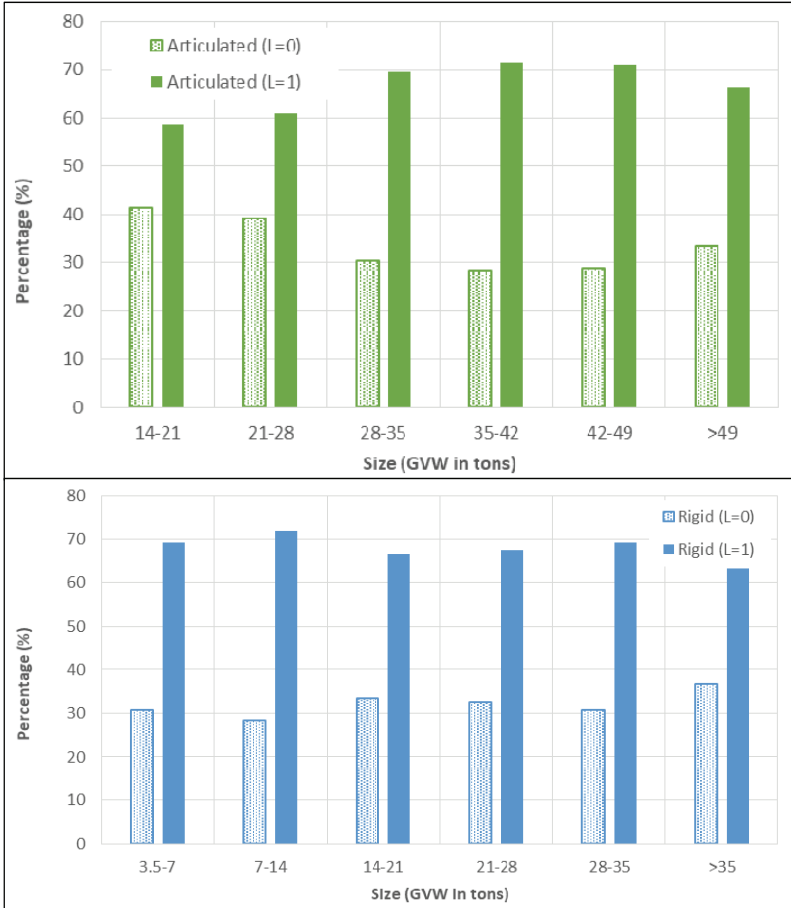


Figure 2 - Market share of loaded and empty trips as a function of truck size (Gross Vehicle Weight)

### 2.1. Limitations

One of the limitations in the data is that the surveys are conducted on state roads, which include mainly intercity freight transportation. As a result, the analysis is not representative for intra-city freight transport, which will behave differently. Secondly, information about trip chains, truck tours, and warehouses, as well as loading and unloading at transitional hubs, is not included in our data set. This would have made possible performing a longitudinal study. Lastly, unavailability of driver's salary related data. One major difference from

developed country as the cost of the truck, although with a comparatively older age, and fuel cost is not dramatically different but the driver salary is. This could explain issues like the readiness to wait for a balancing load and the labor that can be rationally spent on loading the truck carefully. However, it should be noted that the data set is more disaggregated than is usual for a developing country. The data is almost similar to the Danish data set used in Abate [10] and to data in other EU countries. But, as mentioned above, more recent data (data after 2011) did not provide required details preventing their use in the modeling step of this study. Nevertheless, for many of the developing countries, data are hard to obtain due to privacy reasons; most importantly, in the case of developing countries, such data are almost never available due to lack of surveys and resources. However, beyond roadside survey data, it is necessary to collect commodity flow data at a national scale as in the USA.

### 3. METHODOLOGY

#### 3.1. Factors Affecting the Load Factor

The load factor can be defined as the ratio of the commodity weight to the maximum carrying capacity of the truck, expressed in percentage. To evaluate capacity utilization with respect to the load factor, weight and volume (density) of the respective commodities should ideally be taken into account. However, such density data are often not available (as in this study). As a result, this study applied a weight-based load factor percentage ( $LF_i$ ). This is defined as the ratio between commodity weight ( $CW_i$ ) and maximum carrying capacity ( $MC_i$ ).

$$LF_i(\%) = 100 \times \frac{CW_i}{MC_i} \quad (1)$$

Factors affecting  $LF_i$  include distance, commodity groups (to account for heterogeneity of potential high density cargo versus low density cargo), and truck size.

#### 3.2. Factors Affecting Market Access

The market access decision refers to the probability that the truck is loaded rather than empty. Generally, as the distance increases, the percentage of empty runs is expected to go down. It is not clear why there is a linear relationship between the distance and the probability that the truck is loaded. It may well be that there is a marginal decreasing effect of distance. To account for this, a Box-Cox model is also worked out. In fact, it is revealed that the Turkish transport market essentially constitutes two quite different transport markets, one for articulated trucks and another for rigid trucks. For articulated trucks, which are generally larger, the probability that the truck is loaded increases with size. The reverse is true for rigid trucks because rigid trucks are relatively small. Market access decisions may also depend on the freight movement balance, i.e., whether the region is a net-importing or net-exporting region [10]. The probability that the truck is loaded is usually higher for trips towards a net-importing region because it is not easy to find matching return trips. Conversely, because of a higher probability of matching return load from net-exporting regions empty trucks are more likely to be observed towards such regions [9]. Net exporting and net importing zones have been classified from overall freight movement in Turkey and this paper has introduced a dummy variable indicating whether the trip was made to a net-importing or net-exporting

region. Furthermore, truck traffic to and from the very large provinces, i.e., İstanbul, Ankara, and İzmir, will tend to have different characteristics, which in turn may affect the probability of loading. To capture this effect, another dummy variable was introduced for traffic going to and from these large regions. The age of the truck is usually assumed to be negatively related to the market access decision. It should be noted that, for Turkey, the truck age profile is quite different from those in the EU and other developed countries. The average age of trucks is higher and there is a relatively large segment of trucks older than 12-14 years, most of which are rigid. Hence, in some cases, the market access decision might not be altered by the age of the truck, particularly for rigid trucks.

### 3.3. Model Specification

Truck utilization can be modeled as a sample selection problem because the dependent variable (the load factor) is observed only when the truck is loaded. This was proposed by Abate [10], applying the standard Heckman model [12] to the decision to drive with the truck loaded or empty (henceforth referred to as the market access decision) and the load-factor percentage conditional on the truck carrying a load. Clearly, companies and freight carriers prefer to minimize empty running as much as possible, while also trying to maximize the load factor for loaded trucks to maximize profits.

The decision to load the truck is essentially connected to the market access of the carrier and may relate to characteristics of the carrier, the trip and the type of truck. The load-factor percentage will largely depend on the same variables; however, for identification purposes, it is required that some variables are used as exclusion restrictions in the market access decision equation but not in the load factor. If we let  $LF_i$  represent the load-factor percentage for observation  $i$  and  $L_i$  the corresponding binary choice of loading the truck, the model can be written as:

$$L_i^* = f(X_{1i}, \theta) + v_i \tag{2}$$

$$L_i = \begin{cases} 1, & \text{if } L_i^* > 0 \\ 0, & \text{if } L_i^* \leq 0 \end{cases} \tag{3}$$

$$LF_i^* = g(X_{2i}, \beta) + \varepsilon_i, \text{ if } L_i = 1 \tag{4}$$

$$LF_i = \begin{cases} LF_i^*, & \text{if } LF_i^* > 0 \\ 0, & \text{if } LF_i^* \leq 0 \end{cases} \tag{5}$$

$X_{1i}$  and  $X_{2i}$ , respectively, represent vectors of explanatory variables for the binary choice of  $L_i$  and the load-factor percentage equation ( $LF_i$ );  $\theta$  and  $\beta$  represent the structural parameters. A loaded truck is represented by  $L_i=1$ .  $X_{2i}$  contains all the variables in  $X_{1i}$  and additional variables for identification (exclusion restrictions). The residuals  $v_i$  and  $\varepsilon_i$  are jointly normal with zero mean. The assumptions for estimation are as follows [17]:

- 1)  $(X_{2i}, L_i)$  are always observed, but  $LF_i$  is only observed when  $L_i=1$  (sample selection);
- 2)  $(v_i, \varepsilon_i)$  is independent of  $X_{2i}$  with zero mean;



- 3)  $v_i \sim \text{normal}(0,1)$  (distributional assumption); and
- 4)  $E(v_i|\epsilon_i) = \gamma_2\epsilon_i$  (residuals are jointly normal and may be correlated).

In addition to these assumptions, it is necessary to remember that the  $LF_i$  distribution is essentially truncated at 0 (cannot have a negative load factor). Hence, a truncated version of the Heckman model is considered (see Eq. 3-4). It should be said, however, that considering the truncated version of the  $LF_i$  distribution makes relatively little difference in this example, and the problem of considering an un-truncated version seems to be relatively small [10].

The simplest way of estimating the model would be to assume that the load-factor percentage could be modeled by considering only the loaded trips. However, as documented by Heckman [12] and others, this will lead to a sample selection problem, which in turn leads to biased estimates. For instance, it is possible that trucks opt to run empty instead of loaded because the cost of running loaded exceeds the revenue [9]. A joint estimation is needed to model the load factor to understand its determinants at a population level. Therefore, a joint estimation model developed by Heckman implemented in this study, which has been widely used in many areas of science including transportation [18-20]. A more elaborate discussion of Heckman model was provided by de Jong [19]. As a result, this paper estimates the parameters of Eq. 1-4 by full information maximum likelihood (FIML), to account for joint correlation between the two endogenous components, using the QLIM procedure in the SAS Software. The complete likelihood function is given by:

$$\begin{aligned}
 LL(\beta, \theta\lambda) = & \sum_{i|L=0} \ln(1 - \Phi(f(X_{1i'}, \theta))) \\
 & + \sum_{i|L=1} \left[ \ln\left(\phi\left(\frac{LF_i - g(X_{2i'}, \beta)}{\sigma}\right)\right) - \ln\sigma \right. \\
 & + \ln\left[\Phi\left(\frac{f(X_{1i'}, \theta) + \rho\left(\frac{LF_i - g(X_{2i'}, \beta)}{\sigma}\right)}{\sqrt{1 - \rho^2}}\right)\right] \\
 & \left. + \ln\left(\Phi\left(\frac{f(X_{1i'}, \theta)}{\sigma}\right)\right)\right]
 \end{aligned} \tag{6}$$

According to Eq. 6, the first part of the function represents the censored part, where  $L_i = 0$ . The un-censored part consists of the equation resulting from the  $LF_i$  model (adjusted for truncation) and the choice of  $L_i$ . The model shown in Eq. 6 with linear-in-parameter specifications for  $f$  and  $g$  is considered a base model and is referred to as Model 0. The base model do not differentiate between types of trucks but estimate the Heckman model [12] for the whole sample, which reflects the model structure applied in Abate [10]. However, different truck types could behave differently in regard to the load-factor percentage and the share of empty runs, therefore, two alternative versions of Model 1 are estimated, one for rigid trucks and one for articulated trucks, in order to measure differences between the types

of trucks. Finally, these two models are formulated in more comprehensive versions, in Model 2, where main modifications are introduced.

Firstly, Model 2 allows for a richer parameterization of the distance and size variables in the LF model in order to account for differences in how commodity groups are described. This turns out to be relatively important as there are big differences in how size and distance affect the load-factors percentage for different commodities. Secondly, for the  $LF_i$  model, Model 2 account for unobserved heteroscedasticity in the error-term  $\varepsilon_i$  by representing the variance as a function of the distance. Hence, the variance in the  $LF_i$  model is now given by;

$$E(\varepsilon_i^2) = \sigma_i^2 = \sigma_i^2(1 + \gamma \ln(Dist_i)) \tag{7}$$

Generally, it is important to account for heteroscedasticity because the standard Heckman model can produce unreliable results if not accounted for, as shown in Arabmazar and Schmidt [21]. The robustness of the estimators for the sample selection models has been taken into account by [22-23]. Finally, Model 2 relax the linear-in-parameter assumption for the  $L_i$  model by estimating a Box-Cox model for the distance variable. Hence,

$$BC(Dist_i) = \begin{cases} \frac{Dist_i^\lambda - 1}{\lambda} & \text{if } \lambda \neq 0 \\ \ln(dist_i) & \text{if } \lambda = 0 \end{cases} \tag{8}$$

The linear-in-parameter specification, which is a special case of the Box-Cox parameter [24-26] when  $\lambda = 1$ , is clearly rejected.

#### 4. RESULTS

Table 3 and 4 presents the main results for the two different model segments for the rigid and articulated truck. Note that *t*-values at vehicle levels are included in parenthesis. As mentioned above, Model 0 in Table 3 and 4 is the simple base model without segmentation, whereas Model 1 includes one model for rigid and one for articulated trucks. Hence, the assumption of identical parameters for different truck types across the sample is relaxed. Model 2 includes the three additional econometric improvements. For the market access equation, three exclusion restrictions represented by age, net-import (freight movement balance), and dummies for large provinces are included (see Table 4). These variables have significant effects on the market access model, with empty running negatively related to age and positively related to the “net-import” variable and the large provinces dummies. On the other hand, in Model 2, the effect of age is positively significant but almost equal to zero. There are two reasons. First, there are more rigid trucks than articulated trucks in Turkey. Secondly, in this data-set, trucks older than 15 years are 10 percent of the whole rigid truck fleet, while trucks older than 10 years correspond to 28 percent, which is a substantial proportion and can’t be ignored. Although the new road freight transport regulations in 2003 favor moving freight by articulated trucks, it will take a long time to replace older rigid trucks with newer articulated trucks. Besides, the overall age distribution of heavy goods vehicles in Turkey is older compared to the EU or other developed countries; this is a common trend in developing countries. The significance of the correlation coefficient between the residual

terms,  $\rho$ , implies that the joint estimation is appropriate and that the sub-sample of loaded trips is not selected randomly.

Table 3 - FIML Heckman Model estimates for LF equation

Variables	Combined Model 0	Rigid Model 1	Articulated Model 1	Rigid Model 2	Articulated Model 2
Distance	-0.0097 (-11.43)	-0.0098 (-7.18)	-0.0117 (-11.54)		
Distance 1 (Distance×Commodity group 1)				0.0089 (4.35)	0.0043 (1.22)
Distance 2 (Distance×Commodity group 2)				0.0059 (1.53)	0.0161 (4.78)
Distance 3 (Distance×Commodity group 3)				0.0205 (12.62)	0.0068 (2.49)
Distance 4 (Distance×Commodity group 4)				0.0305 (11.79)	0.0152 (4.81)
Distance 5 (Distance×Commodity group 5)				0.0123 (5.66)	-0.0018 (-1.02)
Distance 6 (Distance×Commodity group 6)				0.0216 (8.71)	0.0044 (2.09)
Size	-0.0020 (-6.84)	-0.0030 (-6.51)	-0.0339 (-37.84)		
Size 1 (Size×Commodity group 1)				0.0036 (6.11)	-0.0001 (0.13)
Size 2 (Size×Commodity group 2)				0.0088 (16.08)	0.0003 (0.68)
Size 3 (Size×Commodity group 3)				0.0019 (3.49)	-0.0006 (-1.01)
Size 4 (Size×Commodity group 4)				-0.0057 (-7.44)	-0.0076 (-8.99)
Size 5 (Size×Commodity group 5)				0.0041 (9.09)	0.0010 (2.34)
Size 6 (Size×Commodity group 6)				-0.0054 (-7.95)	-0.0040 (-6.94)

Table 3 - FIML Heckman Model estimates for LF equation (continue)

	Combined	Rigid	Articulated	Rigid	Articulated
Variables	Model 0	Model 1	Model 1	Model 2	Model 2
Voluminous	-0.1290 (-22.58)	-0.1378 (-18.57)	-0.1135 (-14.35)	-0.0514 (-7.95)	-0.0742 (-7.77)
Constant	1.0224 (86.69)	1.0250 (59.25)	2.3124 (61.65)	0.5285 (63.33)	0.7482 (67.07)
Rho	-0.0330 (-1.20)	-0.0129 (-0.31)	-0.0419 (-0.76)	0.9706 (530.29)	0.5556 (87.67)
Sigma	0.4764 (240.17)	0.5071 (196.31)	0.3748 (135.84)	1.0509 (91.88)	0.8227 (46.02)
Hetero Log. Distance				-0.1149 (-128.62)	-0.1097 (-62.54)
No. of observations	53327	36321	17006	36321	17006

Table 4 - FIML Heckman Model estimates for L (market access decision) equation

	Combined	Rigid	Articulated	Rigid	Articulated
Variables	Model 0	Model 1	Model 1	Model 2	Model 2
Distance	0.0861 (38.70)	0.1101 (34.76)	0.0590 (18.54)	0.1102 (12.47)	0.1420 (7.13)
Size	-0.0082 (-10.38)	-0.0084 (-7.60)	0.0008 (0.29)	-0.0038 (-6.39)	0.0191 (8.12)
Age	-0.0134 (-9.45)	-0.0101 (-5.96)	-0.0242 (-8.62)	0.0020 (2.48)	-0.0038 (-1.80)
Net-Import	0.2286 (9.37)	0.2115 (7.15)	0.2737 (6.34)	0.0849 (4.79)	0.1909 (5.20)
Large Provinces	0.3204 (18.32)	0.3641 (16.23)	0.2014 (6.92)	0.1355 (9.41)	0.2370 (10.17)
Lambda2.L (distance)				0.7262 (12.68)	0.4590 (4.97)
Constant	0.5643 (20.07)	0.4688 (13.75)	0.4138 (3.62)	0.4689 (26.36)	-0.4114 (-4.37)
No. of observations	53327	36321	17006	36321	17006

In Models 0 and 1 in Table 3, trip distance appears to be negatively associated with the load factor, i.e., if the trip distance increases, the load factor decreases for both rigid and articulated trucks. This finding does not comply with earlier findings in the literature. However, for the improved Model 2, which accounts for unobserved heterogeneity and expand the parameterization of distance and size, the signs for all distance variables become positive (except for commodity group 5 for the articulated trucks but is not significant). There are significant differences between commodity groups. For commodity groups 3, 4 and 6, distance tends to be quite significant. This fits well with the fact that these commodities are usually among the most “flexible” in the sense that they do not require special equipment or handling. Hence, there is a higher chance of an increase in the overall load factor. Furthermore, the voluminous dummy variable, which indicates whether or not a commodity is voluminous, tends to reduce the load-factor. It is clear from that distance matters more for rigid trucks than for articulated trucks. Furthermore, the negative and highly significant values of Hetero Log distance imply that there is higher variability in the load factor for short distances, whereas the variability in the load decreases for longer distances.

The effect of trip distance is positive and significant on the market access decision in all models, indicating that loaded trucks are preferred for longer distance, irrespective of the type of truck. Size has a negative effect on the load factor in model 1 for both rigid and articulated trucks, which is similar to the finding of Abate [10]. However, when truck size was further classified according to the commodity groups in Model 2, it appears that size has a significantly negative effect on the load factor for commodity groups 4 and 6. This implies that increase in size leads to a decrease in the load factor. The interpretation is that these commodities are difficult to merge and load into a larger truck, while it is easy to fill the smaller trucks with such goods. On the contrary, commodity group 2 is very easy to collect and compact in the form of a concrete mass, and hence for larger trucks, it is much easier to combine this cargo and hence increase the overall load factor. The effect of size on the market access decision is different for rigid and articulated trucks. It has been shown that size has a negative and significant effect for rigid trucks (in model 1 and 2), but a positive and significant effect for articulated trucks (in model 2). This replicates the results of Abate [10], where size has a negative effect on the market access decision, while size-squared has a positive effect; this implies that carriers prefer high capacity trucks to be loaded.

Mostly, studies based on selection models find positive signs for “rho”. Although in Model 0 and 1 rho sign is negative, in Model 2 where three extensions were introduced the rho sign has become positive and quite significant; infers the credibility of Model 2. A detailed explanation is given in Abate [10] for why one might get a positive or negative sign. The positive rho sign might occur if unobserved effects that make a vehicle more likely to be loaded also cause the vehicle to have a higher load factor. A negative sign implies the opposite.

Elasticity for selected variables was estimated (see Table 5 and Figure 3), for the market access equation for Model 1 and Model 2. The elasticities show the proportionate change in the probability that a truck is loaded, for a proportionate change in the continuous explanatory variables. The effect of the distance is positive, i.e., the probability that the truck is loaded increases if the distance increases. For example, a 1 percent increase in distance will increase the probability that the truck is loaded by 0.13% for rigid trucks in Model 1. The elasticity estimates of size show that for rigid trucks, the probability that the truck is loaded increases

if the size decreases; the opposite is true in the case of articulated trucks. However, in Model 1, the effect of size is not significant for the articulated trucks. In Model 2, for example, a 1% increase in the size of rigid trucks decreases the probability that the truck is loaded by 0.17%, while a 1% increase in the size of articulated trucks increases the probability that the truck is loaded by 0.68%. For articulated trucks, the elasticity effect of the age of the truck is such that an increase in age will lead to more empty runs. This is in line with expectations. For the articulated trucks in Model 2, however, the effect of age is not significant.

Table 5 - Elasticity estimates based on truck type for Market Access Equation (L)

Variables	Model 1		Model 2	
	Rigid	Articulated	Rigid	Articulated
Distance	0.1687	0.1286	0.1518	0.1381
Size	-0.1452	0.0087	-0.1713	0.6552
Age	-0.0852	-0.0894	0.0357	-0.0764

Note: The elasticities are calculated by the Delta method. The elasticities show percentage changes in the probability that the truck is loaded or empty for a proportionate change in the explanatory variables.

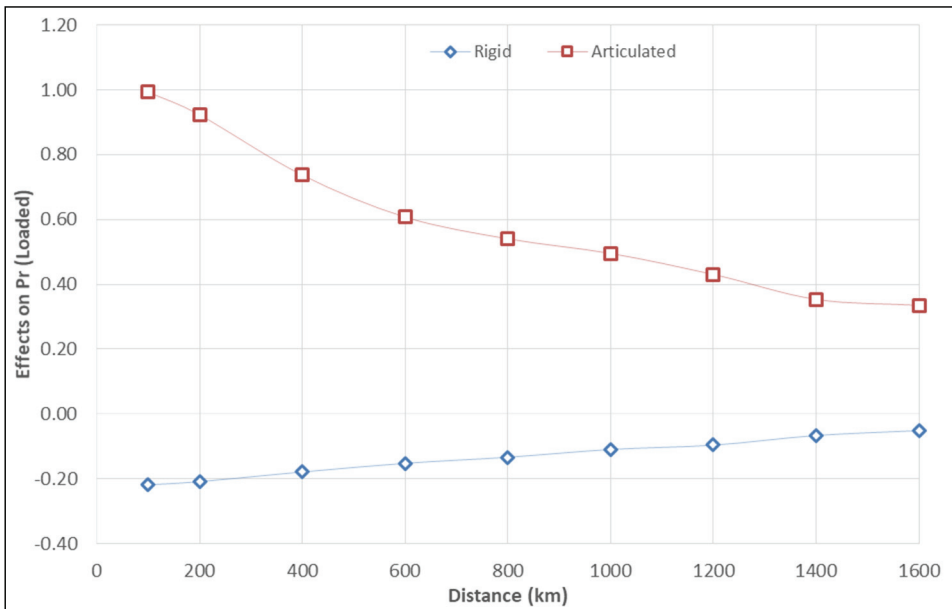


Figure 3 - Average marginal effects of truck size on the probability that the truck is loaded based on Model 2, shown as a function of distance

### 4.1. Validation

Figure 4 and 5 presents a validation of Model 2. To do this, the modeled probability distribution that the truck is loaded against distance and size is plotted, and compared this with the observed distributions (in the survey). It is seen that model closely predicted the observed distributions. As discussed above, the average number of loaded trips increases with increasing trip length. This is the case for rigid as well as for articulated trucks. The corresponding probability distribution with respect to the size of the truck shows an interesting pattern. For articulated trucks, the probability that the truck is loaded increases with size, whereas it is the other way round for rigid trucks. Figure 4 and 5 underlines the fact that the transport markets for rigid and articulated trucks are widely different.

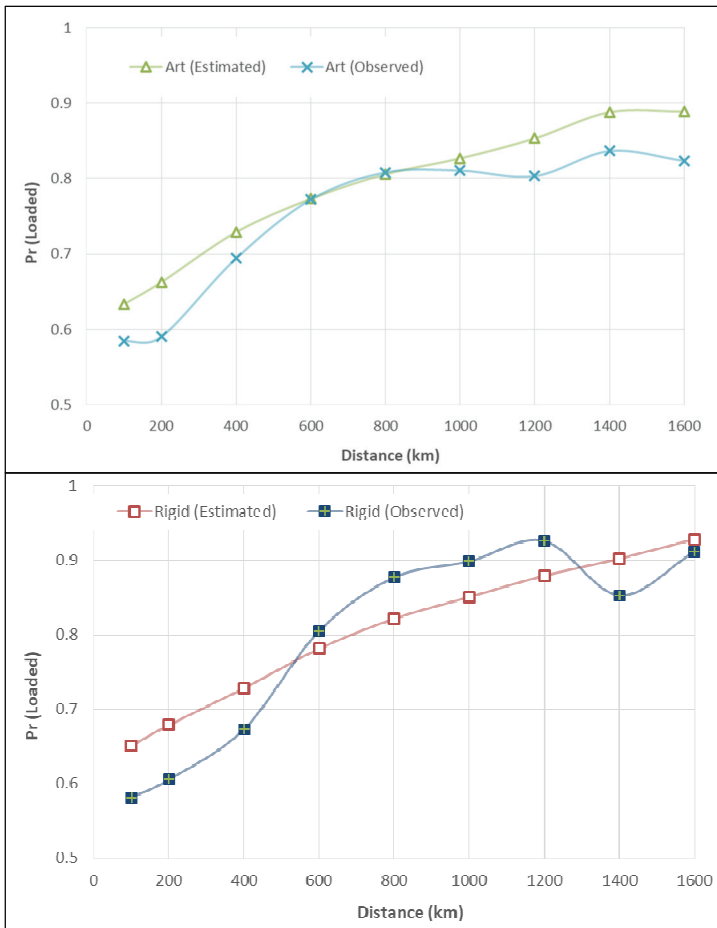


Figure 4 - Probability that the truck is loaded as a function of distance

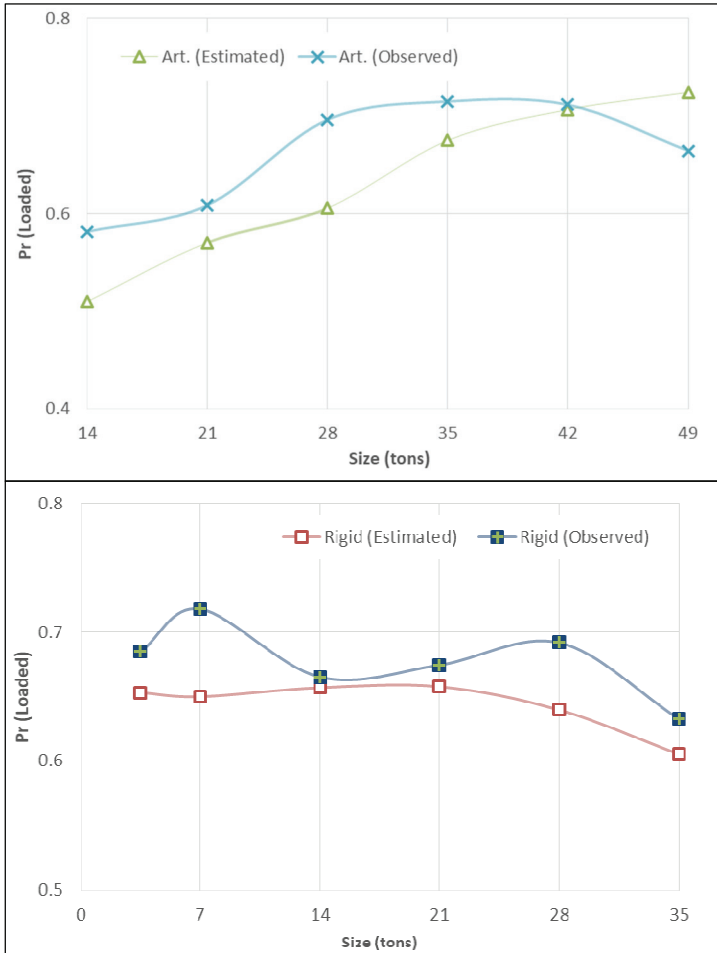


Figure 5 - Probability that the truck is loaded as a function of truck size

## 5. CONCLUSION AND DISCUSSION

Road freight transport corresponds to 89% of overall freight transportation in Turkey, which is very imbalanced, but is not expected to change significantly for many years to come, as alternative modes are not developed properly. This obviously increases the importance of efficiency considerations for the trucking industry because this is a main entry point to change the environmental footprint and economic impact in the country as a whole. The objective of this study was to investigate the main factors affecting capacity utilization in truck freight transport, in Turkey. The current paper extends the work by Abate [10] in three ways. Firstly, this study has introduced econometric extensions by partly relaxing the assumption of linearity of distance effects in the empty run model and by relaxing the assumption of homogeneity in the specification of the variance in the load factor model. Secondly, this paper



has presented a more explicit representation of commodity groups for which we estimate separate parameters with respect to distance and size effects in the load-factor model. Thirdly, this study considers different models for rigid and articulated trucks, which, in turn, are found to be important.

The results suggest that trip distance positively influences both the load factor and the probability of loaded trips. In the baseline model, the size of the truck was shown to have a negative relationship with the load factor and the probability that the truck is loaded. However, in a more advanced model where models are segmented according to the type of truck, it was shown that rigid and articulated trucks behave differently. Hence, the probability that an articulated truck is loaded is positively affected by size, whereas the opposite is true for rigid trucks. The fact that these two types represent essentially different transport markets is interesting for Turkey as well as other developing countries where the same mix of vehicles exist. It may well signal that transport policies aimed at improving efficiency for developing countries should be able to differentiate between the type of truck and possibly even be targeted toward different weight categories. Another important finding is the importance of the commodity type. As it can be seen in the modeling results, the capacity utilization pattern is quite different between commodity groups when combined with size and distance separately. While some commodity types resulted in a net positive relation for size, others had negative ones suggesting that probability of being loaded was higher for them. Thus, any further plans focusing on minimization of the empty or inefficient movements on highways must be developed considering commodity types and their flow characteristics on the intercity corridors in Turkey.

### **Further Recommendations**

Although the paper is a step toward better understanding capacity utilization in the trucking industry, the current findings may have some generic value. For instance, the findings about how truck size influences capacity utilization may have implications for legislation related to allowing mega trucks. Hence, as the size of trucks increases, it may be possible to derive indicative measures on how this will impact the utilization of capacity. Secondly, increase in loading probability by transport distance is a simple reflection of increase in total cost, if truck is empty, which is simply pressurizing drivers to find load for longer trips, while they can risk traveling empty on shorter distances. A complete understanding of capacity utilization for road transport requires further research. Most importantly, more data is needed, as described below:

- A correct representation of freight density and value density is important.
- A better description of the freight operator in order to take account of the heterogeneity between small and large companies is necessary.
- A more detailed representation of logistic activities is needed. Such data includes information about delivery routing, trip chains, warehouses, loading and unloading at transitional locations, as well as restrictions in terms of time windows and dependence on other operators and modes. These hidden constraints may potentially influence and reduce the potential for capacity utilization.

- A more detailed monitoring of Turkish trucks is needed, including intra-regional trips, and not just trips between regions.

Thus, it is necessary to collect national commodity flow data (CFD) as in some developed countries. A detailed discussion of such a national freight data framework was proposed in TRB [27], with many aspects discussed more recently in detail in a workshop [28] to improve the quality of commodities to be transported and to improve commodity flow data [29]. Similarly, the need for building the origin-destination freight (and passenger) matrices for pan-European region was the focus of the project titled “Methodological Framework for Modeling European Passenger and Freight Transport (MYSTIC)”, which also discussed involvement of various parties in freight transportation, such as carriers, intermodal platforms, etc. During this European Union Harmonization Process, as for a candidate country, it is also necessary to push for more integrated and complete freight data in Turkey, bringing public and private parties together to share their data.

Furthermore, freight transportation is an important factor for the development of the national economies. It is always desirable to minimize total cost of freight transportation (private and external costs), while goods are efficiently transported. Private costs are the direct expenses incurred by providers of freight transportation. Freight transportation has also costs that are paid by others. Such costs are referred as external costs (i.e., accidents, emissions, and other costs associated with the operation and maintenance of roads and bridges). If policies are adopted to charge the full cost of freight transportation, transportation users start to consider these costs when making shipping decisions. This may cause major structural changes to reduce external costs. One impact would be reduction of hauling distances by bringing production centers closer to consumption ones [30,31]. Another one would be combining inefficient movements to achieve higher level of capacity utilization in freight. However, to manage the freight sector in a more efficient and successful way, it is also necessary to invest on Intelligent Transportation Systems (ITS) in data collection as well as freight planning and management [32].

## References

- [1] T.C. Ulaştırma Denizcilik ve Haberleşme Bakanlığı, İstatistiklerle Ulaştırma Denizcilik ve Haberleşme: 2003-2016, Strateji Geliştirme Başkanlığı, 2016.
- [2] Organisation for Economic Co-operation and Development/International Transport Forum, ITF Transport Outlook, OECD Publishing, 2017.
- [3] Eurostat, Freight transport statistics, 2017. [http://ec.europa.eu/eurostat/statistics-explained/index.php/Freight\\_transport\\_statistics](http://ec.europa.eu/eurostat/statistics-explained/index.php/Freight_transport_statistics) Last Access: April 1, 2018.
- [4] European Commission, White Paper on Transport: Roadmap to a Single European Transport Area: Towards a Competitive and Resource-efficient Transport System. Publications Office of the European Union, 2011.
- [5] Abate, M., Kveiborg, O., Capacity Utilisation of Vehicles for Road Freight Transport, in: Ben-Akiva, M.E., Meersman, H., Voorde, E. van de (Eds.), Freight Transport Modelling. Emerald Group Publishing, Bingley, UK, 287–304, 2013.

- [6] Ozen, M., Tuydes-Yaman, H. Evaluation of emission cost of inefficiency in road freight transportation in Turkey. *Energy Policy*, 62, 625–636, 2013.
- [7] Piecyk, M. I., Analysis of Long-Term Freight Transport, Logistics and Related CO<sub>2</sub> Trends on a Business-as-Usual Basis. Ph.D. Thesis, Heriot-Watt University, 2010.
- [8] McKinnon, A. C., Benchmarking Road Freight Transport. Review of a Government-Sponsored Programme. *Benchmarking: An International Journal*, 16 (5), 640 – 656, 2009.
- [9] Beilock, R., Kilmer, R. L.. The Determinants of Full-Empty Truck Movements. *American Journal of Agricultural Economics*, 68(1), 67–76, 1986.
- [10] Abate, M., Determinants of Capacity Utilisation in Road Freight Transportation. *Journal of Transport Economics and Policy*, 48(1), 137–152, 2014.
- [11] Abate, M., de Jong, G., The Optimal Shipment Size and Truck Size Choice – The Allocation of Trucks Across Hauls. *Transportation Research Part A: Policy and Practice*, 59, 262–277, 2014.
- [12] Heckman, J. J., Sample Selection Bias as a Specification Error. *Econometrica*, 47(1), 153–161, 1979.
- [13] Ozen, M., Estimation of Road Freight Transportation Emissions in Turkey. Ph.D Thesis. Middle East Technical University, 2013.
- [14] Turkish General Directorate of Highways, Karayollarında Ağır Taşıt Trafiğinin ve Yük Taşımacılığının Özellikleri ve Eğilimleri: 2007-2008-2009 Yılı Etüt Sonuçları. Turkish General Directorate of Highways, 2011.
- [15] Turkish General Directorate of Highways, Karayollarında Ağır Taşıt Trafiğinin ve Yük Taşımacılığının Özellikleri ve Eğilimleri: 2010-2011-2012-2013-2014 Yılı Etüt Sonuçları. Turkish General Directorate of Highways, 2011.
- [16] Unal, L., Modeling of Freight Transportation on Turkish Highways. Ph.D Thesis. Middle East Technical University, 2009.
- [17] Wooldridge, J. M., *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Mass: The MIT Press, 2010.
- [18] Portugal-Perez, A., Wilson, J. S., Export Performance and Trade Facilitation Reform: Hard and Soft Infrastructure. *World Development*, 40(7), 1295-1307, 2012.
- [19] de Jong, G., Tanner, R., Rich, J., Thorhauge, M., Nielsen, O. A., Bates, J., Modelling Production-Consumption Flows of Goods in Europe: The Trade Model Within Transtools3. *Journal of Shipping and Trade*, 2(1), 5, 2017.
- [20] Tsotsotso, K., Montshiwa, E., Tirivanhu, P., Fish, T., Sibiya, S., Mlangeni, T., Moloi, M., Mahlangu, N., Determinants of Skills Demand in A State-Intervening Labour Market: The Case Oof South African Transport Sector. *Higher Education, Skills and Work-Based Learning*, 7(4), 408-422, 2017.
- [21] Arabmazar, A., Schmidt, P., An Investigation of the Robustness of the Tobit Estimator to Non-Normality. *Econometrica*, 50(4), 1055–1063, 1982.

- [22] Maddala, G. S., Nelson, F. D., Specification Errors in Limited Dependent Variable Models. NBER Working Paper No: 0096. National Bureau of Economic Research, Inc., 1975.
- [23] Nelson, F. D., The Effect of and a Test for Misspecification in the Censored-Normal Model. Working Paper No: 291. California Institute of Technology, Division of the Humanities and Social Sciences, 1979.
- [24] Hurd, M., Estimation in Truncated Samples when There is Heteroscedasticity. *Journal of Econometrics*, 11(2–3), 247–258, 1979.
- [25] Sakia, R. M., The Box-Cox Transformation Technique: A Review. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 41(2), 169–178, 1992.
- [26] Box, G. E. P., Cox, D. R., An Analysis of Transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, 26(2), 211–252, 1964.
- [27] Transportation Research Board of the National Academies, Special Report 276: A Concept for a National Freight Data Program. Transportation Research Board of the National Academies, 2003.
- [28] Transportation Research Board of the National Academies, Transportation Research Circular: Commodity Flow Survey Conference. Transportation Research Board of the National Academies, 2006.
- [29] US Census, Commodity Flow Survey, 2015. <https://www.census.gov/programs-surveys/cfs/data/datasets.html> Last Access: September 20, 2018.
- [30] Forkenbrock, D. J., External Costs of Intercity Truck Freight Transportation. *Transportation Research Part A: Policy and Practice*, 33, 505–526, 1999.
- [31] Forkenbrock, D. J., Comparison of External Costs of Rail and Truck Freight Transportation. *Transportation Research Part A: Policy and Practice*, 35, 321–337, 2001.
- [32] Crainic, T. G., Gendreau, M., Potvin, J. Y., Intelligent Freight-Transportation Systems: Assessment and the Contribution of Operations Research. *Transportation Research Part C: Emerging Technologies*, 17(6), 541–557, 1999.