Issue 22(2), 2022, pp.22-52
https://doi.org/10.18757/ejtir.2022.22.2.5662

TU Delft OPEN

EJTIR
ISSN: 1567-7141
http://ejtir.tudelft.nl/

Building latent segments of goods to improve shipment size modelling: Confirmatory evidence from France

Raphael Piendl[1]
Institute of Transport Research, German Aerospace Center (DLR), Germany.

Martin Koning[2]
AME – SPLOTT, Gustave Eiffel University, France.

François Combes[3]
AME – SPLOTT, Gustave Eiffel University, France.

Gernot Liedtke[4]
Institute of Transport Research, German Aerospace Center (DLR), Germany.

Freight transport demand models are generally based on administrative commodity type segmentation which are usually not tailored to behavioral freight transport demand modelling. Recent literature has explored new approaches to segment freight transport demand, notably based on latent class analysis, with promising results. In particular, empirical evidence from road freight transport modelling in Germany hints at the importance of conditioning and handling constraints as a sound basis for segmentation. However, this literature is currently sparse and based on small samples. Before it can be accepted that conditioning should be integrated in the state-of-the-art doctrine of freight data collection and model specification, more evidence is required. The objective of this article is to contribute to the issue. Using detailed data on shipments transported in France, a model of choice of shipment size with latent classes is estimated. The choice of shipment size is modelled as a process of total logistic cost minimization. Latent class analysis leverages the wide range of variables available in the dataset, to provide five categories of shipments which are both contrasted, internally homogenous, and directly usable to update freight collection protocols. The groups are: "‘*standard temperature-controlled food products*"’, "‘*special transports*"’, "‘*bulk cargo*"’, "‘*miscellaneous standard cargo in bags*"’, "‘*palletised standard cargo*"’. This segmentation is highly consistent with the empirical evidence from Germany and also leads to better estimates of shipment size choices than administrative segmentation. As a conclusion, the finding that conditioning and handling information is essential to understanding and modelling freight transport can be regarded as more robust.

Publishing history

Submitted: 20 March 2021
Accepted: 2 February 2022
Published: 11 April 2022

Cite as

Piendl, R., Koning, M., Combes, F. & Liedtke, G. (2022). Building latent segments of goods to improve shipment size modelling: Confirmatory evidence from France. *European Journal of Transport and Infrastructure Research, Issue 22(2)*, 22-52.

Keywords: *freight transport, shipment size choice, total logistics cost, segmentation universality*

[1] A: Institute of Transport Research, German Aerospace Center (DLR), Rutherfordstraße 2, 12489 Berlin, Germany.
E: raphael.piendl@gmail.com
[2] A: AME – SPLOTT, Gustave Eiffel University, 14-20 Bvd Newton, 77447 – Marne-la-Vallée, France. E: martin.koning@univ-eiffel.fr
[3] A: AME – SPLOTT, Gustave Eiffel University, 14-20 Bvd Newton, 77447 – Marne-la-Vallée, France. E: francois.combes@univ-eiffel.fr
[4] A: Institute of Transport Research, German Aerospace Center (DLR), Rutherfordstraße 2, 12489 Berlin, Germany.
E: gernot.liedtke@dlr.de

1. Introduction

Freight transport is a heterogenous market. Freight transport demand consists of very diverse firms, from small businesses dispatching parcels directly to customers, to ten thousand employees' industrial firms procuring thousands of tons of raw materials to produce steel, cars, or computer components. This heterogeneity is a major challenge for freight transport demand modelling, because different firms, set in different geographical and economic environments will not react homogenously to changes in infrastructure networks or transport prices, and should not be modelled as such. The overwhelmingly preferred approach to address this heterogeneity is through discrete explicit segmentations, such as the economic activities of shippers, or administrative commodity types (Tavasszy and de Jong, 2014).[1]

From the perspective of freight data production, a segmentation per commodity type is both easy to implement and statistically sound. Indeed, one can easily derive correlations between the characteristics of sending and receiving firms or regions and the corresponding freight flows for each type of commodity (a necessary step in the specification and calibration of a freight demand model, which essentially computes freight flows as functions of geographic variables, which are both exogenous, and relatively easy to forecast). The current standard for freight data production in Europe is based on a commodity type segmentation, the NST 2007 (European Parliament and Council, 2012). However, an ideal classification should be concise (not too many segments), practical (data production should be feasible and reliable) and relevant (provide useful information for diagnosis, avoid imprecision and biases in models). From this perspective, a segmentation based on commodity type is not perfect (Liedtke and Schepperle, 2004). To illustrate the issue, consider the example of milk: at the upper stages of the supply chain, it is typically transported in bulk, in large, dedicated trucks. On the contrary, at the lower stages, it is typically transported in palettised packs or bottles. The underlying freight transport decisions are very different; and yet, based on a commodity type classification, the two situations are statistically identical.

One promising direction towards a better segmentation is to look at logistics variables, such as handling and conditioning. This is particularly visible when considering models of choice of shipment size. The choice of shipment size is an interesting decision to model, because it lies at the core of the trade off shippers must make when organizing freight transportation – such as the trade-off between transport costs and inventory costs. A segmentation relevant for modelling the choice of shipment size is relevant for freight modelling in general.

Piendl et al. (2017) estimated a model of discrete choice of shipment size based on a Total Logistic Cost specification, with a segmentation derived by Latent Class Analysis (LCA). Based on handling and conditioning variables, four segments were identified: "‘*(temperature-controlled) food products*"’, "‘*miscellaneous standard cargo*"’, "‘*special goods*"’ (meaning shipments with specific handling constraints, such as specific dimensions, etc.), and "‘*non-packed bulk goods*"’. The fit of the model shipment size choice is much better with this data-driven segmentation than with a standard commodity type based one, with only a fraction of segments. The limited number of segments is a direct result of the LCA approach:

[1] In this article the term "‘commodity type"’ refers to classification schemes which are used in official administrative statistics.

in shipper surveys, shipments are typically described with several categorical variables, and there is no trivial way to summarize efficiently that information. Another approach to represent heterogeneity would be to use a mixed logit approach (Train, 2008) which, it is worth reminding, can approximate any random utility model to any degree of accuracy (McFadden and Train, 2000). However, doing so would be missing the opportunity to leverage the information available in shipper surveys, and would provide little information regarding how freight data collection protocols should be improved in the future: this is one of these situations where improving the statistical fit of models and improving their practical usefulness are two distinct objectives. Finally, the approach comes with corollary qualities: in a companion article, Piendl et al. (2019) showed how these segments could be matched with classical commodity flow data, with a machine learning approach, so as to simulate the impact of exogenous changes (information and communication technology in their article) on shipment sizes and transport operation.

Now, before the introduction of handling and conditioning variables into freight database production and freight modelling state-of-the-art can be seriously considered, two questions need answers. First: are the results discussed above robust? Second, which modalities should be observed? Conditioning and handling characteristics are extremely varied. It is not the same thing to assert that it is sufficient to distinguish parcels from pallets and bulk, than to claim that the gross and net weight, dimensions, and corrugated cardboard width are to be precisely measured for every unit in every sampled vehicle. From that perspective, the conclusions of Piendl et al. (2017) are not yet robust enough to proceed. This is mainly due to the sample size and characteristics: the model is estimated on the basis of a German shipper survey, consisting of 487 shipments. The survey was focused on long haulage distance transport, with shipment size above one ton, dispatched by German firms of which the main activity is related to raw material extraction, manufacturing, or wholesale trade.

Against this background, the main objective of this article is to contribute to identifying a more efficient segmentation for freight data collection and freight transport modelling. To do so, we apply the methodology of Piendl et al. (2017) to the French commodity flow survey ECHO (Guilbault and Soppé, 2009), which consists of a large sample (finally 2219 observations) of very diverse shipments with detailed information about shipper-receiver relationships and logistical attributes. The first practical objective is to confirm that a sound segmentation, identified with LCA, strongly improves the quality of a model of choice of shipment size, as compared to conventional classifications per administrative commodity type. This is examined by comparing the models' fits, and by a ten-fold cross-validation approach to assess the model's predictive accuracy. The second objective is to compare the segmentation obtained on the basis of the ECHO dataset to that obtained from the German dataset and determine whether there is some comparability, and, therefore, the degree of generality of these results. Note that the ECHO dataset, which contains many variables, offers the opportunity to further refine the specification of the shipment size: additional factors related to storage costs, and to fixed costs of transport or ordering, are considered.

The remainder of this article is organised as follows: Section 2 reviews the scientific literature relevant to shipment size modelling and dealing with the heterogeneity in the context of freight research. Section 3 presents the model's theoretical basis and specifications, as well as the main principles of the LCA. Section 4 presents the ECHO dataset whereas Section 5 focuses on the empirical analyses. Section 6 synthesizes our results and calls for further research.

2. Literature review

Introducing logistics dimensions in freight transport models is generally limited by the availability of individual or structural data (firms, flows), even at a meso-level. Thus, models can't have an excessive number of parameters and need be able to predict the behaviours of stakeholders. In this respect, shipment size is an attractive opportunity: On the one hand, it is located in between intra-company logistics processes and inventory decisions; on the other hand, it makes explicit the relationships with transport markets (see Islam et al. (2013) for a overview of logistics and its interacting elements).

2.1 The choice of shipment size

The seminal shipment size model is given by the normative Economic Order Quantity (EOQ) model (Harris, 1913) which determines optimal shipment size by finding a suitable trade-off between fixed costs of transport/ production (i.e. the part of transport and production costs which does not depend on shipment size) and inventory holding costs. The idea of introducing inventory considerations in freight transport modelling is taken up again by Baumol and Vinod (1970) who analytically extend the EOQ-model to mode choice. Moreover, they put forward the concept of TLC. The inputs of the TLC are the total commodity flow between shippers and receivers, the storage costs, the shipping costs and the fixed costs of ordering/transport. Hall (1985) analytically examines the dependence of shipment size and vehicle types. He concludes that the optimal shipment size is a discontinuous function of the commodity flow between shippers and receivers (this discontinuity is discussed in depth in the recent literature; see Piendl et al. (2017)).

Overall, there are many options to model shipment size in the empirical literature. Some authors consider shipment size as a continuous variable (Sakai et al., 2020; Koning et al., 2018; Abate and de Jong, 2014; Combes, 2009; Holguín-Veras, 2002; Abdelwahab and Sargious, 1992), others as discrete (Keya et al., 2019; Abate et al., 2019; Piendl et al., 2017; Pourabdollahi et al., 2013; Windisch et al., 2010; de Jong and Ben-Akiva, 2007). Some articles assume shipment size is unaffected by mode choice (Sakai et al., 2020; Kawamura et al., 2010; Wisetjindawat et al., 2005) or conditional on it (mode-specific influences are included as binary indicators (Combes, 2012)), whereas others account for the interdependencies between decisions, either implicitly or explicitly (Keya et al., 2019; Abate et al., 2019; Irannezhad et al., 2017; Combes and Tavasszy, 2016; Abate and de Jong, 2014; Pourabdollahi et al., 2013). Some authors opt for an empirical approach where all potentially relevant explanatory variables are tested (Keya et al., 2019; Pourabdollahi et al., 2013; Windisch et al., 2010; Holguín-Veras, 2002); as opposed to others who start from a microeconomic analytical model – typically the TLC framework – to derive a structural form which is then empirically tested (Sakai et al., 2020; Abate et al., 2019; Koning et al., 2018; Combes and Tavasszy, 2016;

Abate and de Jong, 2014; Liedtke, 2012; de Jong and Ben-Akiva, 2007).

In summary, introducing behaviour-sensitive shipment size choice in freight transport demand models is a fruitful research direction. However, substantial barriers remain including estimation and operational practicality, as well as the representation of the heterogeneity due to the multitude of involved actors and shipments.

2.2 Heterogeneity of logistics requirements

Apart from the variation in shipment sizes originating in the trade-off between transport costs and inventory costs, heterogeneity comes alongside various logistics aspects like handling requirements, packaging or storage conditions (called *logistical requirements* hereinafter). Some articles approximate their impacts with information about the shipper and/or the receiver (Sakai et al., 2020; Pourabdollahi et al., 2013; Holguín-Veras, 2002), or about the transported commodity itself (Sakai et al., 2020; Abate et al., 2019; Keya et al., 2019; Koning et al., 2018; Abate and de Jong, 2014; de Jong and Ben-Akiva, 2007).

The simplest way to reduce heterogeneity is to introduce an exogenous segmentation (Samimi et al., 2014; de Jong et al., 2013, 2010), especially the one describing the nature of the commodities. In Europe, the Standard Goods Classification for Transport Statistics (NST) – introduced at the beginning of the 1960's and revised in 1968 to become the NST/R (10 chapters, 52 main groups, 176 headings) – was based on criteria such as physical properties of commodities, processing stages, and transport methods. In 2007, the NST/R was replaced by the NST 2007 (20 chapters) based on the European Classification of Products by Activity (CPA). These nomenclatures are readily available in every country. However, they mix different dimensions and still, a large amount of heterogeneity remains.

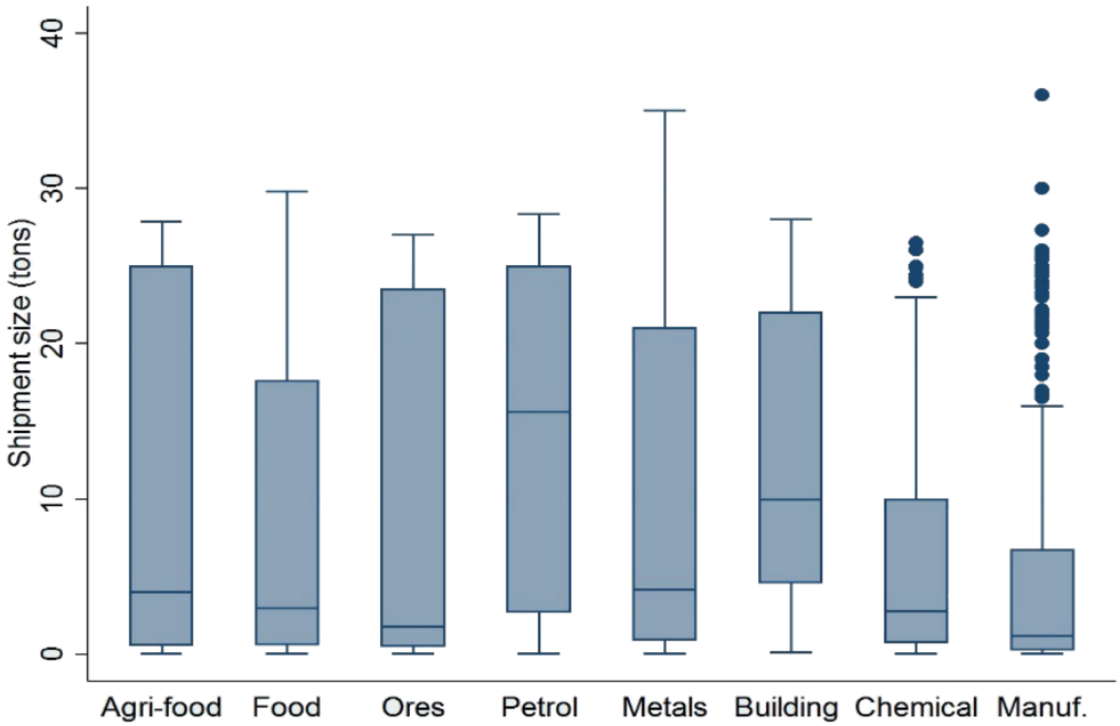
Consider, for example, the distributions of shipment size with respect to commodity types found in the French ECHO dataset. Figure 1 shows a boxplot of shipment size distributions across the NST chapters proposed in the ECHO database. Each box ranges from the 0.25 to the 0.75 quantile also containing a 0.5 quantile (median) bar in it. So called whiskers of the boxes indicate the values which are located in between 1.5 times the interquartile range. [2] Not only the variation within each chapter is large, but some chapters display very similar distributions. As a consequence, standard goods classifications may be suboptimal regarding the description of logistical requirements and shipment sizes (Piendl et al., 2017; Arunotayanun and Polak, 2011). [3]

Another approach which is applied in aggregate freight transport demand models is given by so-called logistic families Tavasszy (2000); Tavasszy et al. (1998). Hereby, categories are built rather heuristically with deterministic thresholds and thus also suboptimal classification schemes might occur.

[2] Figure 1 is based on the obsolete NST/R nomenclature. However, each chapter of the NST 2007 still contains a high degree of heterogeneous actors and transport operations due to the above-mentioned theoretical aspects

[3] Figure 1 is used for illustration purposes. Of course, the variation in the NST chapters might be to some extent explained by key influencing variables identified from the Total Logistics Costs framework. Nonetheless, no systematic differentiation between the NST chapters is observable.

Figure 1: Boxplots of shipment sizes carried by road according to NST/R chapters proposed in the ECHO dataset.



There are only 8 chapters in the ECHO dataset instead of the 10 NST/R chapters. Chapters 4 and 5 ("'Ores and metal waste"' and "'Metal products"') were merged into a specific "'Metals"' chapter, and Chapters 7 and 8 ("'Fertilizers and Chemicals"') were merged into a specific "'Chemical"' chapter. For consistency with the rest of the article, this graph only presents shipments carried by road transport.

In order to deal efficiently with heterogeneity, a well-known method in transportation science is to use mixed logit models which assume predefined distributions of the estimated parameters among the population under study (McFadden and Train, 2000). Since the variables related to logistical requirements are mostly categorical, the natural direction would be to use discrete mixing distributions. Such kind of models categorize the observations into segments and estimate the related choice models (shipment size, transport mode) simultaneously whilst minimising the heterogeneity within each segment and maximising the heterogeneity between the segments. Examples of applications are Greene and Hensher (2003) or Arunotayanun and Polak (2011). The advantage is that the heterogeneity will be efficiently taken into account with an optimal number of segments. However, these models come at a cost: First, a full set of parameters is estimated for each segment. While the numerous degrees of freedom theoretically provide an optimal fit to the data, it cannot be ensured that the model shows desirable theoretical characteristics, such as an intuitively acceptable behaviour of the TLC function. Second as stated by Train (2008), the estimation of a full set of parameters quickly reaches the limit of gradient descent approaches with respect to the computational complexity. Thus, freight transport seems to be a big challenge due to its heterogeneity and how it can be measured as already indicated by Piendl et al.

(2019). When forecasting, the demand also needs to be assigned to each segment (including the firm-to-firm commodity flows). Given that this classification has a probabilistic nature, possible misclassifications might cause significant volatility of the outcomes when comparing projects or proceeding to sensitivity tests, as the underlying parameter estimates of the segments can greatly differ. When a freight transport model is used in an operational context, the results need to be interpreted and disseminated; as a consequence, the robustness of results may be prioritized over a perfect statistical fit.

An intermediate approach between using official segmentations and the simultaneous estimation of segments and choice models through mixed logit models is provided by the LCA. This exogenous segmentation method uses categorical indicators as input data to group observations on the basis of similar response patterns and is e.g. applied by Shelat et al. (2018) to obtain prototypical users of the combined bicycle and transit mode. Furthermore, the resulting segments can be applied subsequently which we show for shipment size models.

Within this context the general approach presented by Piendl et al. (2017) provides a promising basis in order to integrate shipment size choice as suitable, powerful and highly explanatory logistics decision into freight transport demand models. At the same time, the approach offers low complexity as well as low data intensity which supports the operational practicality. However, the article at hand now provides additional relevant implications to the behavioural-sensitive explanation and operational practicality of shipment size choice in the context of freight transport demand modelling. In more detail, the same decisive drivers of shipment size choice reveal robust across data sources and countries supporting their mandatory consideration in shipment size choice models. Moreover, homogenous segments gathered from indicators describing economic activities, handling and conditioning reveal superior to standard classification schemes regarding the explanatory and predictive power. Finally, the composition and effects of these segments can be also considered robust providing linkages to efficient freight database production which is a serious challenge in state-of-the-art freight transport demand modelling.

3. Modelling framework

Our model is based on a TLC formulation combined with LCA for segmentation. These methodologies are described shortly in the two following subsections. Piendl et al. (2019, 2017) provide a more in-depth presentation.

3.1 The optimal shipment size

The TLC is defined for a shipper denoted by n, over a given period (in the following considerations: per year). It consists of the three following components: the fixed costs of transport and/or ordering $C_{n,fix}$, the variable transport costs $C_{n,variable}$, and the storage costs $C_{n,storage}$:[4]

[4] In theory, pipeline inventory cost should be included (Baumol and Vinod, 1970). It is neglected here because pipeline inventory costs have no influence on shipment size in the EOQ model, and also because empirical analysis in Piendl et al. (2017) and in this article find no significant effect of pipeline inventory

$$C_n = C_{n,fix} + C_{n,variable} + C_{n,storage}. \tag{1}$$

The fixed costs $C_{n,fix}$ incurs each time a shipment is dispatched, regardless of its size. It describes the resources required to pass and to monitor the orders, to book a loading dock in warehouses, to pay the vehicle's insurance, etc. It is equal to the total amount of shipments dispatched per year, i.e. the ratio of the flow of goods Q_n [ton/year] sent to a given receiver and the shipment size q_n [ton/shipment], times the unit fixed transport costs F_n [€/shipment].

The variable transport costs $C_{n,variable}$ are assumed to be a function $c_n(q_n)$ [€/ton], expressed on a per ton basis, times the commodity flow rate.

In order to calculate the storage costs $C_{n,storage}$, a constant consumption rate is assumed; the average stock is then $q_n/2$. This average stock is multiplied by the unit storage costs s_n [€/ton] which includes warehousing costs as well as capital costs.

Accordingly, the refined and parametrized TLC is:

$$C_n(q_n) = \frac{Q_n}{q_n} F_n + Q_n c_n(q_n) + \frac{q_n}{2} s_n. \tag{2}$$

Assuming a constant per ton rate $c_n(q_n)$ independent of the shipment size, the optimal shipment size q_n^* is found by minimizing Eq. 2 with respect to q_n:

$$q_n^* = \sqrt{\frac{2F_nQ_n}{s_n}}. \tag{3}$$

This is the classical EOQ formula: The optimum shipment size increases with the commodity flow rate and with the fixed costs; it decreases with increasing unit storage costs. However, this model has some limitations: (i) there are strong assumptions regarding the structure of the transport costs, (ii) the storage costs are not distinguished from inventory cost, (iii) the vehicle capacity constraints are not accounted for, and (iv) different ways of organising transports (e.g. direct vs. consolidation) are not distinguished.

A pragmatic approach to address these issues is to consider the choice of shipment size as a discrete problem, instead of a continuous one. Denote by q_i the i-th possible discrete shipment size category. The shipment size dependent transport costs function c_n is now assumed variable and nonlinear (Liedtke, 2012): The per ton-kilometer variable transport costs, for a shipment size q_i, is $t_{q_i} d_n$, where t_{q_i} is a coefficient varying with shipment size category and d_n describes the haulage distance.

Second, the unit storage costs s_n can be detailed with, on one hand, the unit warehousing

cost on shipment size.

costs w_n and, on the other hand, the unit capital costs rv_n (where r is an interest rate and v_n the value density of the commodities). After normalizing for commodity flow rate, the TLC of shipment size category q_i for agent n is:

$$\frac{C_n\left(q_i\right)}{Q_n}=\frac{F_n}{q_i}+t_{q_i}d_n+\frac{q_i}{2}\frac{w_n}{Q_n}+\frac{q_i r}{2}\frac{v_n}{Q_n}. \tag{4}$$

Before presenting the econometric implementation of the specification implied by Eq. 4, the issue of the integration of logistical requirements using LCA is addressed.

3.2 Latent Class Analysis

LCA is a classification method which assigns observations to latent classes, or segments, based on observed variables (Collins and Lanza, 2010). Consider $j \in \{1, \ldots, J\}$ observable factors with nominal scale and $r_j \in \{1, \ldots, R_j\}$ the response categories for each variable j. S distinct latent segments need to be inferred. Denote by $\mathbf{y} = (r_j, \ldots, r_J)$ the observation's responses to the J discrete variables. Define by L the latent variable including the latent segments $s = 1, \ldots, S$ and denote by $I\left(y_j = r_j\right)$ an indicator equal to 1 if the response to variable j is r_j (0 otherwise). The probability $P\left(\mathbf{Y} = \mathbf{y}\right)$ of observing a particular response pattern is:

$$P\left(\mathbf{Y}=\mathbf{y}\right)=\sum_{s=1}^{S}\gamma_s\prod_{j=1}^{J}\prod_{r_j=1}^{R_j}\rho_{j,r_j|s}^{I(y_j=r_j)} \tag{5}$$

Two unknown sets of parameters are included in Eq. 5: γ_s, the latent class prevalence of class s; and $\rho_{j,r_j|s}$, the probability of observing response r_j to variable j conditional on the membership to class s. The individual membership probabilities (or posterior probabilities) can then be derived by applying Bayes' Theorem:

$$P\left(S=s|\mathbf{Y}=\mathbf{y}\right)=\frac{P\left(\mathbf{Y}=\mathbf{y}|S=s\right)P\left(S=s\right)}{P\left(\mathbf{Y}=\mathbf{y}\right)}=\frac{\left(\prod_{j=1}^{J}\prod_{r_j=1}^{R_j}\rho_{j,r_j|s}^{I(y_j=r_j)}\right)\gamma_s}{\sum_{s=1}^{S}\gamma_s\prod_{j=1}^{J}\prod_{r_j=1}^{R_j}\rho_{j,r_j|s}^{I(y_j=r_j)}}. \tag{6}$$

Accordingly, each observation is assigned to the class with the maximum posterior probability.[5] The global result of the LCA is the definition of latent segments made of observations sharing similar response patterns.

[5]Beside this approach, Lanza et al. (2013) investigate two different approaches to assign individual observations to segments. Nevertheless, the appropriateness of the maximum membership probability can be supported by having a high value for the entropy measure (close to 1) and an odds of correct classification (OCC) value higher than 5 (Collins and Lanza, 2010; Nagin, 2005)

LCA is a particularly relevant approach to model shipment size choices because logistical requirements are measured by a large set of categorical variables (see Section 4). The issue is to integrate these variables in the model, in order to identify their influence, and to improve the model's explanatory power without making it overly complex. Put differently, LCA summarises the information provided by various logistical requirement variables into a limited set of dummy variables so that the model is both simple and relevant.

3.3 Econometric specification

The specification of the discrete shipment size model which will be estimated is based on Eq. 4. In general, the TLC function is interpreted as a negative utility and we add an extreme value type I i.i.d error term $\epsilon_{q_i,n}$, thus resulting in a multinomial logit model.[6] The utility of shipment size q_i for agent n is given by:

$$\begin{aligned} -U_{q_i,n} &= \frac{C_n\left(q_i\right)}{Q_n} + \epsilon_{q_i,n} \\ &= \alpha_{q_i} + \beta_{q_i,1} \cdot \mathbf{F}_n + \beta_{q_i,2} \cdot d_n + \beta_{q_i,3} \cdot \frac{w_n}{Q_n} + \beta_{q_i,4} \cdot \frac{v_n}{Q_n} + \beta_{q_i,S} \cdot \mathbf{S}_n + \epsilon_{q_i,n}. \end{aligned} \tag{7}$$

As in every empirical model, not all components are directly measurable and thus suitable approximations need to be found: fixed costs of ordering/transport do not appear directly in the model because they cannot be directly measured. Instead, we will consider a vector of dummies $\mathbf{F}_n$ which capture some of the corresponding heterogeneity. More precisely, they are approximated by the dichotomous variables explained in Table 1 "'Dest_foreign"', "'No_application"', "'Half_day_arrival"' and "'Multi_trips"'.

Storage costs are also not directly measured in the data. In accordance with Eq. 4 and Eq. 7, they are represented by two variables: The first one is the inverse of the density of the commodities $[\mathrm{m}^3/\mathrm{t}]$ and corresponds to the variable "'Vol_ton"' from Table 1. This variable intuitively approximates warehousing costs w_n, given the fact that they capture the need for warehousing volume. The second one is the value density v_n, as in Eq. 4. Both variables are here normalized by the commodity flow rate Q_n. The distance dependent component of transport costs is kept in the model by variable d_n. Above all, we add the vector of dummies $\mathbf{S}_n$ which represent the membership to segments $s = 1, \ldots, S$, found through the LCA. The size, sign and significance of the $\beta_{q_i,S}$ parameters are of prime interest because they inform us about differentiated impacts of the data-driven segments on the optimal shipment size.

4. Data

The data comes from the ECHO survey (Guilbault and Soppé, 2009), collected in 2004-2005 and describing 10,462 shipments dispatched by about 3,000 shippers. The ECHO survey con-

[6]We have also tested mixed logit specifications, which did not improve the results. See the discussion in subsection 5.3.

sists of four distinct parts focused on: the shipper, the shipment and the shipper-receiver relationship, the transport operation, and the firms involved in the transport operation. As a consequence, each shipment is described with many variables (about 700). The main distinction between the ECHO survey and other commodity flow surveys is that the former observes the shipper-receiver relationship and the transport operation very detailed. Particularly, the ECHO survey includes information about the most informative variable for shipment size choice: the total volume in tonnes per year to be transported between a shipper and its recipient (Koning et al., 2018).

In order to facilitate the comparison of results with those proposed for Germany by Piendl et al. (2017), we here focus only on road shipments. It is worth noting that main variables such as the total shipper-receiver commodity flow rate or the value density are not available for all shipments. Moreover, shipments below 30kg were removed because they belong to parcel services, which is a very specific type of road transport (outside the scope of this article). Finally, the ECHO dataset assumes that all shipments of less than 0.5m^3 are 0.5m^3. In order to avoid a possible biased estimate, shipments with a volume of 0.5m^3 are excluded from the estimations. All in all, a final dataset with 2,219 observations is used for model estimations.

Table 1: Summary statistics of the TLC variables (2,219 observations).

Variable	Definition	Mean	Std.dev.	Min	Max
q	Shipment size (tons)	7.06	8.94	0.03	36
Q	Flow of goods (tons/year)	2,171.85	11,206.10	0.05	250,000
v	Value density (€/ton)	11,990.6	99,167.43	0.39	3,102,703
d	Haulage distance (km)	274.58	281.25	0.000095	2,713.42
Vol	Volume of the shipment (m^3)	20.23	30.54	1.00	490
Vol_ton	Volume density (m^3/ton)	9.63	42.72	0.093	1,002.05
Dest_foreign	1 if destination of shipment is abroad	0.14	0.35	0	1
No_application	1 if no software applications are used for transaction	0.36	0.48	0	1
Half_day_arrival	1 if delivery time is decided half a day or more before delivery	0.62	0.49	0	1
Multi_trips	1 if shipment made multiple trips	0.33	0.47	0	1

Table 1 presents the summary statistics of the variables of interest for this paper. All variables describing the standard components of the TLC function are self-explanatory, except the volume density ("‘*Vol_ton*"’) which is the ratio of the shipment’s volume and its tonnage and which is rarely integrated in shipment size models due to the lack of information. In the same vein, we take advantage of the many variables available in the ECHO survey in order to approximate the influence of the fixed costs of ordering and/or transport on shipment size.

The *Dest_foreign* variable aims at describing potential border effects or administrative works implied by transnational transactions. The variable *No_application* is also a proxy for transaction costs since the use of ICTs may reduce communication and ink costs between commercial partners. The *Half_day_arrival* variable could approximate an increase in the fixed costs of dispatching a shipment: Due to the asked time pressure for the shipper, it is impossible to carry several shipments together, as opposed to the *Multi_trips* variable (which may additionally reduce some of the fixed part of driver wages or docking expenditures).

Table 2: Summary statistics of the LCA variables (2,219 observations).

Variable	Definition	Mean	Std.dev.
Ship_interm_prod	1 if ship. activity is production of intermediary prod.	0.25	0.43
Ship_interm_sales	1 if ship. activity is wholesale of intermediary prod.	0.16	0.37
Ship_manuf_prod	1 if ship. activity is production of manufactured prod.	0.13	0.33
Ship_manuf_sales	1 if ship. activity is wholesale of manufactured prod.	0.02	0.15
Ship_food_prod	1 if ship. activity is production of food prod.	0.12	0.33
Ship_food_sales	1 if ship. activity is wholesale of food prod.	0.09	0.28
Ship_consump_prod	1 if ship. activity is production of consumption prod.	0.18	0.38
Ship_consump_sales	1 if ship. activity is wholesale of consumption prod.	0.03	0.16
Ship_warehouse	1 if ship. activity is warehousing	0.02	0.15
Bulk	1 if commodity is transported as unpacked bulk	0.18	0.39
Bags	1 if commodity is transported in bags	0.18	0.38
Pallets	1 if commodity is transported on pallets	0.58	0.49
Dangerous	1 if commodity is inflamm., explosive, poisonous, etc.	0.03	0.18
Temperature	1 if commodity needs temperature-controlled handling	0.08	0.28
Fragile	1 if commodity is fragile	0.11	0.32
Voluminous	1 if commodity is voluminous	0.07	0.26
Great_dim	1 if commodity is oversized	0.03	0.17
Unconstrained	1 if commodity is unconstrained	0.70	0.46

Note: For the shipper's activity variables exact one variable takes the value 1. All other variables can be arbitrary either 0 or 1.

Table 2 presents the summary statistics of the logistical requirements used for the LCA. Each shipper belongs to exactly one among the nine activity groups.[7] The other variables are related to the shipments. About 70% of the shipments do not have any particular handling constraint. Besides, 60% of the goods are packaged onto pallets. As will be clear, these two variables play a major role.

Lastly, Table 3 shows the distribution of the shipment size categories used for the multinomial logit. The discretization is similar to the one made by Piendl et al. (2017) relying on the existing transport markets "‘general cargo"’ for shipments smaller than 3 tons, "‘partial loads"’ for shipments between 3 and 12 tons and "‘full truck loads"’ for shipments bigger than 12 tons. In contrast to that there are many shipments of less than one ton in the French

[7]These activity groups are based on the main activity of the shippers, classified along the Statistical Classification of Economic Activities in the European Community, or NACE rev.1 (Eurostat, 2015). On the basis of this classification, shippers were grouped into the nine categories shown in Table 2 (Guilbault and Soppé, 2009)

Table 3: Shipment size categories (2,219 observations).

Shipment size category	Frequency	Percentage	Cumulative
q_1: < 1t	890	40.11	40.11
q_2: Up to 3t	368	16.58	56.69
q_3: 3t–12t	427	19.24	75.94
q_4: > 12t	534	24.06	100.00
Total	2,219	100.00	

dataset and thus 2 categories are distinguished for shipments below 3 tons. Shipment sizes are well spread in the sample, with a large number of small shipments, but no category with a disproportionately low number of observations.

5. Results

This section first presents the results from the LCA.[8] Then, we estimate a discrete choice model of shipment size which integrates the corresponding latent segments of goods. Finally, we discuss the robustness of our main findings. In particular, we apply a 10-fold cross-validation in order to check the accuracy of shipment size predictions.

5.1 Homogeneous segments

The optimal number of classes to be integrated into the shipment size choice model is in the first instance not predefined. When working with LCA, the usual practice is to estimate several models with an increasing number of classes. Those models are then compared on the basis of information criteria, which balance likelihood improvements with the number of parameters. Examples of such criteria are the Bayes Information Criteria (BIC; see Schwarz (1978)) and the Akaike Information Criteria (AIC; see Akaike (1998)).

Table 4: Information criteria of competing LCA models.

No. classes	Log-Likelihood	BIC	Adjusted BIC	AIC
2	-10,882.915	4,715.9859	4,604.7855	4,516.3175
3	-10,238.274	3,565.3901	3,397.001	3,263.0351
4	-9,806.1061	2,839.7405	2,614.1626	2,434.6989
5	-9,479.9048	2,326.0245	2,043.2579	1,818.2962
6	-9,285.6757	2,076.2528	1,736.2975	1,465.838
7	-9,172.5686	1,988.7252	1,591.5811	1,275.6237
8	-9,081.682	1,945.6388	1,491.3059	1,129.8506

Note: Only the results of 8 classes are illustrated as more classes did not lead to an improvement of the shipment size choice model.

[8] Estimates are executed by using the STATA (V13) software and its LCA plugin (Lanza et al., 2015).

Table 4 presents the estimation of several models, with the number of segments growing from 2 to 8. None of the information criteria stops decreasing as the number of classes increases. With the sole objective to maximize the likelihood, the number of segments should consequently be increased even more. In the context of this article, however, we have decided to work with 5 segments because the goal is to find the most suitable segmentation with respect to the shipment size modelling. Since both models are estimated independently, there is no joint function which is optimized and thus we refrain from purely sticking to the results of the information criteria.

Instead, when looking at goodness of fit statistics for the respective shipment size model in Table 5, it appears that there is no benefit to have more than 5 segments.[9]

Table 5: Comparison of goodness of fit statistics for discrete shipment size choice model with 4, 5, 6, 7 and 8 latent classes.

No. classes	Log-Likelihood	McFadden R^2	BIC
4	-1,816.83	0.3694	3,911.03
5	-1,784.96	0.3793	3,870.41
6	-1,781.89	0.3793	3,887.37
7	-1,773.53	0.3812	3,893.77
8	-1,771.29	0.3809	3,912.41

Comparing the adjusted McFadden R^2 (for models between 5 and 8 classes, less classes perform even worse), the values are more or less the same. In contrast, the 5 class segmentation is clearly preferable according to the BIC of the discrete shipment size models (3870.41 for 5 segments in contrast to 3887.37 for 6 segments, 3893.77 for 7 segments and still increasing for more segments). Even if one might conclude from the pure LCA classification that working with more than 5 segments could be relevant, Table 5 stipulates that shipment size choice models do not improve when doing so. Additionally, it can be seen in Section 5.4 that 5 segments perform best with respect to prediction accuracy.

Table 6 shows the results of the LCA with the preferred 5 segments in detail. For an easier interpretation of the latent classes, the conditional probabilities which are both higher than the other probabilities in the same class and from the same indicators across the other classes are highlighted in bold. In general, the high conditional response probabilities reveal a good model separation, with clearly defined segments. This is also confirmed by the high relative entropy of 0.987 (i.e. the weighted average of the assigned posterior probabilities) which originally ranges between 0 and 1, with larger values indicating better separation (Collins and Lanza, 2010). Based on these conditional probabilities, it is possible to label the 5 segments as follows:

[9]When using the goodness of fit statistics for the respective discrete shipment size choice models, it is implicitly assumed that the loss of degree of freedoms caused by the segmentation is not taken into account.

Table 6: LCA results for the preferred 5 classes.

	Class 1	Class 2	Class 3	Class 4	Class 5
γ_c	0.08255	0.17821	0.17066	0.12852	0.44006
	Conditional probability of answering "'Yes"'				
Ship_interm_prod	0.02748	0.29590	0.15748	0.26110	0.29789
Ship_interm_sales	0.00563	0.10805	**0.45763**	0.09891	0.11852
Ship_manuf_prod	0.00014	0.17480	0.10540	0.10676	0.14902
Ship_manuf_sales	0.00002	0.01280	0.03151	0.03952	0.02123
Ship_food_prod	**0.36600**	0.07779	0.15000	0.13473	0.08390
Ship_food_sales	**0.52957**	0.04472	0.00797	0.08847	0.05121
Ship_consump_prod	0.01656	0.24254	0.06665	0.23194	0.21474
Ship_consump_sales	0.00003	0.03033	0.01586	0.03856	0.02765
Ship_warehouse	0.05456	0.01307	0.00749	0.00002	0.03584
Bulk	0.03292	0.06479	**0.99882**	0.00013	0.00004
Bags	0.19685	0.19481	0.00009	**0.98786**	0.00004
Pallets	**0.72655**	**0.63442**	0.00031	0.00041	**0.92978**
Dangerous	0.00004	**0.13976**	0.05467	0.00004	0.00001
Temperature	**0.99793**	**0.00010**	0.00005	0.00006	0.00002
Fragile	0.10965	**0.57603**	0.00014	0.00009	0.00002
Voluminous	0.01645	**0.29314**	0.10305	0.00005	0.00002
Great_dim	0.00003	**0.13309**	0.03261	0.00002	0.00001
Unconstrained	**0.00076**	**0.00044**	**0.78218**	**0.9825**	**0.99992**
Observations: 2,219					
Parameter estimated: 89					
Entropy: 45.47140					
Relative Entropy: 0.9873					

Note: To foster the interpretation of the LCA outcome, the parameters which are characteristic for the respective class are marked bold. They are considered as characteristic items for the class with respect to high values in the class itself as well as across the classes.

- *Standard temperature-controlled food products* (8% of the total number of shipments): Shipments mainly from the food sector (production or wholesale), mostly conditioned onto pallets, transported and warehoused in temperature-controlled environments. Typical representatives are: meat, fruit and vegetables, dairy products.
- *Special transport* (18%): Shipments with a specific transport constraint, except for the need of a temperature-controlled environment. This includes: hazardous materials, voluminous or fragile shipments, etc. Typical representatives: fertilizers, furniture, wine bottles, drugs.
- *Bulk cargo* (17%): Mostly bulk-packed shipments, mainly associated with production or sale of intermediate products. Typical representatives: cereals, scrap metals, sand.
- *Miscellaneous standard cargo in bags* (13%): Shipments in bags or parcels, without any constraint on the transport. Typical representatives: leaflets, clothes.
- *Palletised standard cargo* (44%): Palletised shipments, without constraints. Typical representatives: parquet flooring, dishware, chocolate.

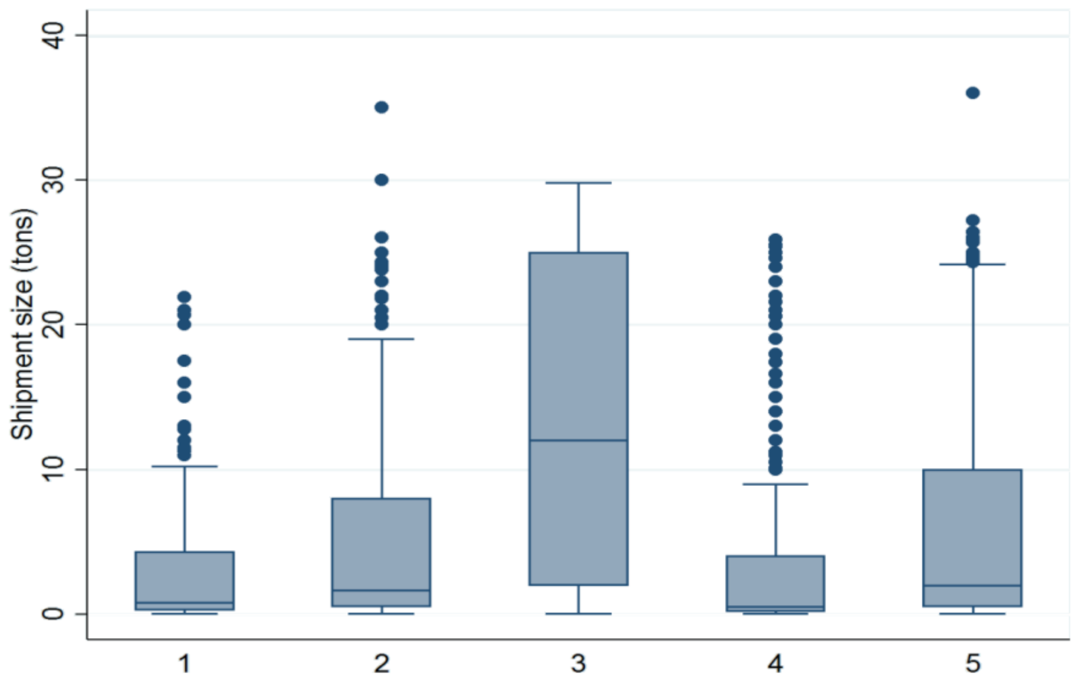
Several conclusions can be drawn from these results: First, the economic activity of the shipper is not particularly discriminatory[10], except for the food sector and, to a lesser extent, for intermediary products. By contrast, conditioning variables (bulk, pallets, bags and parcels) play a major role, in combination with the transport constraints (controlled temperature, hazardous materials, fragile or voluminous shipments, etc.). Second, this segmentation is clearly interpretable and shows a nice behaviour regarding the shipment sizes. Figure 2 proposes a boxplot of the distribution of the shipment sizes within each segment. The comparison with Figure 1 clearly highlights the improvements brought by LCA over the NST/R classification: With fewer segments, the heterogeneity of shipment sizes is much better captured.

In order to compare this segmentation with that obtained by Piendl et al. (2017), let remind the reader of the 4 segments found on German data (see also table 11 in the appendix):

- *(Temperature controlled) food products* (10% of the total number of shipments): These shipments do not have specific transport constraints and are often conditioned onto pallets.
- *Miscellaneous standard cargo* (36% of the total number of shipments): Mostly standard shipments and to a certain extent dangerous shipments.
- *Special goods* (35% of the total number of shipments): Shipments with specific constraints, such as being fragile, dangerous, or particularly valuable.
- *Unpacked bulk goods* (19% of the total number of shipments): Mostly bulk or liquid shipments, without other constraints.

[10]In addition to the structure of the LCA presented in this article, another LCA was realized with a distinct segmentation by economic activity. This gave more weight to economic activity in the identification of the segments, but performed worse in the shipment size model.

Figure 2: Boxplots of shipment sizes for derived segments.



These segments show remarkable similarities with the 5 segments identified on French data. In both cases, there is a food and temperature controlled segment with similar characteristics and size. In both cases, there is also a bulk segment, similar in characteristics as well as in size. In both cases, there is a special goods segment grouping shipments with a variety of handling constraints. Those segments differ in size though, and do not have exactly the same parameter. In particular, dangerous shipments belong to the standard shipments category in the German dataset, whereas in the French data there are two standard shipment groups, with a distinction between palletised shipments on the one hand and bags and parcels shipments on the other hand. Also, the prevalence of these last segments are not directly comparable: The German special goods segment has a much larger importance (35%) than the French one (18%). This can be explained by the difference in the nature of shipments between the two datasets: Shipments from the mechanical engineering and manufactured goods sectors are more frequent in the German dataset than in the ECHO survey.

Neglecting these minor differences, the two segmentations are remarkably similar, especially since the two datasets have been collected in different years, different countries and using a different survey design and method. It is already possible to conclude from these first findings that the LCA yields robust results as well as transferable results and that handling constraints and conditioning should be given the adequate attention in freight transport statistics.

5.2 Discrete shipment size model

We now present the estimation of the multinomial logit model investigating the determinants of shipment sizes. The smallest size category in Table 3 is taken as reference indicating that all coefficients are set to 0. Every shipment of the dataset is assigned exclusively to the latent class with maximum probability calculated according to Eq. 6. The impacts of the latent segments are estimated with "'*palletised standard cargo*"' segment as reference value. We first discuss the results without any reference to the segments derived from the LCA.

The variables derived from the TLC expression are significant and the signs of the coefficients are consistent with the theory. Taking the natural logarithm of the respective variables increases the goodness of fit significantly. A similar observation was made by Piendl et al. (2017) who explain it by the fact that the shipment size categories have distinct widths and by the damping effect of the natural logarithm. Given the positive and increasing coefficients of the haulage distance, the variable transport costs have a monotonically increasing and concave slope. This is in line with several empirical analyses (Liedtke, 2012; McCann, 2001). Possible explanations for this effect are the less than proportional fuel/time costs for larger vehicles (Abate and de Jong, 2014), a higher unreliability of transports with bigger haulage distance (implying higher inventories) (Piendl et al., 2017) or the decoupling of production locations and of regional retail centres by regional distribution centres (Combes, 2009). The two variables related to the storage costs have negative effects. Consistently with the EOQ model, an increase in the shipper-receiver commodity flow rate has a positive impact on shipment size. Conversely, an increase in the value density of goods has a negative impact on shipment size. Importantly, the *Vol_ton* variable also reveals a negative influence on the shipment size: A higher cubic density, theoretically associated with higher warehousing costs, is empirically associated with smaller shipments.

Looking then at the dummies used to control for the fixed costs of ordering and/or of transport, the *Dest_foreign* variable positively affects shipment size, revealing the significant influence of distinct cultures, trade barriers or the higher degree of administrative work going along with transnational transports. Also, the increased uncertainty linked to longer distances could explain higher shipment sizes, protecting the firms from stock shortages. Except for shipment size class q_2: 1t – 3t, we notice that the shipment size increases if the shippers do not use any software application to monitor the commercial process. This can be explained by an increase of the transaction costs, which raises the fixed costs of dispatching a shipment. Regarding the *Half_day_arrival* variable, it is associated with larger shipment sizes. The resulting lack of flexibility comes at a cost for carriers, in a way which can imply an increase in the fixed costs of dispatching a shipment (the need to take margins to arrive on time, the greater costs of being early or late, the lack of possibility to carry more shipments together). Lastly, transshipments are associated with the consolidation of small shipments and cross-docking operations, in order to reduce the fixed costs of transporting shipments (Combes and Tavasszy, 2016). As a consequence, the *Multi_trips* variable should be associated with smaller shipments, which is empirically confirmed.

Now questioning the effects of the latent segments derived from the LCA, it is worth noting that results presented in Table 7 have been estimated in reference to "'*palletised standard cargo*"'. Shipments assigned to the "'*standard temperature-controlled food products*"' class

Table 7: Discrete shipment size choice model with latent classes.

	1t – 3t	3t –12t	> 12t
Constant	-1.4562^{***} (0.3802)	-1.3622^{***} (0.4212)	-5.0282^{***} (0.6913)
Dest_foreign	0.5395^{**} (0.2466)	1.1334^{***} (0.2809)	1.3836^{***} (0.3338)
No_application	-0.0116^{ns} (0.1497)	0.3358^{**} (0.1712)	0.6774^{***} (0.2064)
Half_day_arrival	0.2763^{*} (0.1462)	0.4106^{**} (0.1658)	0.7121^{***} (0.2098)
Multi_trips	-0.7058^{***} (0.1511)	-1.5019^{***} (0.1902)	-1.9025^{***} (0.2474)
$\ln(d_n)$	0.1406^{***} (0.0443)	0.1880^{***} (0.0446)	0.4671^{***} (0.1011)
$\ln(vol\_ton_n/Q_n)$	-0.3885^{***} (0.0550)	-0.5264^{***} (0.0686)	-0.9003^{***} (0.0870)
$\ln(v_n/Q_n)$	-0.1042^{**} (0.0501)	-0.3631^{***} (0.0621)	-0.5447^{***} (0.0784)
Temp_food	-1.2105^{***} (0.2625)	-1.6556^{***} (0.2755)	-2.8567^{***} (0.3667)
Special_cargo	0.2436^{ns} (0.1938)	0.4276^{*} (0.2198)	0.6326^{**} (0.2855)
Unpacked_bulk_cargo	0.8525^{***} (0.2286)	0.8498^{***} (0.2685)	1.5211^{***} (0.2988)
Miscellaneous_cargo_bags	-0.8262^{***} (0.2153)	-1.4679^{***} (0.2625)	-1.0095^{***} (0.3107)

No. observations: 2,219
Final Log-Likelihood: -1,794.044
McFadden R^2 (ρ^2): 0.3895
Likelihood ratio test: χ^2 = 586.19

Note: Robust standard errors in parentheses. Smallest shipment size category q_1: < 1t is the base category. Significance levels: * p < 0.1; ** p < 0.05; *** p < 0.01; ^{ns} non-significant.

tend to be smaller, all other things equal. This can be explained by the higher transport costs, warehousing costs, and with the perishability of these commodities. By contrast, "‘*special transports*"’ tend to have bigger shipment sizes. One possible explanation is the increased transport costs associated with voluminous, fragile and dangerous products. Another possible explanation is the indivisible size of some shipments, something which is not accounted by the EOQ model. Quite intuitively, the "‘*unpacked bulk cargo*"’ has the highest positive effect on the shipment size choice. Low warehousing costs and low perishability might be one possible explanation. Also, consolidation is not easy for goods moved as bulk (it is for example not possible to transport simultaneously two kinds of wheat, or milk, in the same truck and then separate them, except if it has been specifically designed to do so.) Additionally, the segment has a proportion of 60% on intermediary products which are often used in the process industry. Running out of stock in the process industry goes along with enormously high costs and therefore a higher need for safety is required. Finally, the "‘*miscellaneous standard cargo in bags*"’ corresponds to shipments conditioned either in bags or in parcels. This conditioning generally goes with smaller shipment sizes: The estimation's results are intuitively correct.

Table 8: Comparison of goodness of fit statistics for discrete shipment size choice models with no segmentation, administrative segmentation and 5 latent segments.

Model	No. parameters	Log-Likelihood	McFadden R^2	BIC
No segmentation	27	-1,863.01	0.3660	3,934.05
Administrative segmentation	48	-1,815.03	0.3824	3,999.89
Latent class segmentation	39	-1,784.96	0.3926	3,870.41

In order to stress the relevance of using segments derived from LCA, let us highlight that the same estimates, but without the 5 latent classes, yield results with a significantly lower explanatory power (even if they are quite similar regarding the effects of other explanatory variables). Table 8 shows the key performance indicators of shipment size choice models with no segmentation, administrative segmentation and latent class segmentation with preferred 5 segments. The latter model is superior to the other models with respect Log-Likelihood, McFadden R^2 and the Bayesian Information Criterion (BIC). As confirmed by a Likelihood-Ratio (LR) test, the 80 units increase in the log-likelihood suggests that the model with the homogeneous segments (in LR-tests referred to as unconstrained model) should be favoured over the model without latent segments (in LR-tests referred to as constrained model). Above all, considering the NST-R classification rather than the latent segments leads to less precise estimates. Even if no formal test can be used to contrast non-nested models, Table 13 in the appendix shows that the log-likelihood is 30 units lower with NST-R, despite a larger number of explanatory variables. Recognizing that Piendl et al. (2017) reach the same conclusions, the signs and the magnitude of the coefficients estimated for the segments correspond to a great extent to those found on German data, as presented in table 12 of the appendix. These econometric results support the transferability of the LCA approach as a way to reduce the heterogeneity in shipment size modelling.

5.3 Robustness tests

The authors have run various econometric tests (not presented here but available upon request) in order to question the robustness of our main empirical findings.

In line with other authors (e.g. Abate and de Jong (2014); Combes (2009); Holguín-Veras (2002)), we also treat shipment size as a continuous variable. Doing so, we find very similar results: All coefficients are significant, and their signs are consistent with those shown in Table 7, especially for the volume density whose consideration represents a novelty with respect to the previous literature. When adding the segments derived from LCA, it appears that their effects are significant in comparison to the "'*miscellaneous standard cargo on pallets*"' category. All other things equal, shipments belonging to the segment "'*special cargo*"' or to "'*unpacked bulk cargo*"' are significantly larger, but not significantly different from each other. The "'*miscellaneous standard cargo in bags*"' shipments are smaller, and the "'*standard temperature-controlled food products*"' even more. Again, all these results are consistent with those in Table 7, even if they provide a bit less information due to the smaller amount of degrees of freedom. Moreover, the R^2 increases substantially (from 0.71 to 0.73) when adding these latent segments into the vector of explanatory variables. Also, it is confirmed that introducing the NST classes does not yield the same improvements in precision of the estimates.

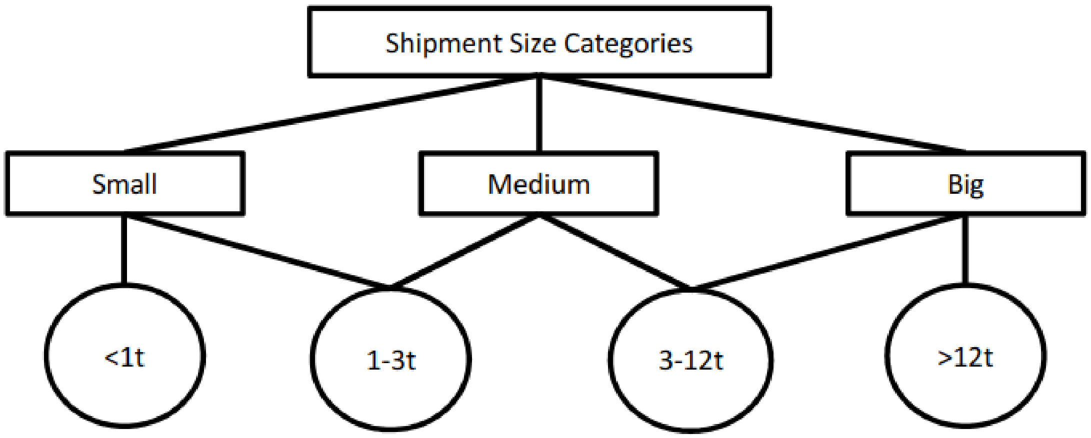
Additionally, a stark restriction of the shipment size choice model is the assumption of independence of irrelevant alternatives (IIA). This assumption implies an inflexible substitution patterns between shipment size categories although they are in a natural order. Thus, we checked first if the IIA assumption is violated for the model with 5 latent classes by applying the Hausman-McFadden test and the Small-Hsiao test. The former test reveals a dependency between the smallest two shipment size categories whereas the latter one shows none of the shipment size categories to be dependent on a 5% significance level. As stated by Long and Freese (2014), the tests often produce unreliable results requiring an additional robustness check. In order to do so, we tried different mixed logit formulations enabling flexible substitution patterns between the shipment size categories.

First, we estimated a model where the alternative-specific constants have a normal mixing distribution which coincides with a flexible substitution of the alternatives. After normalizing the alternative with the lowest variance to zero (see Walker et al. (2007)), an improvement in model fit caused by a mixed logit formulation can't be observed.

As the approach of applying mixing distributions to the alternative specific constants doesn't provide convincing results, we additionally applied a mixed logit formulation to build different nests of shipment size categories and thus also allow for a flexible substitution across the alternatives. First, we build two nests for the smaller shipment sizes ($<$ 1t, 1 – 3t) and the bigger shipment sizes (3 – 12t, $>$ 12t) in order to account for diverging substitution patterns. After normalizing arbitrarily error variances of one nest, the remaining variance shows up insignificant without an increase of the final log-likelihood value (final log-likelihood increases by 0.002).

Furthermore, we applied a cross-nested logit approach as shown in Figure 3. Within that model formulation, it is assumed that the medium shipment size categories (1 – 3t, 3 – 12t)

Figure 3: Structure of the cross-nested logit model.



are allowed to belong to a nest which is related to their lower or their upper bound and thus have a more flexible substitution pattern. After checking the order and rank conditions, only two variances in this mixed logit model are identifiable. After normalizing the smallest of the three variances, the estimated model again doesn't show an improvement in its final log-likelihood. In summary, the flexible modelling of shipment size choice by applying mixed logit models doesn't improve the model fit which in our view supports the usage of a standard multinomial logit model. This behaviour of the choice model might be caused by the explanatory variables — especially the annual flow variable -- which accounts for a big portion of the heterogeneity of firms and shipments and thus further correlations which would violate the IIA assumption are not observable.[11]

5.4 Predicting shipment sizes

Once one considers the performance of a model, two dimensions are crucial. Firstly, the ability of explaining the actual observations, which is usually measured by the final log-likelihood and its related indicators, is considered. Secondly, it is important to know how the model performs if it is applied to data not used for model estimation. The latter one deals with the external validity of our empirical strategy regarding the accuracy of shipment size predictions. To do so, we here follow a k-fold cross-validation approach: First, we randomly divide the initial dataset into $k = 10$ stratified (keeping the share of categories of the original data set in each subset) subsamples of equal sizes. Assuming that the segmentation is fixed due to its exogenous character with respect to the shipment size

[11] In addition to the checks presented in this section we additionally conducted rather informal tests addressing the drop of a substantial amount of observations from the initial ECHO survey caused by our sampling strategy (see Section 4). Although no anomalies were detected by increasing the sample in various ways, it is pointed to the fact that a potential bias still might exist. Further let us emphasize that several interaction terms have been introduced in order to control for a possible influence of the "'*Dest_foreign*"' variable on the balance between transport costs and volume/value of the shipment. Since all of these interaction terms are not statistically different from zero, the model presented in Table 4 is the most relevant: It allows controlling for the non-transport costs linked to international trade.

choice model, we will only evaluate the changes of prediction accuracy for the shipment size choice models conditioned on a predefined segmentation.

Within that framework, $k - 1 = 9$ subsamples are used for estimating shipment size determinants and the omitted subsample serves in a second step as validity check for prediction performance. In general, each observation of the data is used in total $k - 1$ times for estimation and one time for prediction. In order to ensure a robust model assessment, the presented k-fold cross validation approach is repeated solely 10 times (10 times stratification with subsequent altering model estimations and applications) as the models and their outputs are on average stable and thus additional runs don't improve the analysis. In general, the results are first averaged according to the folds and then according to the repetitions of the approach.

Table 9: Overall accuracy of shipment size choice models.

	Internal accuracy in % (training data)	External accuracy in % (test data)
No Segmentation	63.69	63.11
NST classification	65.00	63.76
4 latent segments	64.95	64.07
5 latent segments	65.70	64.69
6 latent segments	65.30	64.22
7 latent segments	65.12	63.96
8 latent segments	65.49	64.43

Table 9 presents the average accuracy for different segmentation approaches which is measured as share of correctly predicted shipment size categories in the total amount of observations used for prediction. As one can see, the overall accuracy of the models is good but not overwhelming using the maximum probability as classification criterion. A maximum of 64.69% correctly predicted external (not used for model estimation) shipments to the respective shipment size categories is reached for 5 latent segment. As a reference: the most trivial model with highest average accuracy – every shipment belongs to the smallest shipment size category – has an overall accuracy of 40.11%. Additionally, the developed models don't suffer from overfitting as the accuracy of the test data is almost as high as the accuracy of the training data.

In general, one can state that the segmentation approaches have only a small effect on the prediction accuracy. Nonetheless, the segmentation with five latent classes provides the best external fit although it has less estimated parameters in the models than the NST classification for example.

Furthermore, Table 10 provides the prediction accuracy of the 5 class model differentiated by shipment size categories. The accuracy values in this context are also known as sensitivity or recall of a classifier which are defined as share of correctly classified shipments conditioned on a specific shipment size category.

Obviously, the shipment size choice model has problems correctly predicting the second

Table 10: Correctly predicted test shipments conditioned on shipment size category for the model with 5 segments.

	Accuracy in % q_1: < 1t	Accuracy in % q_2: 1t – 3t	Accuracy in % q_3: 3t – 12t	Accuracy in % q_4: > 12t	Accuracy in % Total
5 segments	89.05	11.15	42.19	78.97	64.69

shipment size category and with some limitations also the third shipment size category. Nonetheless, especially the first and the fourth category reveal high accuracy. Such a behaviour doesn't seem to be unusual once one considers imbalanced data or multiple outcomes (e.g. Wang and Ross (2018); Samimi et al. (2017)). A possible strategy to increase the prediction accuracy could be to combine the first and the second shipment size to a single category but this would lead to a higher data imbalance and thus potentially negatively affect prediction accuracy of the third or the fourth shipment size category.

6. Summary and further research

The general objective of this paper was to assess the robustness of the empirical strategy proposed by Piendl et al. (2017) for Germany by using the French commodity flow survey ECHO. Our results indicate that the combination of homogeneous segments derived from LCA and a rational core of shipment size model based on a Total Logistics Costs (TLC) function is a promising approach to address the prevalent heterogeneity in freight activities, to establish a behaviour-sensitive multi-step freight transport model and to ensure the aggregation in a comprehensive freight transport demand model framework.

6.1 Results of the empirical analysis

The dataset used in this paper includes 2,219 shipments transported in France exclusively by road, of weight over 30kg, and of size larger than 0.5m^3. The variables used for estimations come from the Economic Order Quantity (EOQ) microeconomic model, and they additionnally describe a set of logistics requirements. The econometric estimations have reached our initial objectives.

In general, the results obtained with the French data are consistent with those found for Germany. In particular, the segments obtained by LCA are both consistent with Piendl et al. (2017) and significantly improve the explanatory power of the shipment size model. These segments also show intuitive impacts on the chosen shipment size. This goes a long way in demonstrating the value of the proposed classification approach. and is especially true as Germany and France are responsible for about 25% (in tkm, year: 2017) of total European road freight transport (Source: Eurostat). This also indicates that useful data for freight transport modelling should not be limited to administrative commodity types, as most current national and international transport surveys are based upon. As a matter of fact, our estimations show that segments built upon information about handling constraints and conditioning bring more explanatory power than standardised NST classes.

Beside the robustness of the goods segments, the empirical models are consistent with the theoretical EOQ model. It should be noted that the introduction of the volume density variable as a proxy of warehousing costs (commodities occupying more space per ton are more expensive to store) is empirically valid and improves the understanding of shipment size variation. In addition, the large amount of information available in the ECHO dataset made it possible to decompose the fixed costs of transport and/or ordering into transactional variables. These different findings are stable for both discrete choice and continuous models; and they are confirmed by the various robustness tests presented in this article. Additionally, the predictive power has been assessed by a cross-validation approach revealing acceptable performance of the shipment size choice models and somehow restricted effects of the segmentations. This is not surprising as they enter the models by categorical variables often having restricted explanatory power. Nonetheless, we showed that the NST classification was outperformed in almost all cases (except the internal accuracy for 4 latent segments) although it uses more parameters.

6.2 Applicability and restrictions of the modelling approach

The scope of the model presented in this article is given by a better behaviour-sensitive prediction of shipment sizes for truck transports. Thus, the study is focused on the examination of direct shipments vs. hub & spoke networks for long distance road transports, rather than the relationship between shipment size choice and mode choice. Based on this model, the assessment of possible future developments are imaginable such as the market penetration of longer semi-trailers or trucks with extraordinary high volumes, advanced multimodal concepts or other innovative logistical initiatives.

Nonetheless, an enhancement of our empirical approach to other modes still has to be done. Having such a methodological framework would make policy analysis such as given by Kleist and Doll (2005) also realistic with respect to logistical adaptions and thus enable deeper understanding of stakeholder reactions. Adjustments on the ECHO dataset, re-estimations of the LCA and the application of an econometric framework for a joint decision on shipment size and mode may be viewed as prerequisites. Concerning the latter objective, multiple empirical approaches do exist in the literature. First, if the shipment size is treated continuously, the estimation of discrete-continuous choice models becomes relevant (e.g. Abate and de Jong (2014)). Treating the shipment size as discrete, the combination might be expressed by a simple multinomial logit model for shipment size choice and mode choice or by (cross) nested logit models (e.g. Stinson et al. (2017); de Jong and Ben-Akiva (2007)). As the flexibility of the dependency structures might be restricted in the named approaches, it may be relevant to employ copulas techniques in this respect (e.g. Irannezhad et al. (2017); Pourabdollahi et al. (2013)).

6.3 Wider implications of the transferability analysis

The transferability analysis provided in this article supports the generalization of the estimated segments and its impacts on the shipment size choice. Although using slightly different attributes for the segmentation approach, similar results have been obtained for France and Germany. Even if the analysis has been restricted to only two countries, which calls for further research in other places (e.g. Sweden or the US with their comprehensive commodity flow surveys), theoretical statements about its wider applicability are possible. When comparing surveys on a global level, the attributes referring to the commodities are predominantly describing similar characteristics. Finding no extraordinary differences regarding these attributes might consequently support the applicability of the proposed goods segments, even in other countries than Germany and France. Furthermore, the attributes describing commodities are more or less standardised, so that they potentially can be integrated into shipment related surveys.

An open question is given by the prospective usage of the estimated latent segments. Several distinct approaches are possible in this case. On the one hand, one might use the estimated segments in shipment surveys so that respondents allocate the shipments directly to one group. Since such an approach requires a hard allocation rule and could be biased when commodities are suitable for multiple segments, the implementation is at least questionable. Another possibility might be given by the application of fuzzy allocation methods based on (even just some of) the logistics attributes. More precisely, this would be comparable to the so-called "‘Sinus-Milieus"’ concept, a typology of societal target groups which are similar in their lifestyle (Edwards and Calmbach, 2020). Finally, if there is no information about shipments attributes, the usability of the segments is limited. Nevertheless, if surveys contain information on the standardised classifications, one might apply a probabilistic assignment approach which accounts for the relation between standard classifications and latent segments, as proposed by Piendl et al. (2019). If there is no information about correspondences at all, simple probabilistic assignments according to the estimated segment proportions are an option. In the two latter cases, however, the volatility of the modelling outcomes would obviously become a major concern.

Acknowledgements

The authors are thankful are various participants of the ICPLT – Interdisciplinary Conference on Production, Logistics and Traffic for their valuable comments to the presentation. This article emerged within the French-German virtual research institute DISTRANS, which enables bi-national collaborations in the research area of distribution, logistics & transport between IFSTTAR-SPLOTT (France) and the German Aerospace Center (DLR).

References

Abate, M. and de Jong, G. (2014). The optimal shipment size and truck size choice - the allocation of trucks across hauls. *Transportation Research Part A: Policy and Practice*, 59:262–277.

Abate, M. A., Vierth, I., Karlsson, R., de Jong, G., and Baak, J. (2019). A disaggregate stochastic freight transport model for sweden. *Transportation*, 46:671–696.

Abdelwahab, W. and Sargious, M. (1992). Modelling the demand for freight transport: A new approach. *Journal of Transport Economics and Policy*, 26(1):49–70.

Akaike, H. (1998). *Information Theory and an Extension of the Maximum Likelihood Principle*, pages 199–213. Springer New York, New York, NY.

Arunotayanun, K. and Polak, J. W. (2011). Taste heterogeneity and market segmentation in freight shippers' mode choice behaviour. *Transportation Research Part E: Logistics and Transportation Review*, 47(2):138–148.

Baumol, W. J. and Vinod, H. D. (1970). An inventory theoretic model of freight transport demand. *Management Science*, 16(7):413–421.

Collins, L. M. and Lanza, S. T. (2010). *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences*, volume 718 of *Wiley Series in Probability and Statistics*. John Wiley & Sons.

Combes, F. (2009). *The choice of shipment size in freight transport*. PhD thesis, Université Paris-Est.

Combes, F. (2012). Empirical evaluation of economic order quantity model for choice of shipment size in freight transport. *Transportation Research Record: Journal of the Transportation Research Board*, 2269:92–98.

Combes, F. and Tavasszy, L. A. (2016). Inventory theory, mode choice and network structure in freight transport. *European Journal of Transport & Infrastructure Research*, 16(1).

de Jong, G. and Ben-Akiva, M. (2007). A micro-simulation model of shipment size and transport chain choice. *Transportation Research Part B: Methodological*, 41(9):950–965. Behavioural insights into the Modelling of Freight Transportation and Distribution Systems.

de Jong, G., Ben-Akiva, M., and Baak, J. (2010). Method report - logistics model in the swedish national freight model system (version 2). Technical report.

de Jong, G., Ben-Akiva, M., Baak, J., and Grønland, S. E. (2013). Method report - logistics model in the norwegian national freight model system (version 3). Technical report.

Edwards, J. and Calmbach, M. (2020). More signal, less noise: Targeted content marketing with sinus-milieus. In *Content gekonnt*, pages 91–111. Springer.

European Parliament and Council (2012). Regulation (eu) no 70/2012 of the european parliament and of the council of 18 january 2012 on statistical returns in respect of the carriage of goods by road. *Official Journal of the European Union*.

Greene, W. H. and Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit. *Transportation Research Part B: Methodological*, 37(8):681–698.

Guilbault, M. and Soppé, M. (2009). *Grandes tendances d'évolution des pratiques de transport et logistique au travers des enquêtes chargeurs 1988 et ECHO 2004*, volume 121, pages 33–56. Les collections de l'INRETS, actes inrets edition.

Hall, R. W. (1985). Dependence between shipment size and mode in freight transportation. *Transportation Science*, 19(4):436–444.

Harris, F. W. (1913). How many parts to make at once. *Factory, The Magazine of Management*, 10(2):135–136.

Holguín-Veras, J. (2002). Revealed preference analysis of commercial vehicle choice process. *Journal of Transportation Engineering*, 128(4):336–346.

Irannezhad, E., Prato, C. G., Hickman, M., and Mohaymany, A. S. (2017). Copula-based joint discrete-continuous model of road vehicle type and shipment size. *Transportation Research Record: Journal of the Transportation Research Board*, 2610:87–96.

Islam, D. M. Z., Meier, J. F., Aditjandra, P. T., Zunder, T. H., and Pace, G. (2013). Logistics and supply chain management. *Research in Transportation Economics*, 41(1):3–16. An Intensive Programme in Railway and Logistics.

Kawamura, K., Mohammadian, A. K., Samimi, A., Pourabdollahi, Z., and Klekotka, M. (2010). Development of freight policy analysis tool for northeastern illinois and the united states. Technical report.

Keya, N., Anowar, S., and Eluru, N. (2019). Joint model of freight mode choice and shipment size: A copula-based random regret minimization framework. *Transportation Research Part E: Logistics and Transportation Review*, 125:97–115.

Kleist, L. and Doll, C. (2005). 9. economic and environmental impacts of road tolls for hgvs in europe. *Research in Transportation Economics*, 11:153–192. Interurban road charging for trucks in Europe.

Koning, M., Combes, F., Piendl, R., and Liedtke, G. (2018). Transferability of models for logistics behaviors: A cross-country comparison between france and germany for shipment size choice. *RTS - Recherche Transports Sécurité*.

Lanza, S. T., Dziak, J. J., Huang, L., Wagner, A. T., and Collins, L. M. (2015). Lca stata plugin users' guide (version 1.2). Technical report.

Lanza, S. T., Tan, X., and Bray, B. C. (2013). Latent class analysis with distal outcomes: A flexible model-based approach. *Structural equation modeling: a multidisciplinary journal*, 20(1):1–26.

Liedtke, G. (2012). Estimation of the benefits for shippers from a multimodal transport network. *Logistics Research*, 4(3-4):113–125.

Liedtke, G. and Schepperle, H. (2004). Segmentation of the transportation market with regard to activity-based freight transport modelling. *International Journal of Logistics Research and Applications*, 7(3):199–218.

Long, J. S. and Freese, J. (2014). *Regression Models for Categorical Dependent Variables Using Stata*. Stata press, 3 edition.

McCann, P. (2001). A proof of the relationship between optimal vehicle size, haulage length and the structure of distance-transport costs. *Transportation Research Part A: Policy and Practice*, 35(8):671–693.

McFadden, D. and Train, K. E. (2000). Mixed mnl models for discrete response. *Journal of applied Econometrics*, 15(5):447–470.

Nagin, D. (2005). *Group-based modeling of development*. Harvard University Press.

Piendl, R., Liedtke, G., and Matteis, T. (2017). A logit model for shipment size choice with latent classes – empirical findings for germany. *Transportation Research Part A: Policy and Practice*, 102:188–201.

Piendl, R., Matteis, T., and Liedtke, G. (2019). A machine learning approach for the operationalization of latent classes in a discrete shipment size choice model. *Transportation Research Part E: Logistics and Transportation Review*, 121:149–161.

Pourabdollahi, Z., Karimi, B., and Mohammadian, A. (2013). Joint model of freight mode and shipment size choice. *Transportation Research Record: Journal of the Transportation Research Board*, 2378(1):84–91.

Sakai, T., Alho, A., Hyodo, T., and Ben-Akiva, M. (2020). Empirical shipment size model for urban freight and its implications. *Transportation Research Record*, 2674(5):12–21.

Samimi, A., Mohammadian, A., Kawamura, K., and Pourabdollahi, Z. (2014). An activity-based freight mode choice microsimulation model. *Transportation Letters*, 6(3):142–151.

Samimi, A., Razi-Ardakani, H., and Nohekhan, A. (2017). A comparison between different data mining algorithms in freight mode choice. *American Journal of Applied Sciences*, 14(2):204–216.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464.

Shelat, S., Huisman, R., and van Oort, N. (2018). Analysing the trip and user characteristics of the combined bicycle and transit mode. *Research in Transportation Economics*, 69:68–76. Competition and Ownership in Land Passenger Transport (selected papers from the Thredbo 15 conference).

Stinson, M., Pourabdollahi, Z., Livshits, V., Kyunghwi, J., Nippani, S., and Zhu, H. (2017). A joint model of mode and shipment size choice using the first generation of commodity flow survey public use microdata. *International Journal of Transportation Science and Technology*, 6(4):330–343. Special Issue on Urban Spatiotemporal Behavior and Network Assignment.

Tavasszy, L. A. (2000). On the demand-side segmentation of freight transport markets with logistical families. Think-up project deliverable, European Commission, Brussels.

Tavasszy, L. A. and de Jong, G. (2014). *Modelling freight transport.* Elsevier.

Tavasszy, L. A., Smeenk, B., and Ruijgrok, C. J. (1998). A dss for modelling logistic chains in freight transport policy analysis. *International Transactions in Operational Research*, 5(6):447–459.

Train, K. E. (2008). EM algorithms for nonparametric estimation of mixing distributions. *Journal of Choice Modelling*, 1(1):40–69.

Walker, J. L., Ben-Akiva, M., and Bolduc, D. (2007). Identification of parameters in normal error component logit-mixture (neclm) models. *Journal of Applied Econometrics*, 22(6):1095–1125.

Wang, F. and Ross, C. L. (2018). Machine learning travel mode choices: Comparing the performance of an extreme gradient boosting model with a multinomial logit model. *Transportation Research Record: Journal of the Transportation Research Board.* In Press.

Windisch, E., de Jong, G., Van Nes, R., and Hoogendoorn, S. P. (2010). A disaggregate freight transport model of transport chain and shipment size choice. In *Proceedings of the European Transport Conference.*

Wisetjindawat, W., Sano, K., and Matsumoto, S. (2005). Supply chain simulation for modeling the interactions in freight movement. *Journal of the Eastern Asia Society for Transportation Studies*, 6:2991–3004.

A. Appendix

The appendix consists of two different components. First, the modelling results of Piendl et al. (2017), which are given by the segmentation model and the multinomial logit model, are presented for the sake of comparison.

Table 11: Results of the LCA presented by Piendl et al. (2017).

	Class 1	Class 2	Class 3	Class 4
γ_c	0.1070	0.3600	0.3475	0.1854
	Conditional probability of answering "‘Yes"’			
Fragile	0.1727	0.0634	**0.3302**	**0.0000**
Valuable	0.2631	0.2395	**0.6373**	0.1266
Bulky	**0.0000**	0.2376	**0.4942**	0.2734
Temperature	**0.6331**	0.0855	0.0000	0.0000
Food	**1.0000**	0.0000	0.0000	0.0651
Dangerous	**0.0000**	**0.2079**	0.0599	0.1707
Bulk cargo	0.0398	0.0000	**0.0000**	**0.3647**
Liquid	0.0555	0.0000	**0.0000**	**0.1800**
Standard	**0.5550**	**0.8113**	0.2308	**0.0305**
Custom	0.0384	0.0841	**0.3974**	**0.0000**
Accumulation	0.4781	0.4134	0.3992	**0.0557**
Observations: 487				
Parameter estimated: 47				

Second, we present the results of the discrete shipment size model using the NST/R categories. The modelling framework remains the same as compared to the estimates shown in Table 7, solely the latent segments are replaced.

Table 12: Results of the discrete shipment size choice model presented by Piendl et al. (2017).

	q_2: 3t-12t	q_3:> 12t
Constant	-5.5253^{***} (1.3670)	-10.027^{***} (1.7849)
d_n	0.0013^{***} (0.0005)	0.0027^{***} (0.0006)
$\ln(1/Q_n)$	-0.8900^{***} (0.1933)	-1.4964^{***} (0.2451)
$\ln(v_n/Q_n)$	-0.2239^{*} (0.1175)	-0.7521^{***} (0.1516)
Miscellaneous standard cargo	0.9168^{*} (0.4884)	1.3603^{**} (0.5799)
Special goods	1.2750^{**} (0.5004)	1.9700^{***} (0.6004)
Unpacked bulk goods	1.1386^{**} (0.5776)	2.4094^{***} (0.6741)
Final log-likelihood: -321.15		
McFadden R^2 (ρ^2): 0.3711		
Adj. McFadden $\bar{R}^2$ ($\bar{\rho}^2$): 0.3482		
Likelihood ratio test: χ^2 = 379 (p value = <2.22e-16)		

Note: Significance levels: * p < 0.1; ** p < 0.05; *** p < 0.01; 487 observations; standard errors are given in brackets.

Table 13: Discrete shipment size choice model with NST/R categories.

	1t – 3t	3t -12t	> 12t
Constant	-2.1521^{***} (0.4189)	-3.1540^{***} (0.4891)	-6.0996^{***} (0.6773)
Dest_foreign	0.5713^{**} (0.2452)	1.1660^{***} (0.2824)	1.5488^{***} (0.3272)
No_application	0.0696^{ns} (0.1477)	0.4727^{***} (0.1732)	0.8286^{***} (0.2055)
Half_day_arrival	0.2979^{**} (0.1458)	0.3316^{**} (0.1654)	0.6540^{***} (0.2050)
Multi_trips	-0.7666^{***} (0.1494)	-1.5579^{***} (0.1882)	-2.0581^{***} (0.2405)
$\ln(d_n)$	0.1159^{**} (0.0449)	0.2005^{***} (0.0456)	0.4783^{***} (0.1028)
$\ln(vol\_ton_n/Q_n)$	-0.3535^{***} (0.0547)	-0.4727^{***} (0.0661)	-0.8003^{***} (0.0858)
$\ln(v_n/Q_n)$	-0.1150^{**} (0.0509)	-0.3854^{***} (0.0622)	-0.5918^{***} (0.0818)
Food	0.4051^{ns} (0.3221)	1.2447^{***} (0.3745)	1.2300^{***} (0.3736)
Ores	1.0859^{**} (0.4858)	0.5120^{ns} (0.8913)	1.8044^{***} (0.6347)
Petrol	2.0204^{***} (0.5892)	1.7529^{**} (0.7665)	2.1330^{***} (0.6023)
Metals	1.0265^{**} (0.4255)	1.7745^{***} (0.5158)	1.6571^{***} (0.5285)
Building	0.6530^{ns} (0.6033)	3.0932^{***} (0.5623)	2.8836^{***} (0.6132)
Chemical	0.9789^{***} (0.3630)	2.2018^{***} (0.4081)	1.4815^{***} (0.4378)
Manufactured	0.9749^{***} (0.3042)	2.0499^{***} (0.3663)	1.3169^{***} (0.3654)

No. observations: 2,219
Final Log-Likelihood: -1,822.3363
McFadden R^2 (ρ^2): 0.3799
Likelihood ratio test: χ^2 = 633.73

Note: Robust standard errors in parentheses. Smallest shipment size category q_1: < 1t is the base category. Impact of NST/R categories estimated with "'*Agri-food*"'-category as reference. Significance levels: * p < 0.1; ** p < 0.05; *** p < 0.01; ns non-significant.