

# **Exploring multiple-discreteness in freight transport. A Multiple Discrete Extreme Value Model application for grain consolidators in Argentina**

## **Rodrigo J. Tapia, Corresponding Author**

Laboratorio de Sistemas de Transportes  
Engenharia de Produção  
Universidade Federal do Rio Grande do Sul, Brazil,  
Av. Osvaldo Aranha, 99, Porto Alegre, 90035-190, Brazil  
Email: rodrigo.tapia@ufrgs.br

## **Gerard de Jong**

Economics and Appraisal group  
Institute for Transport Studies  
University of Leeds, UK  
36-40 University Rd, Leeds LS2 9JT  
Email: G.C.DeJong@its.leeds.ac.uk

## **Ana M. Larranaga**

Laboratorio de Sistemas de Transportes  
Engenharia de Produção  
Universidade Federal do Rio Grande do Sul, Brazil,  
Av. Osvaldo Aranha, 99, Porto Alegre, 90035-190, Brazil  
Email: analarra@producao.ufrgs.br

## **Helena B. Bettella Cybis**

Laboratorio de Sistemas de Transportes  
Engenharia de Produção  
Universidade Federal do Rio Grande do Sul, Brazil,  
Av. Osvaldo Aranha, 99, Porto Alegre, 90035-190, Brazil  
Email: helenabc@producao.ufrgs.br

**ABSTRACT**

There are some examples where freight choices may be of a multiple discrete nature, especially the ones at more tactical levels of planning. Nevertheless, this has not been investigated in the literature, although several discrete-continuous models for mode/vehicle type and shipment size choice have been developed in freight transport. In this work, we propose that the decision of port and mode of the grain consolidators in Argentina is of a discrete-continuous nature, where they can choose more than one alternative and how much of their production to send by each mode. The Multiple Discrete Extreme Value Model (MDCEV) framework was applied to a stated preference data set with a response variable that allowed this multiple-discreteness. To our knowledge, this is the only application of the MDCEV in regional freight context. Free alongside ship price, freight transport cost, lead-time and travel time were included in the utility function and observed and random heterogeneity was captured by the interaction with the consolidator's characteristics and random coefficients. In addition, different discrete choice models were used to compare the forecasting performance, willingness to pay measures and structure of the utility function against.

*Keywords:* Freight modelling, Multiple Discrete Extreme Value Model, Willingness to pay

## INTRODUCTION

Freight transport modelling is shifting from an aggregated perspective towards a behavioural and disaggregate one (Tavasszy and de Jong, 2014). This is happening not only because of new modelling techniques, but also because of the need to address more complex behaviour of the supply chain. The data constraints has been more severe in freight rather than passengers, resulting in the later adoption of a behavioural framework (Brooks and Trifts, 2008).

Except for a few very recent studies (for example (Khan and Machemehl, 2017a, 2017b)(Khan and Machemehl, 2017a)(Khan and Machemehl, 2017a)(Khan and Machemehl, 2017a) and Rashidi and Roorda (2018), all of them in urban context) choices in the freight context have been treated as mutually exclusive alternatives within the traditional discrete framework (Chow et al., 2010; Danielis and Marcucci, 2007; de Jong et al., 2013, 2000; De Jong et al., 2014; Rich et al., 2009; Vellay and de Jong, 2003). This carries the underlying assumption that all alternatives are extensive and mutually exclusive, meaning that choosing one fulfils the need that triggered the choice and they are perfect substitutes of each other.

Discrete-continuous models are used when a decision-maker has the possibility of choosing between multiple alternatives and the intensity of that choice. There have been some efforts to model these situations in the past, starting with applications in passenger transport or electricity demand (Dubin and McFadden, 1984; Hanemann, 1984; de Jong, 1990).

Some researchers have applied this discrete-continuous model framework in regional freight transport. Examples with the choice of mode or road vehicle type as the discrete dependent variable and shipment size as the continuous dependent variable are: Abdelwahab (1998), Abdelwahab and Sargious (1992), de Jong and Johnson (2009), Holguín-Veras (2002) and McFadden et al., (1986). All these models however, use a single continuous dependent variable and the idea is that only one mode or vehicle type is chosen in reality, that gets all the continuous units (so the alternatives are mutually exclusive). Moreover, these models are not or only loosely based on the micro-economic theory of utility (or profit) maximisation under a budget constraint. Also, the estimation methods used are mostly sequential (two-step): first one choice is estimated, and then in the second step the other choice conditional on the estimation results of the first. However, simultaneous estimation is statistically more efficient than two-step estimation.

Most choices in freight transport are mutually exclusive, but there are some where the mutual exclusiveness of the choices is broken. For instance, some companies may use dual sourcing for production (Yu et al., 2009), meaning that they would locate or purchase their production in more than one location and decide to produce/buy different volumes in each one. This would allow them to have a better mix of manufacturing flexibility, costs, risk management and speed of response. The discrete choice models and discrete-continuous models mentioned above do not cover such situations of multiple discreteness, unlike the Multiple Discrete Extreme Value model (MDCEV) (Bhat, 2008, 2005).

The MDCEV has brought an impulse to more complex decision making. The model has its most applications in time use models, where multiple alternatives are chosen and the amount of time spent in each of them. This assumes that the choice is made a priori (e.g. before the day starts) and the decision of all the activities are made simultaneously (horizontal choice) as opposed to a sequential (vertical) decision process.

As the delivery method is mostly planned ahead, the choice of alternatives is done simultaneously and diminishing returns of each alternative is present. In a way, situations that in passenger/massive consumption choices are “vertical”, in managerial/freight context can be “horizontal” meaning that multiple discreteness can be present. No application in regional freight

was found. In cases when a strategic or planning process is performed, which is more common in freight transport than in passenger transport, simultaneous (horizontal) choices of several alternatives can be made, thus breaking the perfect substitution between alternatives that would otherwise hold. However, the papers do not explicitly treat the choices as part of a planning problem.

In the particular case of mode and port choice, there can be cases where multiple-discreteness is present. At a tactical planning level (Tavasszy and de Jong, 2014), where transport tariffs, aggregated volumes and contracts are negotiated, some shippers may decide to plan to send goods through different modes and commercialization channels. The benefits from this would be: having more flexibility, a good average price, quick response, and lower overall supply chain disruption risk.

Although this behaviour may exist, no work has been found that addresses it explicitly. This paper aims to introduce multiple-discreteness analysis in regional freight transport by analysing port and mode choice of Argentinian grain consolidators, making clear the planning nature of the choice.

Agriculture is a dynamic and important sector in Argentina's economy. It is carried out by a large number of scattered producers at a relatively small distance to the export ports. Some authors have pointed out that this short distance is one source of the low competitiveness of the railroad (Barbero, 2010; Regunaga, 2010) and this has a direct impact on the competitiveness of the Argentinian agricultural sector as a whole (Schnepf et al., 2001).

The industries and exporting complexes are mainly located near the three main exporting ports. The most important one is Rosario, where 70% of the grain is exported and the main crushers are located. The second most important is Bahia Blanca, located in the South Atlantic with 16% of the grain exports. The port of Quequén is also on South Atlantic and moves 12% of the exports. Smaller terminals along the Paraná River export the rest of the grains.

The inland transport is heavily road transport oriented. Around 84% of the products are transported this way, 14% by rail and 2% by waterway (Barbero, 2010; Cohan and Costa, 2011). While the road network is dense and covers most of the study area, the railroads have much less coverage and are heavily port oriented.

The grain supply chain is organized as follows. After the harvest, the producers have to decide between using temporary facilities for storage (*silobolsas*), sell directly to port with a broker or to send it to a consolidator for the commercialization.

The consolidators' business model is to sell supplies, such as fertilizers, seeds and agrochemical products, and offer services such as technical advice, storage and sell the grains. The fact that they store and commercialize the grains makes them key agents in the supply chain. They have the capacity to generate greater volumes and in that way to access the railroad. The larger the producer, the less likely it is that they use consolidators as part of their commercialization chain.

As harvests occur once a year per crop, it is feasible to have estimates on the volumes to be commercialized in advance, starting the look for potential buyers. This allows consolidators to plan ahead and negotiate transportation contracts and to analyse where to sell their products. Therefore, mode and destination volumes are probably allocated in advance and involve multiple alternatives. As most of the grain is sold in ports (due to export and industries being large located there), this involves port choice. This gives room to analyse the shipping of goods at a planning level considering the choice as a multiple discrete-continuous variable.

The objectives of this paper are threefold. Firstly, it is to model and describe the port and mode choice of grain consolidators as a multiple discrete variable. Secondly, to compare two

modelling frameworks for this situation. The first is the traditional discrete behavioural modelling, including ranked logit and (Chapman and Staelin, 1982) the Fractional Split (FS) model (Papke and Wooldridge, 1996; Sivakumar and Bhat, 2002). The other framework is based on the Multiple Discrete Extreme Value Model (MDCEV), proposed by (Bhat, 2008). Finally, it is to discuss the possibility of the continuous variable to catch smaller yet important trade-offs between alternatives, and thus enhancing data collection techniques. To do so, a Stated Preference (SP) survey is used with a continuous dependent variable that allows respondents to choose more than one alternative and how intensively they use it.

This paper contributes to the literature by considering the mode and destination choice at a planning level as a multiple discrete-continuous choice. The discrete choice alternatives are combinations of a mode and a destination (the port) and the continuous dimension is the amount of use of these combinations (in tonnes). The application of the MDCEV in this context is novel and can allow researchers to be more flexible in their assumptions when modelling choices at this level. As a consequence, more models will be available and a more accurate depiction of the choice maker behaviour can be obtained. Unlike earlier discrete-continuous models in freight transport we model a situation where multiple alternatives can be used at the same time, the model is directly based on micro-economic theory and simultaneous estimation is used.

Moreover, the paper compares the effect of assuming different types of behaviour in the accuracy and willingness-to-pay (WTP) measures. The former can affect volume predictions, while the latter can impact infrastructure evaluation, since WTP is a key parameter in Cost-Benefit Analysis. Overall, the paper supports the idea that tactical choices can and should have their own framework and models typically used in choice scenarios are not directly transferable.

The rest of the paper is organized as follows. Section 2 makes a brief literature review about multiple-discrete modelling and freight multiple discrete models. Section 3 describes the study region and the data collection process. Section 4 illustrates the method used for modelling and section 5 shows results and analysis. Finally, conclusions are shown in section 6.

## **LITERATURE REVIEW**

The appearance of the Multiple Discrete Extreme Value model (MDCEV) (Bhat, 2008, 2005) has been a landmark in the formulation and usage of discrete-continuous models. Bhat in 2005 proposed a model where, within a given time period, an individual could choose to engage in multiple activities. The choice modelled was which activities to participate in (discrete component) and for how long (continuous component). By breaking the assumption that alternatives are perfect substitutes he developed the MDEV that has a closed form and collapses to a Multinomial Logit (MNL) when only one alternative is chosen. In his paper time use choice is defined as “horizontal”, meaning that consumers choose a variety of alternatives simultaneously rather than as a result of multiple consecutive consumption decisions (“vertical” choices). The model can accommodate diminishing marginal returns. Bhat in 2008 reformulated the utility function of the MDCEV in order to clarify the role of the parameters, allow new correlation structures and discuss identification issues. The 2008 MDCEV version estimate the discrete and continuous component simultaneously, implying a single error term (Eluru et al., 2010). This has led to a new MDCEV proposed in 2018 that separates the baseline utilities from the continuous part, only for the cases where an outside good (i.e. an alternative that is always chosen) is present (Bhat, 2018).

After the first application, the MDCEV was used extensively in time use models (Astroza et al., 2018; Bhat et al., 2006; Calastri et al., 2017a; Copperman and Bhat, 2007; Enam et al., 2018;

Nurul Habib and Miller, 2008; Paleti et al., 2011; Sikder and Pinjari, 2013; Spissu et al., 2009), for week, weekends and holidays destinations. In time use models the budget and outside good are straightforward, since there is a firm and unified limit (e.g. 24 hours per day) and activities that everybody does (e.g. sleep).

Vehicle use is a frequent application of the model (Ahn et al., 2008; Bhat and Sen, 2006; Jian et al., 2017; Shin et al., 2012; Tanner and Bolduc, 2014), where the discrete dimension is which cars to own and the continuous decision is how much it is used (e.g. mileage). In contrast to the time use models, vehicle models do not have a fixed budget or a clear outside good (Eluru et al., 2010). As a result, the mileage has to be given exogenously to the model, making the model not responsive in terms of the total miles driven. Additionally, it was pointed out that in vehicle use models the choices are not necessarily done simultaneously, but as a consequence of several correlated and repeated choices, breaking the “horizontal” principle originally proposed (Bhat, 2005).

The lack of “horizontality” also appears in other applications of the MDCEV, especially when individual consumers are the choice makers. In energy modelling applications (Acharya and Marhold, 2019; Biying et al., 2012; Yu and Zhang, 2015) the choices are (generally) which fuels/energy sources are used and how much they are used. Something similar happens in social media use/interaction (Calastri et al., 2017b; Woo et al., 2014) or in transport expenditure models, where the choices can be understood as a sequence of smaller (discrete) choices related among themselves through an external budget (Anowar et al., 2018; Bhat et al., 2015; Ferdous et al., 2010).

The situation is different in products promotion (i.e. discounts) modelling (Lu et al., 2017; Richards et al., 2012). In this case, the assumption is made that the choice of multiple consumption is simultaneous at a given time, which is when the consumers are in the supermarket. However, it shares the problem that the budget is exogenous.

Regarding company choices, where freight choices are contained, MDCEV can prove to be an interesting tool. Ko and Kim (2017) used the MDCEV to model which transport demand management strategies companies in Seoul use. The MDCEV was used because each organization can be part of several strategies simultaneously that are likely to be conscious and planned. Another use of the MDCEV was to model tour chaining and time of day departure, (Khan and Machemehl, 2017a, 2017b). These are the only applications of the MDCEV in freight so far to the best of our knowledge.

Other attempts in freight of including a continuous variable in choice can be seen in single discrete continuous models, where only one discrete alternative is chosen together with the intensity of such choices, such as Abdelwahab (1998), Abdelwahab and Sargious (1992), de Jong and Johnson (2009), Holguín-Veras (2002) and McFadden et al., (1986).

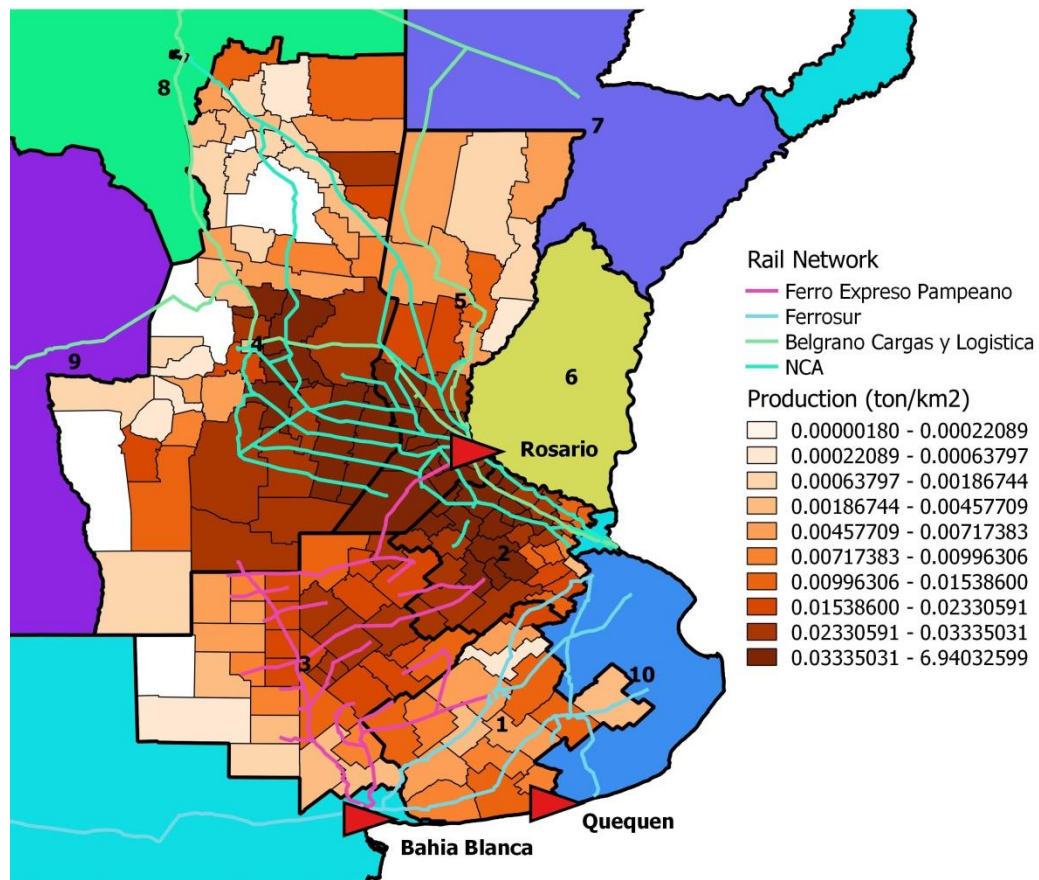
With an intermediate approach between the MDCEV and the single discrete models, Rashidi and Roorda (2018) have used a multiple discrete approach to model the fleet composition of companies for urban freight in Canada. The decision modelled was the vehicle type (car, van or truck) and the fleet size for each type as a discrete variable. The MDCEV approach was not used in this case because the intensity of the choice is also discrete (number of vehicles) and not mileage as in other vehicle use models.

## **METHOD**

### **Data Collection**

Stated preference (SP) surveys present the respondents with hypothetical choice situations in order to collect data for their preferences. The scenarios can be created by the researcher, allowing to test for alternatives that are currently not available. This flexibility is used in this paper to show rail alternatives to all respondents, independently from the actual rail availability in the region. Despite its advantages, SP can be subject to hypothetical bias and other types of respondent induced biases (e.g. non trading, lexicographic or inconsistent behaviour (Hess et al 2010)). The hypothetical nature of the data makes the SP models not reliable for forecasting due to scale effects, but it has been widely used to estimate WTP values.

For this study, the country has been divided into 10 zones taking into account port hinterlands and railway services, as shown in figure 1. Areas that had one dominant mode or one dominant port have been discarded. Zone 1 contains the mixed hinterland between Quequén and Bahía Blanca, with *Ferrosur* as a rail operator. Zone 2 is the area under the influence of the port of Rosario and the terminals in the north of the Province of Buenos Aires. Zone 3 contains the areas where the hinterlands of Rosario and Bahía Blanca coincide, with *FerroExpreso Pampeano* operating the rail lines. This zone is particularly relevant because the rail operator specializes in grain products and is an area where these two ports compete with each other. Zone 4 contains the province of Córdoba, and it is within Rosario's hinterland, and served by *NCA* Railway Company. This zone has the higher grain output of the country. Finally, Zone 5 is the region of the Paraná River close to Rosario with the *Belgrano Cargas y Logística* Railway Company attending them. The rest of the zones were discarded for the study.



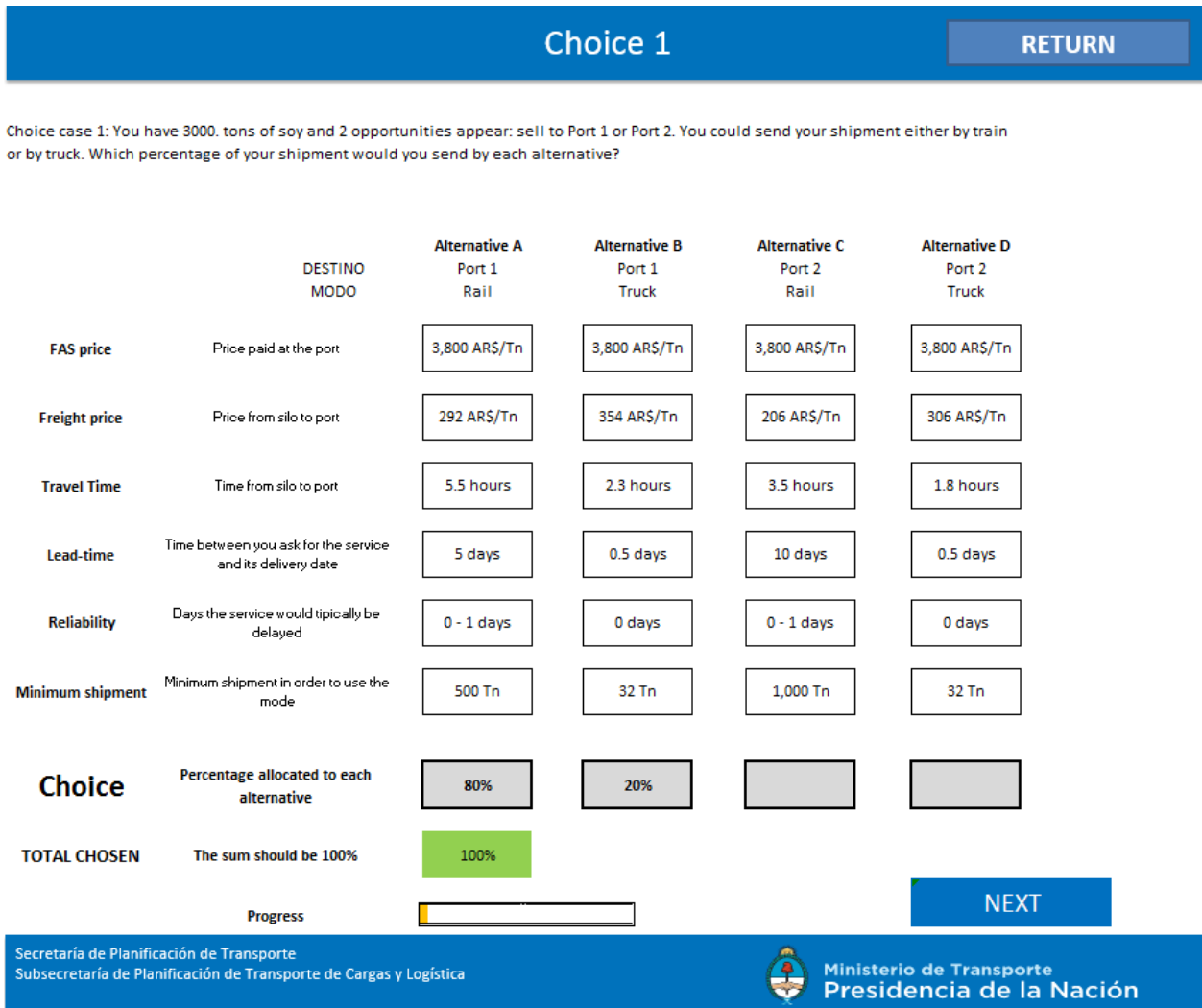
**FIGURE 1. Map of Argentina with the study zones and selected grain production**

The SP survey was conducted between May and September 2017. The commercial departments of grain consolidators of Zones 1 to 5 were targeted. Due to difficulties in acquiring contact information of producers, they were excluded from the survey. The survey was carried out with the support of the undersecretary of Logistic Transport and Planning of the Ministry of Transport of Argentina.

From an original population of 867 possible consolidators, 467 could be contacted by e-mail and 127 of them were contacted by telephone. The differences between these figures are due to incorrect or outdated contact details, rather than a sampling procedure. The telephone calls consisted on an explanation of the objectives, the general method, the survey instrument and the dependent and independent variables. Additionally, it was made clear that the answers will be used for policy making by the Ministry of Transport. This consequentiality of the answers has proven to diminish the bias in the responses (Crastes dit Sourd et al., 2018). 58 questionnaires were successfully completed, which is considered a good sample size in the freight context. In some cases, some observations had to be dropped, but not invalidating the respondent. In all, 670 observations (i.e. choice tasks) were used. The geographical coverage of the sample included all regions identified in figure 1, guaranteeing representation of the whole country's productive areas.

A Bayesian efficient design (Rose and Bliemer, 2009) was used to build a survey instrument with 12 choice tasks. Each choice task consisted of a commercial situation where the respondents had a finite amount of products (depended on the size and the main selling product, both informed by the interviewee) that they had to sell to two unlabelled ports and send them either by train or by truck. These generated four choice options to which the interviewees could freely allocate a percentage of the available cargo. The ports, although unlabelled, made reference to the closest ports to the interviewee in order to make the variable levels closer to reality and thus increase the likeliness of engagement of the respondent. Figure 2 illustrates the choice task card and Annex I shows the experimental design specification.





**FIGURE 2. Choice task of the SP.**

Each option, labelled A, B, C and D, was characterized by the Free Alongside Ship (FAS) price, freight price, travel time, lead-time or headway, reliability and minimum shipment size. All variables were tested at three levels.

There are several studies supporting the inclusion of freight price, travel time, frequency and reliability in mode choices in the freight context (Cullinane and Toy, 2000; Feo-Valero et al., 2011; Feo et al., 2011; Hoffmann, 2003; Shinghal and Fowkes, 2002; Zamparini et al., 2011). Regarding the FAS price, there is less evidence (Tapia et al., 2019), but it is considered important in the grain context because it is the main reference signal made by the exporters to the consolidators. The minimum shipment size was included to increase compatibility with models currently used by the Ministry of transport and because it is considered to be a barrier for the access to the rail mode.

FAS price is the price paid to the producer at the ports. The reference taken was for wheat, sunflower, maize, soy and sorghum in Rosario the 09/06/2017. The alternatives were both ports either had the same price, or the further away paid a 2.5 or a 5% extra.

The trucks freight prices used were the official published ones by the truck drivers' union (CATAC, 2017). Since it is a published price, it was considered unrealistic if they were changed so they remained constant in all choice sets. The prices for the train were set as a 67.5, 75 and 82.5% of the truck prices. These reflected the full price from storage to destination.

Something similar was set for travel time. A reference speed for truck of 65 km/h was taken and fixed for all scenarios. This speed is just a seed value does not imply actual velocity of the trucks. The levels for the train mode had a baseline of 30 km/h with two other levels being 90 and 110% of that value. The time was defined as the travel time between storage and destination.

The lead-time was defined as the maximum time the shipper had to wait for the transport service to arrive. As a reference for the truck, this was set as 0.5 days and for rail ranged between 5,7 and 10 days. Reliability was defined as the number of days that the loading time of the train could be delayed. The values were 0, 1 and 3 days.

Finally, the minimum shipment size was the minimum volume that had to be sent by this mode in order for it to be available. It was fixed as a full truckload for the road mode (32 tons) and ranged between 500 tons to 1,500 tons for rail.

The prior values for travel time, freight price, FAS price and lead-time were used from an earlier version of the model in Tapia et al., (2019), that was a study made with the same target population, only in a smaller region. For reliability, the work carried out by Larranaga et al. (2017) was considered. No prior for the shipment size coefficient was available, so a zero centered value was used.

Normally, SP experiments offer a discrete response variable, but in this study a continuous variable is offered to the respondents. This way, information about small trade-offs are expected to be captured. For instance, a change of parameters that would shift the choice for some alternative from 30% towards 40% would not be necessarily enough to trigger a discrete choice shift, meaning some loss of information if it were a purely discrete choice.

### **Modelling framework**

For modelling this response variable can be treated in five ways: i) following the fractional split model; ii) MDCEV model; iii) discretising on the basis of highest probability (Tapia et al., 2019); iv) discretising using a distribution and; v) using a ranked response.

The response variable in i and ii does not need any treatment in order to be modelled, while iii, iv and v need to go through a discretization process. The discretization to the highest probability consists of allocating a discrete choice on the basis of which alternative has the highest stated value, as it was done in (Tapia et al., 2019). Ties were treated *ad hoc* in order to favour the choice shift from the previous choice. For example, consider the following choice: A = 60%; B = 10%; C = 30%; D = 0%, followed by A=50%; B=0%; C=50%; D=0%. The first choice by the discretization method would be A (the highest) and the second one C (so the information that A has lost to C the preference is not lost). In this paper, this discretization will be called Max.

The other form for discretization is taking draws from a distribution for simulating the choice. In this paper a uniform distribution between 0 and 1 was used. Using the example above, if the uniform draw fell between 0 and 0.6 A would be chosen; if it was between 0.6 and 0.7, alternative B would be chosen and; if it fell between 0.7 and 1 alternative C was picked. For the second choice the intervals would be: A between 0 and 0.5 and C between 0.5 and 1. In this paper, this discretization will be called Uniform. Max and Uniform discretization have the same choices in 67.4% of the times.

The last method used consists in considering the response variable as an ordered response and generating an exploded ranking. Using the values from the example, the ranked choice would be A:1; B:3; C:2; D:4. For the second choice ranking would be A:1; B:2; C:1; D:2. According to the ranking, the choices were exploded according to the exploded ranking procedure (Chapman and Staelin, 1982). The exploded choice the first choice would imply three discrete choices: i) A from the choice set {A;B;C;D} (same situation as in Max); ii) C from the choice set {B;C;D} and; iii) B from the choice set {B;D}. The second choice would be exploded into two: i) A from the choice set {A;B;C} and; ii) C from the choice set {B;C;D}. In this paper, this discretization will be called Ranked.

The Ranked, Uniform and Max model are built on the standard likelihood function of the discrete choice models. The Fractional Split (FS) allows to model modal split data based on the aggregate (zone-to-zone) data, where the dependent variable consists of the market shares of the modes for each origin-destination pair. Besides the application for regional freight in (Sivakumar and Bhat, 2002), applications in accident analysis have used this specification (Eluru et al., 2013; Lee et al., 2018). Please note that in our application of this specification we use data at disaggregate (individual firm) level. In this paper, it will be mentioned as the FS model; with a log-likelihood (LL) function is given by equation 1.

$$LL = \sum Prop_{ji} * \log(P_{ji}) \quad (1)$$

Where  $Prop_{ji}$  is the percentage allocated to each alternative i and  $P_{ji}$  is the probability of that alternative to be chosen in the choice j. This LL collapses to the MNL in cases where a discrete choice (e.g 100% is allocated to one alternative) is made. Nested and non-nested structures were tested for this model.

The second framework used was the MDCEV model. This model consists of a baseline utility with added satiation parameters. Due to the complex formulation of the nested MDCEV, correlations between alternatives were tested using an error component model (Train, 2003). The LL function is given by equation 2.

$$LL = \sum \left( \frac{1}{\sigma^{M-1}} \left[ \prod_{i=1}^M c_i \right] \left[ \sum_{i=1}^M \frac{1}{c_i} \right] \left[ \frac{\prod_{i=1}^M e^{V_k/\sigma}}{\left( \sum_{k=1}^K e^{V_k/\sigma} \right)} \right] \right) (M - 1)! \quad (2)$$

Where  $c_i = \left( \frac{1-\alpha_i}{e_i^* + \gamma_i} \right)$ ,  $V_k = \beta z_k + (\alpha_k - 1) \ln \left( \frac{e_k^*}{\gamma_k} + 1 \right)$ ,  $\beta z_k$  is a vector of coefficients and attributes,  $e_k$  the amount chosen for each available alternative k, M are the chosen alternatives,  $\sigma$  is the error scale parameter and  $\alpha_k$  and  $\gamma_k$  satiation parameters. The satiation parameters allow the model to change the marginal utility of extra consumption of the alternatives. Their effect is to decrease the marginal utility when the alternative is being consumed and thus can be interpreted as the consumer becoming more and more satiated.

The effects of  $\alpha$  and  $\gamma$  as satiation parameters are confounded in the model and cannot be estimated simultaneously. While  $\gamma$  allows corner solutions and zero consumption,  $\alpha$  reflects the marginal satiation rate. In order to estimate them, either one of them have to be fixed. When  $\alpha$  is fixed, a  $\gamma$  profile is estimated, which has the advantage of having a more straightforward forecasting method and interpretation.

In the SP, none of the alternatives was a clear favourite and none of them was always chosen. This makes that the MDCEV formulations that consider an outside good (that is always chosen) were not suitable.

The core difference between the MDCEV and the other five models reside in its approach towards substitution. On one hand, the FS, Ranked, Uniform and Max assume that the alternatives are mutually exclusive, as all traditional single discrete choice models. On the other hand, the MDCEV relaxes this assumption to allow the choice of multiple alternatives simultaneously in a consistent “horizontal” way, as would happen in a planning scenario. The FS in particular has the possibility of accommodating several alternatives, but the mutual exclusiveness assumption can imply that the choices are vertical (i.e. sequential).

For all models, linear and non-linear parameters in the utility function were tested. Several reports highlight the importance of non-linearities in the freight context (Gatta and Marcucci, 2016; Marcucci et al., 2015).

Observed heterogeneity was tested with the inclusion of variables related to the type of consolidator and its location, such as size, truck ownership. The interaction was defined as the product of the coefficient of the variable  $\beta$  and  $(1 + \beta_{interaction} * \delta)$ ,  $\beta_{interaction}$  being the interaction coefficient and  $\delta$  the dummy for the interaction effect. This means that the interpretation of the interaction parameter is that of a magnifying effect. If the dummy takes the value of zero, the overall coefficient is  $\beta$ , and if it is 1 it becomes  $\beta * (1 + \beta_{interaction})$ . When  $\beta_{interaction}$  is positive, it enlarges the effect of the variable, if negative and smaller than one, it diminishes it. For values below -1 the sign of the coefficient is reversed, which usually does not make sense econometrically.

Random heterogeneity was modelled with the inclusion of uniform, log uniform and lognormal distributions of the parameters with 250 Halton draws (Train, 1999). Normal distributions were discarded because of being unbounded distributions, which can give either values with the wrong sign or problems estimating Willingness to Pay (WTP) measures (Daly et al., 2012b). For this matter, lognormal distributions are favoured. Generally, observed heterogeneity is preferred over the random one because it provides better behavioural insights. The panel nature of the data was considered for the estimation and for the Halton draws.

The criteria for choosing the models were not only to have the higher LL. The parameters sign and microeconomic logic and behavioural insights were taken into account too.

The key difference between the MDCEV models and the other models in this paper is the key assumption around the discreteness of the response variable and the mutual exclusiveness of the alternatives. In order to assess the effect of this assumption, the models will be compared to each other.

To compare the models, the data was split into training and testing parts. The training data consisted of 80% of the participants, meaning that the answers for the same interviewee were on the same database. In total 50 rounds of random splitting were done.

For forecasting with the model, firstly, the baseline utility is computed by simulating the Gumbel distributed error, ordered decreasingly and assumed it is consumed. Secondly, the Lagrange multiplier of the utility maximization problem is calculated and compared with the baseline utility function of the next highest alternative. If smaller, the second alternative is chosen and the step will continue to iterate until, there are no more alternatives left, or the Lagrange multiplier is larger than the next baseline utility. When the chosen alternatives are selected, the optimal consumption is estimated. The forecasting algorithm is described and demonstrated in detail in (Pinjari and Bhat, 2011). First the results are averaged across the draws for each hold out sample, obtaining mean absolute error for the continuous part and type I and II

errors for the discrete part. All the measures were then averaged across all 50 hold out samples. The measures used to compare the models are the mean absolute error per alternative (Jäggi et al., 2013).

Additionally, type I and type II errors were obtained from the test sample. A type I error (false positive) occurs when the model allocates some percentage to the alternative but the respondent did not choose the alternative. Type II errors (false negatives) occurs when the model predicts that the alternative is not chosen, but the respondent did select it. For the discrete models, type I errors are equal to the times the alternative is not chosen, and type II equal to zero because it always gives a non zero probability for every alternative. For each test sample, both errors were counted and summarized as a percentage of the amount of choices.

Willingness to pay (WTP) measures were estimated using the delta method (Daly et al., 2012a) adapted for the mixed logit case, presented in Bliemer and Rose (2013). The method takes into account the error of the estimates and simulates the dimension(s) of the random parameter generating a WTP distribution. The main differences to other simulation methods is that the one in Bliemer and Rose provides symmetrical normal distributions, while other methods replicate the shape of the simulated data. Other simulation-based WTP measures were also estimated for comparison.

## **RESULTS AND ANALYSIS**

The continuous response variable of the SP experiment was well accepted by the respondents. Of all, only five of them used only one alternative in all the 12 choices and of the 670 observations, only 100 were for only one alternative. This shows that in the case of the experiment, the alternatives were not perfect substitutes of each other.

Table 1 shows the results of the models for the different response variables. As expected, freight price, travel time and headway have a negative sign and FAS price had a positive influence in the utility function. In both models, the reliability and the minimum shipment size were not significant. Error components for the MDCEV were tested, but were found not to be significant.

**TABLE 1: Model estimates**

	Multiple discrete framework		Discrete choice framework		
	MDCEV	Fractional Split	Uniform	Max	Ranked
<b>sigma</b>	1 (Fixed)	-	-	-	-
<b>Alpha</b>	0 (Fixed) <sup>3</sup>	-	-	-	-
<b>ASC_A</b>	0 (Fixed)	-	-	-	-
<b>ASC_B</b>	-1.29 (-1.54)	-0.92 (-2.17)	-0.62 (-1.1)	-1.22 (-2.05)	-0.29 (-0.4)
<b>ASC_C</b>	0.03 (0.09)	-0.53 (-2.15)	0.03 (0.07)	-0.79 (-1.95)	-1.11 (-2.67)
<b>ASC_D</b>	-1.29 (-1.28)	-0.94 (-1.76)	-0.18 (-0.25)	-1.30 (-1.61)	-0.22 (-0.25)
<b>TIME</b>	-1.04 (-1.8) <sup>1</sup>	-1.37 (-3.42) <sub>1</sub>	-0.90 (-1.86) <sup>1</sup>	-1.2 (-1.99) <sup>1</sup>	-0.93 (-1.2) <sup>1</sup>
<b>LEAD TIME</b>	-0.84 (-3.92) <sup>1</sup>	-0.11 (-4.31)	-0.095 (-2.87)	-0.18 (-4.42)	-0.14 (-3.41)
<b>FREIGHT PRICE</b>	-1.43 (-8.7) <sup>2</sup>	1.47 (9.91) <sup>12</sup>	1.44 (6.85) <sub>12</sub>	1.59 (8.41) <sup>12</sup>	1.86 (11.14) <sub>12</sub>
<b>FAS</b>	-1.20 (-5.47) <sup>2</sup>	-1.60 (-7.25) <sub>2</sub>	-1.55 (-6.2) <sub>2</sub>	0.19 (3.44)	0.26 (3.76)
<b>FAS (Std) *</b>	-1.18 (-4.84)	0.68 (4)	0.74 (4.25)	0 (Fixed)	0 (Fixed)
<b>FREIGHT PRICE (Std)*</b>	0.70 (6.82)	-0.67 (-5.97)	-0.72 (-5.34)	-1.12 (-8.59)	-1.08 (-6.59)
<b>Gamma_A</b>	2.68 (12.45) <sup>3</sup>	-	-	-	-
<b>Gamma_B</b>	2.43 (13.18) <sup>3</sup>	-	-	-	-
<b>Gamma_C</b>	2.90 (14.87) <sup>3</sup>	-	-	-	-
<b>Gamma_D</b>	== Gamma B	-	-	-	-
<b>Gamma_D * Own Trucks</b>	0.555 (1.83) <sup>3</sup>	-	-	-	-
<b>LL(0)</b>	-7850.4	-928.817	-928.817	-928.817	-3976
<b>LL(final)</b>	-4908.23	-757.373	-760.844	-673.837	-2562.08
<sup>1</sup> indicates a logarithmic transformation;					
<sup>2</sup> Mean of the log of the distribution; coefficient sign given in the transformation;					
<sup>3</sup> Values parametrized according to (Bhat, 2008)					
* standard deviation of the log of the distribution;					
t values given in brackets					

The FS model showed that the travel time and freight price have a logarithmic relation with the utility function, while the FAS price and lead-time had a linear parameter. Heterogeneity was only captured by a log uniform distribution for the freight price. Since it is a lognormal distribution, the coefficient sign was included in the utility function by the authors, so the fact that it is positive does not mean that it has a positive impact in the utility.

It is worth mentioning that the ranked model does not have a significant parameter for travel time, neither in the utility space nor the WTP space. Nevertheless, it will be kept for the measurement of the errors in order to make the models more comparable. For the MDCEV and

uniform models the time parameter was significant at a 90% level of confidence and considered to be acceptable.

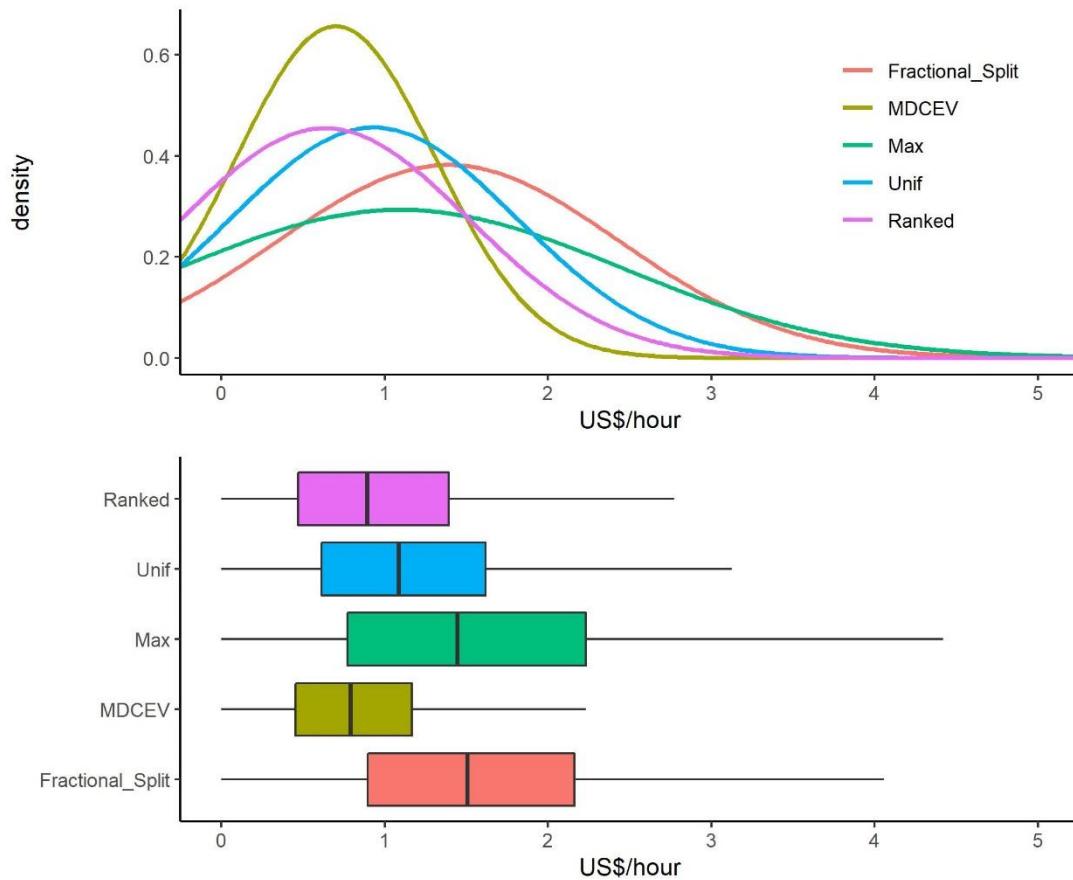
For the MDCEV, the lead-time and the travel time had a logarithmic relation with the baseline utility function. Heterogeneity was captured with the inclusion of an interaction of truck ownership with the lead-time and the satiation parameter for the closest truck alternative and with lognormal coefficients for FAS price and freight price (same appreciations of the coefficient signs as in the FS model).

In the satiation parameter an interesting interpretation is given by the MDCEV. All satiation parameters are similar in magnitude (they were tested to see if they were significantly different), with both truck alternatives having the same satiation level. Both alternatives related to train had larger values, indicating lower satiation. This means that once the train has enough utility to be chosen, it will be used proportionately more than the truck alternatives.

Regarding satiation heterogeneity, if the consolidator owns trucks they would have less satiation for the closest port by road (alternative D). This would mean that if the consolidator were to use trucks, they would use them more intensively to the closest port. This would maximize their utilization rate of the trucks by making them do more trips.

In general, it can be seen that the model structure of the discrete choice models is similar, regarding the non linearities and the way they capture heterogeneity. They show some differences with the MDCEV that can affect the model, depending of the objective. In order to compare the magnitude of the parameters WTP measurements compared.

WTP measures were estimated for the travel time and for the lead-time. The distribution of the price coefficient was simulated through 10.000 draws and non linearities were evaluated at the mean value. The median value was used for the estimation of the delta method because it provides less bias towards outliers (Bliemer and Rose, 2013). Figure 3 shows the resulting distributions and main values of WTP for travel time for train and truck, estimated for all models.

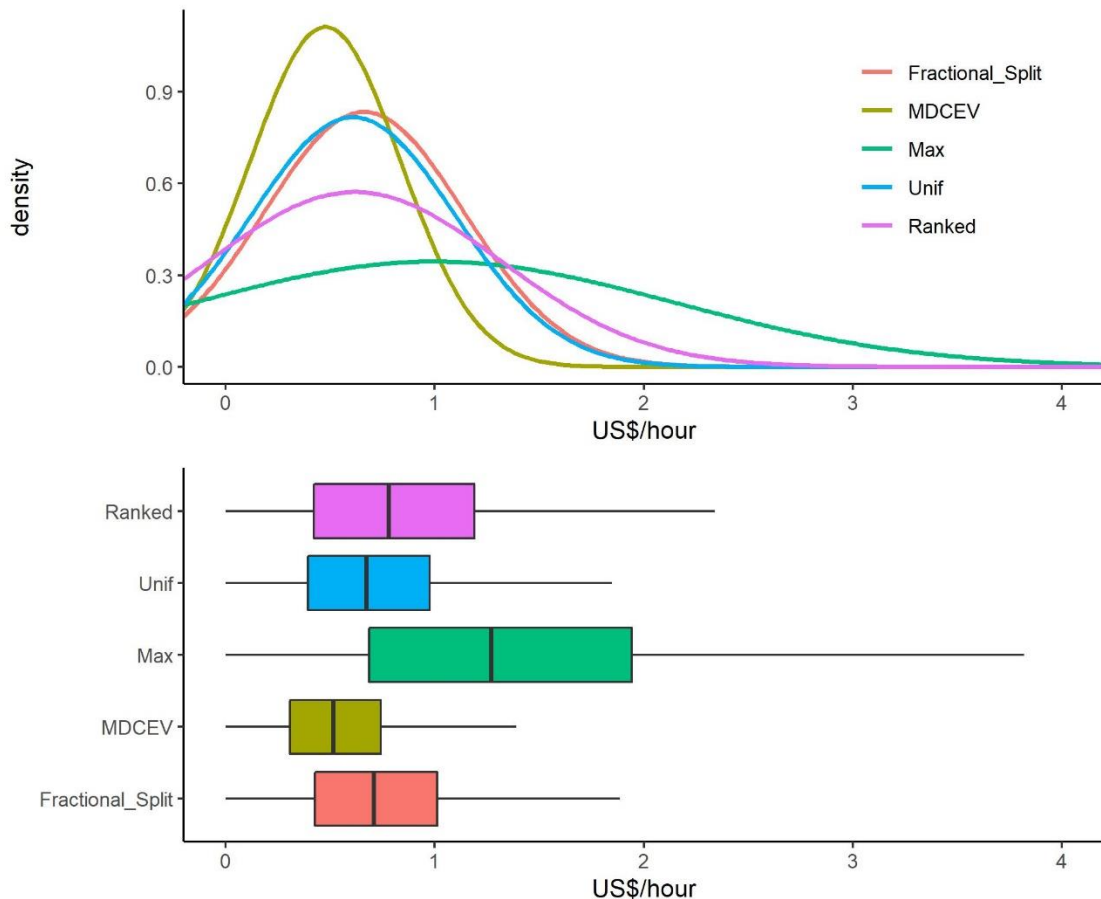


**Figure 3 WTP distributions for travel time**

The MDCEV showed the lowest variance of all the models, shown by the higher peaks and the lower interquartile range. The MDCEV also showed the lowest median value, of 0.70 US\$/ton/hour. The largest values were reported for the FS model, with a median of 1.39 US\$/ton/hour. The largest variability is given by the Max model. The medians for truck ranged from 1.11 to 2.53 US\$/ton/hour and for train from 0.40 to 0.87 US\$/ton/hour. These values are in line with the ones found in the literature, where Tapia et al. (2019) estimated a value of 1.49 US\$/ton/hour for the same context. Larranaga et al. (2017) and de Jong et al. (2013) reported values from the literature in between 0.1 to 3.4 US\$/ton/hour.

The WTP to reduce a day of lead-time can be interpreted as the WTP for a reduction in a day in the inventory, due to characteristics of the product involved (agricultural commodity). As there is a concentration of demand for storage during the harvest, a lower lead time can allow for a better usage of the fixed capacity. Figure 4 shows the distribution and main values for the WTP for lead time.





**Figure 4 WTP distributions for lead time**

As it happened for the WTP for travel time the MDCEV had the lowest variability and values of all the models and the Max the highest. The median values for the MDCEV were 0.48 US\$/ton/day and for the FS 0.66 US\$/ton/day. The ranked model, that did not have a significant time coefficient, showed a larger variability and values than the MDCEV. All the median values found are within range from the 0.99 US\$/ton/day obtained by Tapia et al. (2019).

Table 2 summarises the confidence intervals (CI) for the WTP of travel and lead time using the delta method to obtain the 2.5%, median (50%) and 97.5% quantiles. Even though part of the CI is negative for some models, still the larger part of the distributions have positive values and the objective of this paper is to analyse the impact of the model assumptions in WTP measures, rather than obtaining highly significant WTP measures. It is also worth noting that except for the MDCEV values, the cost coefficient is not linear, making it unsuitable for Cost-Benefit Analysis.

The CIs were also estimated through simulation using a simulation method also suggested in Bliemer and Rose (2013); also see Gatta et al. (2015) specifically for small samples. It relies on the Krinsky and Robb (Krinsky and Robb, 1986, 1990) procedure and the method described in Hensher and Greene (2003) to simulate the WTP distribution and can provide asymmetrical CIs. The results were 50% to 300% wider CIs compared to the delta method. The percentage of values above 0 was larger in most cases. This percentage followed the significance of the time/lead time parameters, since the cost coefficient is constrained to negative values.

**TABLE 2: Confidence intervals of the WTP values**

	Quantiles			Percentage of values >0
	2.5%	50%	97.5%	
<b>WTP Time FS [US\$/ton/hour]</b>	-0.64	1.39	3.41	91
<b>WTP Time MDCEV [US\$/ton/hour]</b>	-0.48	0.70	1.85	88
<b>WTP Time Max [US\$/ton/hour]</b>	-1.55	1.09	3.73	79
<b>WTP Time Unif [US\$/ton/hour]</b>	-0.77	0.93	2.63	86
<b>WTP Time Ranked [US\$/ton/hour]</b>	-1.07	0.63	2.33	77
<b>WTP Head Prop [US\$/ton/day]</b>	-0.27	0.66	1.59	92
<b>WTP Head MDCEV [US\$/ton/day]</b>	-0.22	0.48	1.17	91
<b>WTP Head Max [US\$/ton/day]</b>	-1.25	1.00	3.24	81
<b>WTP Head Unif [US\$/ton/day]</b>	-0.34	0.61	1.56	90
<b>WTP Head Ranked [US\$/ton/day]</b>	-0.74	0.62	1.97	81

It could be argued that the MDCEV's assumptions lie closer to the observed behaviour of the respondents since it has a framework that allows the choice maker to choose multiple alternatives and their intensity. By being able to reflect the behaviour more accurately, the MDCEV is also able to reflect the trade-offs between the variables, and thus the WTP. Moreover, the CI of the MDCEV is included in the CIs of the other four models, supporting this claim further. This could imply that the traditional discrete choices overstate the WTP values. This can bring consequences if the models are used for infrastructure evaluation, since the travel time and headway savings valuation would be lower when using the MDCEV.

The models were also compared on a test sample. Table 3 shows the results of the errors of the forecasting procedures by the models ordered according to their overall accuracy.

**TABLE 3: Mean absolute error of the models**

Alternative	MDCEV	Ranked	Uniform	Max	Fractional Split
<b>A</b>	40.24	25.16	24.56	24.75	24.1
<b>B</b>	30.95	23.87	19.6	21.46	20.02
<b>C</b>	35.6	25.09	24.95	26.17	23.47
<b>D</b>	29.83	17.69	19.16	18.54	18.52
<b>Total</b>	34.16	22.95	22.07	22.73	21.53

It can be seen that the MDCEV shows the lowest accuracy and the FS the highest, being the average absolute error approximately 60% higher from the MDCEV. Nevertheless, all models showed relatively high error rate. The relatively good performance of the FS compared to the

other logit choices is not a surprise because it uses all the information from the choices, including the small trade-offs that are not large enough to trigger a modal shift or even the ranked order. The ranked model was expected to capture better the information from the response variable, but it showed a bad model specification and performance.

An analysis of type I and type II errors of the models can explain the bad performance of the MDCEV. Traditional discrete choice models always allocate some percentage to every alternative. As a result, the amount of false positives (type I) is equal to the times the alternative is not chosen and the false negatives (type II) equal zero. Table 4 shows the results of the hit rate analysis.

**TABLE 4: Type I and Type II errors of the models**

	<b>MDCEV</b>	<b>Ranked</b>	<b>Uniform</b>	<b>Max</b>	<b>Fractional Split</b>
<b>Type I A</b>	20%	62%	62%	62%	62%
<b>Type I B</b>	26%	69%	69%	69%	69%
<b>Type I C</b>	31%	71%	71%	71%	71%
<b>Type I D</b>	30%	73%	73%	73%	73%
<b>Average Type I</b>	27%	69%	69%	69%	69%
<b>Type II A</b>	26%	0	0	0	0
<b>Type II B</b>	27%	0	0	0	0
<b>Type II C</b>	20%	0	0	0	0
<b>Type II D</b>	24%	0	0	0	0
<b>Average Type II</b>	24%	0%	0%	0%	0%

The table above provide some explanations for the lower performance of the MDCEV, shown by the percentage of errors relative to the total amount of choice tasks. False negatives (type II) bring more error than type I errors. False negatives imply that the model assigns a zero quantity of goods when it was effectively chosen. This makes the error to be higher than the discrete choice model, which allocates to every alternative a non-zero probability of being chosen. The proportion of type I and type II errors are balanced in average. Jäggi et al. (2013) suggested that the MDCEV has higher false predictions rate when less multiple choices are observed.

In light of this weakness, Bhat (2018) proposed a new MDCEV model that amends the lack of accuracy in the estimation of the discrete part. So far, it only contemplates cases with an outside good, not suitable for the present work.

Despite the lower accuracy of the MDCEV, it has interesting features that makes it attractive from a normative standpoint. It gives insights that improve the behavioural understanding of satiation parameters, the ones that suggest the existence of multiple-discreteness in this regional freight scenario.

## **CONCLUSION**

Mode and destination choice is normally considered a purely discrete one. Nevertheless, there are examples where this decision could be analysed considering that the choice-maker chooses more than one alternative. This could be the case for choices at a tactical or strategic level such as dual sourcing or the allocation of shipments into different modes. This multiple allocation can bring the organization a greater mix in overall cost, flexibility offered to the clients and reduce

the overall risk of the transaction. Despite the existence of these situations, no work has been found with this focus.

In this paper, mode and port choice was successfully modelled as a discrete-continuous choice, considering the mode/port choice as a discrete choice and the amount of cargo allocated to each alternative as a continuous part. This suggests the possibility of having multiple-discreteness in the regional freight context. The main assumption is that in the case study the mode and port is made first at an aggregated planning level, where the choice maker can decide the mix of delivery times, cost and other level of service characteristics. By extrapolating this assumption, the framework applied in this paper can be used to expand the tools available for modelling freight transport. This can allow researchers to use more flexible frameworks with better behavioural ties to model freight planning, which constitutes the main contribution of the paper. Additionally, the use of a continuous variable can potentially provide an effective way of capturing more information, even when using a discrete choice framework (such as the FS).

A SP survey was carried out with 58 respondents and a wide geographical coverage. A continuous response variable was used, allowing respondents to allocate freely their cargo among four alternatives, each combining two unlabelled ports and two modes (truck and rail). This flexibility was widely used among the interviewees, having most of them choosing more than one alternative at least at one point. This also suggests that the alternatives were not mutually exclusive. Using this continuous variable allowed to extract more information per response because it allowed little trade-offs that would otherwise be masked in a purely discrete choice.

A multiple discrete and several discrete choice frameworks were used for this data. For the MDCEV and FS no treatment of the data was needed and three discretization alternatives were tested (Max, Uniform and Ranked). All models included travel time (not significant for the ranked model), lead-time, freight price and FAS price as explanatory variables. Random heterogeneity was captured by lognormals in the freight cost for all models and in the FAS price for the MDCEV, FS and Uniform. To our knowledge, this is the only application so far of the MDCEV in freight regional context. The models with a discrete choice framework shared the logarithmic transformation for travel time and freight cost, while the MDCEV had transformations for the freight cost and lead time.

The MDCEV offers a framework that can reflect satiation patterns. It is observed that higher satiation occurs for truck compared to rail. The heterogeneity captured in the MDCEV has been related with the ownership of truck. These show a lower satiation if the alternative of the truck to the closest port is chosen.

WTP values were within the expected range, both for time and for lead-time. In general, the WTP values were higher for the discrete choice models than the MDCEV. This paper argues that the assumptions of the MDCEV are closer to the behaviour of the choice makers in this situation and probably better reflects the trade-offs (and thus the WTP) of the situations. Therefore, it could be an indication that the fractional split model overstates the value of the WTP. As a consequence, some of the values used for infrastructure appraisal within the welfare economics could be overstating the benefits of travel time and service savings.

Regarding model performance, the discrete choice models show better accuracy. This could be because of the relatively lower performance in the discrete forecast of the MDCEV. There is evidence in previous research that suggests that a low number of simultaneously chosen alternatives, as is the case here, can be responsible for this. Nevertheless, this model brings interesting insights of the satiation patterns in the context of multiple-discreteness. In this way, it is still a useful tool to understand freight behaviour.

This results highlight the importance of having clear the objectives of the model beforehand. Discrete choice models are simpler and show a (relative) good predictive performance. However, MDCEV has a better theoretical and econometrical performance, giving useful behavioural insights and probably more accurate WTP measures.

Overall, this works uses an SP that allowed to use more information from the respondents by introducing a continuous variable and modelled as a discrete-continuous phenomenon using the MDCEV, a relative innovation in the regional freight modelling studies.

As mentioned before, the shippers can obtain different mixes of level of services, average costs and reduce the overall risks by choosing more than one alternative. However, this work does not address this in particular and the study of how this affects decisions at a planning level is an interesting course of action for future research.

## ACKNOWLEDGMENTS

The authors thank the CAPES and the PSDE program for the research grants, the Lastran for support during the research process and to the under-secretary of cargo planning of the Ministry of Transport of Argentina for the data and support.

## REFERENCES

- Abdelwahab, W.M., 1998. Elasticities of mode choice probabilities and market elasticities of demand: Evidence from a simultaneous mode choice/shipment-size freight transport model. *Transp. Res. Part E Logist. Transp. Rev.* 34E, 257–266. [https://doi.org/10.1016/S1366-5545\(98\)00014-3](https://doi.org/10.1016/S1366-5545(98)00014-3)
- Abdelwahab, W.M., Sargious, M., 1992. Modelling the Demand for Freight Transport: A New Approach. *J. Transp. Econ. Policy* 26, 49–70.
- Acharya, B., Marhold, K., 2019. Determinants of household energy use and fuel switching behavior in Nepal. *Energy* 169, 1132–1138. <https://doi.org/10.1016/j.energy.2018.12.109>
- Ahn, J., Jeong, G., Kim, Y., 2008. A forecast of household ownership and use of alternative fuel vehicles: A multiple discrete-continuous choice approach. *Energy Econ.* 30, 2091–2104. <https://doi.org/10.1016/j.eneco.2007.10.003>
- Anowar, S., Eluru, N., Miranda-Moreno, L.F., 2018. How household transportation expenditures have evolved in Canada: a long term perspective. *Transportation (Amst)*. 45, 1297–1317. <https://doi.org/10.1007/s11116-017-9765-3>
- Astroza, S., Bhat, P.C., Bhat, C.R., Pendyala, R.M., Garikapati, V.M., 2018. Understanding activity engagement across weekdays and weekend days: A multivariate multiple discrete-continuous modeling approach. *J. Choice Model.* 28, 56–70. <https://doi.org/10.1016/j.jocm.2018.05.004>
- Barbero, J.A., 2010. La logística de cargas en América Latina y el Caribe: una agenda para mejorar su desempeño. *Banco Int. Desarro. BID* 1–7.
- Bhat, C.R., 2005. A multiple discrete-continuous extreme value model: Formulation and application to discretionary time-use decisions. *Transp. Res. Part B Methodol.* 39, 679–707. <https://doi.org/10.1016/j.trb.2004.08.003>
- Bhat, C.R., 2008. The multiple discrete-continuous extreme value (MDCEV) model: Role of utility function parameters, identification considerations, and model extensions. *Transp. Res. Part B Methodol.* 42, 274–303. <https://doi.org/10.1016/j.trb.2007.06.002>
- Bhat, C.R., 2018. A new flexible multiple discrete-continuous extreme value (MDCEV) choice model. *Transp. Res. Part B Methodol.* 110, 261–279. <https://doi.org/10.1016/j.trb.2018.02.011>
- Bhat, C.R., Castro, M., Pinjari, A.R., 2015. Allowing for complementarity and rich substitution patterns in multiple discrete-continuous models. *Transp. Res. Part B Methodol.* 81, 59–77. <https://doi.org/10.1016/j.trb.2015.08.009>
- Bhat, C.R., Sen, S., 2006. Household vehicle type holdings and usage: An application of the multiple discrete-continuous extreme value (MDCEV) model. *Transp. Res. Part B Methodol.* 40, 35–53. <https://doi.org/10.1016/j.trb.2005.01.003>
- Bhat, C.R., Srinivasan, S., Sen, S., 2006. A joint model for the perfect and imperfect substitute goods case: Application to activity time-use decisions. *Transp. Res. Part B Methodol.* 40, 827–850. <https://doi.org/10.1016/j.trb.2005.08.004>

- Biying, Y., Zhang, J., Fujiwara, A., 2012. Analysis of the residential location choice and household energy consumption behavior by incorporating multiple self-selection effects. *Energy Policy* 46, 319–334. <https://doi.org/10.1016/j.enpol.2012.03.067>
- Bliemer, M.C.J., Rose, J.M., 2013. Confidence intervals of willingness-to-pay for random coefficient logit models. *Transp. Res. Part B Methodol.* 58, 199–214. <https://doi.org/10.1016/j.trb.2013.09.010>
- Brooks, M.R., Trifts, V., 2008. Short sea shipping in North America: Understanding the requirements of Atlantic Canadian shippers. *Marit. Policy Manag.* 35, 145–158. <https://doi.org/10.1080/03088830801956805>
- Calastri, C., Hess, S., Choudhury, C., Daly, A., Gabrielli, L., 2017a. Mode choice with latent availability and consideration: Theory and a case study. *Transp. Res. Part B Methodol.* 123, 374–385. <https://doi.org/10.1016/j.trb.2017.06.016>
- Calastri, C., Hess, S., Daly, A., Maness, M., Kowald, M., Axhausen, K., 2017b. Modelling contact mode and frequency of interactions with social network members using the multiple discrete–continuous extreme value model. *Transp. Res. Part C Emerg. Technol.* 76, 16–34. <https://doi.org/10.1016/j.trc.2016.12.012>
- CATAC, 2017. Tarifario 2017 [WWW Document]. URL <http://www.catac.org.ar/tarifas.aspx> (accessed 10.25.18).
- Chapman, R., Staelin, R., 1982. Exploiting Rank Ordered Choice Set Data within the Stochastic Utility Model. *J. Mark.* 19, 288–301.
- Chow, J.Y.J., Yang, C.H., Regan, A.C., 2010. State-of-the art of freight forecast modeling: Lessons learned and the road ahead. *Transportation (Amst)*. 37, 1011–1030. <https://doi.org/10.1007/s11116-010-9281-1>
- Cohan, L., Costa, R., 2011. Panorama general de las nuevas formas de organizacion del agro: las principales cadenas agroalimentarias. *Nac. Unidas*.
- Copperman, R.B., Bhat, C.R., 2007. An analysis of the determinants of children’s weekend physical activity participation. *Transportation (Amst)*. 34, 67–87. <https://doi.org/10.1007/s11116-006-0005-5>
- Craetes dit Sourd, R., Zawojka, E., Mahieu, P.A., Louviere, J., 2018. Mitigating strategic misrepresentation of values in open-ended stated preference surveys by using negative reinforcement. *J. Choice Model.* 28, 153–166. <https://doi.org/10.1016/j.jocm.2018.06.001>
- Cullinane, K., Toy, N., 2000. Identifying influential attributes in freight route / mode choice decisions : a content analysis. *Transp. Res. Part E Logist. Transp. Rev.* 36, 41–53. [https://doi.org/10.1016/S1366-5545\(99\)00016-2](https://doi.org/10.1016/S1366-5545(99)00016-2)
- Daly, A., Hess, S., de Jong, G., 2012a. Calculating errors for measures derived from choice modelling estimates. *Transp. Res. Part B Methodol.* 46, 333–341. <https://doi.org/10.1016/j.trb.2011.10.008>
- Daly, A., Hess, S., Train, K., 2012. Assuring finite moments for willingness to pay in random coefficient models. *Transportation (Amst)*. 39, 19–31. <https://doi.org/10.1007/s11116-011-9331-3>
- Daly, A., Hess, S., Train, K., 2012b. Assuring finite moments for willingness to pay in random coefficient models. *Transportation (Amst)*. 39, 19–31. <https://doi.org/10.1007/s11116-011-9331-3>
- Danielis, R., Marcucci, E., 2007. Attribute cut-offs in freight service selection. *Transp. Res. Part E Logist. Transp. Rev.* 43, 506–515. <https://doi.org/10.1016/j.tre.2005.10.002>
- de Jong, G., 1990. An indirect utility model of car ownership and private car use. *Eur. Econ. Rev.* 34, 971–985. [https://doi.org/10.1016/0014-2921\(90\)90018-T](https://doi.org/10.1016/0014-2921(90)90018-T)
- de Jong, G., Gommers, M., Klooster, J., 2000. Time valuation in freight transport: methods and results, in: Ortúzar, J.D.D. (Ed.), *Stated Preferences Modelling Techniques*. PTRC Perspectives Series. pp. 231–242.
- de Jong, G., Johnson, D., 2009. Discrete mode and discrete or continuous shipment size choice in freight transport in Sweden. *Eur. Transp. Conf.* 2009 1–14.
- De Jong, G., Kouwenhoven, M., Bates, J., Koster, P., Verhoef, E., Tavasszy, L., Warffemius, P., 2014. New SP-values of time and reliability for freight transport in the Netherlands. *Transp. Res. Part E Logist. Transp. Rev.* 64, 71–87. <https://doi.org/10.1016/j.tre.2014.01.008>
- de Jong, G., Vierth, I., Tavasszy, L., Ben-Akiva, M., 2013. Recent developments in national and international freight transport models within Europe. *Transportation (Amst)*. 40, 347–371. <https://doi.org/10.1007/s11116-012-9422-9>
- Dubin, J.A., McFadden, D.L., 1984. An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica* 52, 345. <https://doi.org/10.2307/1911493>
- Eluru, N., Bhat, C.R., Pendyala, R.M., Konduri, K.C., 2010. A joint flexible econometric model system of household residential location and vehicle fleet composition/usage choices. *Transportation (Amst)*. 37, 603–626. <https://doi.org/10.1007/s11116-010-9271-3>
- Eluru, N., Chakour, V., Chamberlain, M., Miranda-Moreno, L.F., 2013. Modeling vehicle operating speed on urban roads in Montreal: A panel mixed ordered probit fractional split model. *Accid. Anal. Prev.* 59, 125–134. <https://doi.org/10.1016/j.aap.2013.05.016>
- Enam, A., Konduri, K.C., Eluru, N., Ravulaparthi, S., 2018. Relationship between well-being and daily time use of elderly: evidence from the disabilities and use of time survey. *Transportation (Amst)*. 45, 1783–1810.

- <https://doi.org/10.1007/s11116-017-9821-z>
- Feo, M., Espino, R., García, L., 2011. An stated preference analysis of Spanish freight forwarders modal choice on the south-west Europe Motorway of the Sea. *Transp. Policy* 18, 60–67. <https://doi.org/10.1016/j.tranpol.2010.05.009>
- Feo-Valero, M., García-Menéndez, L., Sáez-Carramolino, L., Furió-Pruñonosa, S., 2011. The importance of the inland leg of containerised maritime shipments: An analysis of modal choice determinants in Spain. *Transp. Res. Part E Logist. Transp. Rev.* 47, 446–460. <https://doi.org/10.1016/j.tre.2010.11.011>
- Ferdous, N., Pinjari, A.R., Bhat, C.R., Pendyala, R.M., 2010. A comprehensive analysis of household transportation expenditures relative to other goods and services : an application to United States consumer expenditure data. *Transportation (Amst)*. 37, 363–390. <https://doi.org/10.1007/s11116-010-9264-2>
- Gatta, V., Marcucci, E., 2016. Behavioural implications of non-linear effects on urban freight transport policies: The case of retailers and transport providers in Rome. *Case Stud. Transp. Policy* 4, 22–28. <https://doi.org/10.1016/j.cstp.2015.08.001>
- Gatta, V., Marcucci, E., Scaccia, L., 2015. On finite sample performance of confidence intervals methods for willingness to pay measures, *Transportation Research Part A* 82, 169-192.
- Hanemann, M., 1984. Discrete/Continuous Models of Consumer Demand. *Econometrica* 52, 541–561.
- Hess, S., Rose, J.M., Polak, J., 2010. Non-trading, lexicographic and inconsistent behaviour in stated choice data. *Transp. Res. Part D Transp. Environ.* 15, 405–417. <https://doi.org/10.1016/j.trd.2010.04.008>
- Hoffmann, P.J., 2003. Corridors of the Sea : An investigation into liner shipping connectivity Introduction : Transport networks and connectivity 257–270.
- Holguín-Veras, J., 2002. Revealed Preference Analysis of Commercial Vehicle Choice Process. *J. Transp. Eng.* 128, 336–346.
- Jäggi, B., Weis, C., Axhausen, K.W., 2013. Stated response and multiple discrete-continuous choice models: Analyses of residuals. *J. Choice Model.* 6, 44–59. <https://doi.org/10.1016/j.jocm.2013.04.005>
- Jian, S., Rashidi, T.H., Dixit, V., 2017. An analysis of carsharing vehicle choice and utilization patterns using multiple discrete-continuous extreme value (MDCEV) models. *Transp. Res. Part A Policy Pract.* 103, 362–376. <https://doi.org/10.1016/j.tra.2017.06.012>
- Khan, M., Machemehl, R., 2017a. Commercial vehicles time of day choice behavior in urban areas. *Transp. Res. Part A Policy Pract.* 102, 68–83. <https://doi.org/10.1016/j.tra.2016.08.024>
- Khan, M., Machemehl, R., 2017b. Analyzing tour chaining patterns of urban commercial vehicles. *Transp. Res. Part A Policy Pract.* 102, 84–97. <https://doi.org/10.1016/j.tra.2016.08.014>
- Kim, J., Allenby, G.M., Rossi, P.E., 2002. Modeling Consumer Demand for Variety. *Mark. Sci.* 21, 229–250. <https://doi.org/10.1287/mksc.21.3.229.143>
- Ko, J., Kim, D., 2017. Employer-based travel demand management program: Employer’s choice and effectiveness. *Transp. Policy* 59, 1–9. <https://doi.org/10.1016/j.tranpol.2017.06.003>
- Larranaga, A.M., Arellana, J., Senna, L.A., 2017. Encouraging intermodality: A stated preference analysis of freight mode choice in Rio Grande do Sul. *Transp. Res. Part A Policy Pract.* 102, 202–211. <https://doi.org/10.1016/j.tra.2016.10.028>
- Lee, J., Yasmin, S., Eluru, N., Abdel-Aty, M., Cai, Q., 2018. Analysis of crash proportion by vehicle type at traffic analysis zone level: A mixed fractional split multinomial logit modeling approach with spatial effects. *Accid. Anal. Prev.* 111, 12–22. <https://doi.org/10.1016/j.aap.2017.11.017>
- Lu, H., Hess, S., Daly, A., Rohr, C., 2017. Measuring the impact of alcohol multi-buy promotions on consumers’ purchase behaviour. *J. Choice Model.* 24, 75–95. <https://doi.org/10.1016/j.jocm.2016.05.001>
- Marcucci, E., Gatta, V., Scaccia, L., 2015. Urban freight, parking and pricing policies: An evaluation from a transport providers’ perspective. *Transp. Res. Part A Policy Pract.* 74, 239–249. <https://doi.org/10.1016/j.tra.2015.02.011>
- McFadden, D., Winston, C., Boersch-Supan, A., 1986. Joint estimation of freight transportation decisions under nonrandom sampling. *Anal. Stud. Transp. Econ.* 137–157. <https://doi.org/10.1017/cbo9780511895913.007>
- Nurul Habib, K.M., Miller, E.J., 2008. Modelling daily activity program generation considering within-day and day-to-day dynamics in activity-travel behaviour. *Transportation (Amst)*. 35, 467–484. <https://doi.org/10.1007/s11116-008-9166-8>
- Paleti, R., Copperman, R.B., Bhat, C.R., 2011. An empirical analysis of children’s after school out-of-home activity-location engagement patterns and time allocation. *Transportation (Amst)*. 38, 273–303. <https://doi.org/10.1007/s11116-010-9300-2>
- Papke, L.E., Wooldridge, J.M., 1996. Econometric Methods for Fractional Response Variables with an Application to 401(k) Plan Participation Rates. *J. Econom.* 11, 619–632.

- Pinjari, A.R., Bhat, C., 2011. An Efficient Forecasting Procedure for Kuhn-Tucker Consumer Demand Model Systems: Application to Residential Energy Consumption Analysis.
- Rashidi, T.H., Roorda, M.J., 2018. A business establishment fleet ownership and composition model. *Transportation (Amst)*. 45, 971–987. <https://doi.org/10.1007/s11116-017-9758-2>
- Regunaga, M., 2010. Implications of the Organization of the Commodity Production and Processing Industry: The Soybean Chain in Argentina. *LCSSD Occas. Pap. Ser. Food Prices* 1–32.
- Rich, J., Holmblad, P.M., Hansen, C.O., 2009. A weighted logit freight mode-choice model. *Transp. Res. Part E Logist. Transp. Rev.* 45, 1006–1019. <https://doi.org/10.1016/j.tre.2009.02.001>
- Richards, T.J., Gómez, M.I., Pofahl, G., 2012. A Multiple-discrete/Continuous Model of Price Promotion. *J. Retail.* 88, 206–225. <https://doi.org/10.1016/j.jretai.2012.01.002>
- Rose, J.M., Bliemer, M.C.J., 2009. Constructing efficient stated choice experimental designs. *Transp. Rev.* 29, 587–617. <https://doi.org/10.1080/01441640902827623>
- Schnepf, R.D., Dohlman, E., Bolling, C., 2001. Agriculture in Brazil and Argentina: Developments and Prospects for Major Field Crops, Market and Trade Economics Division, Economic Research Service, U.S. Department of Agriculture, Agriculture and Trade.
- Shin, J., Hong, J., Jeong, G., Lee, J., 2012. Impact of electric vehicles on existing car usage: A mixed multiple discrete-continuous extreme value model approach. *Transp. Res. Part D Transp. Environ.* 17, 138–144. <https://doi.org/10.1016/j.trd.2011.10.004>
- Shinghal, N., Fowkes, T., 2002. Freight mode choice and adaptive stated preferences. *Transp. Res. Part E Logist. Transp. Rev.* 38, 367–378. [https://doi.org/10.1016/S1366-5545\(02\)00012-1](https://doi.org/10.1016/S1366-5545(02)00012-1)
- Sikder, S., Pinjari, A.R., 2013. The benefits of allowing heteroscedastic stochastic distributions in multiple discrete-continuous choice models. *J. Choice Model.* 9, 39–56. <https://doi.org/10.1016/j.jocm.2013.12.003>
- Sivakumar, A., Bhat, C., 2002. Fractional Split-Distribution Model for Statewide Commodity-Flow Analysis. *Transp. Res. Rec.* 1790, 80–88. <https://doi.org/10.3141/1790-10>
- Spissu, E., Pinjari, A.R., Bhat, C.R., Pendyala, R.M., Axhausen, K.W., 2009. An analysis of weekly out-of-home discretionary activity participation and time-use behavior. *Transportation (Amst)*. 36, 483–510. <https://doi.org/10.1007/s11116-009-9200-5>
- Tanner, R., Bolduc, D., 2014. The Multiple Discrete-continuous Extreme Value Model (MDCEV) with Fixed Costs. *Procedia - Soc. Behav. Sci.* 111, 390–399. <https://doi.org/10.1016/j.sbspro.2014.01.072>
- Tapia, R.J., dos Santos Senna, L.A., Larranaga, A.M., Cybis, H.B.B., 2019. Joint mode and port choice for soy production in Buenos Aires province, Argentina. *Transp. Res. Part E Logist. Transp. Rev.* 121, 100–118. <https://doi.org/10.1016/j.tre.2018.04.010>
- Tavasszy, L., de Jong, G., 2014. *Modelling Freight Transport, Modelling Freight Transport*. Elsevier Inc.
- Train, K., 1999. Halton Sequences for Mixed Logit. *Work. Pap.* 1–18.
- Vellay, C., de Jong, G., 2003. A simultaneous SP/RP Analysis of Mode Choice in Freight Transport in the Region Nord-Pas-de-Calais.
- Woo, J.R., Choi, J.Y., Shin, J., Lee, J., 2014. The effect of new media on consumer media usage: An empirical study in South Korea. *Technol. Forecast. Soc. Change* 89, 3–11. <https://doi.org/10.1016/j.techfore.2014.09.001>
- Yu, B., Zhang, J., 2015. Modeling household energy consumption behavior: A comparative analysis. *Transp. Res. Part D Transp. Environ.* 39, 126–140. <https://doi.org/10.1016/j.trd.2015.07.001>
- Yu, H., Zeng, A.Z., Zhao, L., 2009. Single or dual sourcing: decision-making in the presence of supply chain disruption risks. *Omega* 37, 788–800. <https://doi.org/10.1016/j.omega.2008.05.006>
- Zamparini, L., Layaa, J., Dullaert, W., 2011. Monetary values of freight transport quality attributes: A sample of Tanzanian firms. *J. Transp. Geogr.* 19, 1222–1234. <https://doi.org/10.1016/j.jtrangeo.2011.01.002>

## Appendix I

The values used for the experimental design were coded from 0 to 2 with the values shown in Table I.1 for mode variables and Table I.2 for port variables. The values that were sent to the interviewees were customized using the actual values for the closest port. Table I.3 shows the design of each of the scenarios for the train alternatives (A and C) and Table I.4 shows them for



truck alternatives (B and D). It should be noted that the truck option was used as the reference alternative for mode and the closest port as reference for FAS price.

Table I.1

	Freight Price for train	Freight Price for truck	Travel Time for rail	Travel Time for truck	Lead Time for rail	Lead Time for truck	Reliability for rail	Reliability for Truck	Minimum shipment size for rail (tn)	Minimum shipment size for truck (tn)
<b>0</b>	67.5% of current Truck price	100% of truck price	90% of travel time at 30km/h	100% of travel time at 65km/h	5 days	0.5 days	0 days	0 days	500	32
<b>1</b>	75% of current price	100% of truck price	100% of travel time at 30km/h	100% of travel time at 65km/h	7 days	0.5 days	0 - 1 days	0 days	1000	32
<b>2</b>	82.5% of current price	100% of truck price	110% of travel time at 30km/h	100% of travel time at 65km/h	10 days	0.5 days	0 - 3 days	0 days	1500	32

Table I.2

	FAS price for closest port	FAS price for further port
<b>0</b>	0% less of the current price in Rosario	5% more of the current price in Rosario
<b>1</b>	0% less of the current price in Rosario	2,5% more of the current price in Rosario
<b>2</b>	0% less of the current price in Rosario	0% more of the current price in Rosario

Table I.3

Choice situation	A						C					
	Lead Time	Time	Freight Price	FAS price	Shipment Size	Reliability	Lead Time	Time	Freight Price	FAS price	Shipment Size	Reliability
<b>Example</b>	1	1	1	1	1	2	2	1	1	1	1	1
<b>1</b>	0	2	2	0	0	1	2	0	0	1	1	1
<b>2</b>	1	0	2	2	0	2	1	2	0	1	1	0
<b>3</b>	0	0	2	0	1	1	2	2	0	1	0	1
<b>4</b>	2	1	2	1	0	0	0	1	0	1	1	2
<b>5</b>	2	1	1	0	1	0	0	1	1	1	0	2
<b>6</b>	1	0	1	2	1	1	1	2	1	1	0	0
<b>7</b>	1	2	1	2	0	0	1	0	1	1	1	2
<b>8</b>	1	2	0	1	0	2	1	0	2	1	1	0
<b>9</b>	2	0	0	0	0	2	0	2	2	1	1	0
<b>10</b>	0	1	1	2	1	1	2	1	1	1	0	1
<b>11</b>	0	2	0	1	1	2	2	0	2	1	0	1
<b>12</b>	2	1	0	1	1	0	0	1	2	1	0	2

