

Urban freight distribution policies: joint accounting of non-linear attribute effects and discrete mixture heterogeneity

Valerio Gatta, DIPES/CREI, University of Roma Tre¹, Rome Italy

Edoardo Marcucci, DIPES/CREI, University of Roma Tre, Rome Italy

ABSTRACT

This paper jointly addresses non-linear attribute effects and discrete mixture heterogeneity in the case of urban freight transport policy by adopting an agent-specific perspective.

The innovative contributions are threefold.

The first relates to the *data set* specifically collected during a research project conducted for Volvo Research Foundation in 2009.

The second refers to the agent-specific perspective adopted in the paper. In particular, it investigates, via a stated ranking experiment, the preferences transport providers have for different transport policies.

Thirdly the policy relevant issues are addressed in an innovative way. The paper, in fact, tests for *non-linear effects* in attribute level variations to evaluate the implications on willingness to pay measures and also investigates for the presence of *heterogeneity among transport providers* by estimating a latent class model on the assumption that the heterogeneity can aptly be summarized via discrete mixtures.

The results obtained underline the relevance of non-linear effects for our sample along with the presence of two separate classes of agents that have distinguishably different preferences for the policies tested. The presence of non-linear effects suggests policy makers should carefully consider the effects induced by the specific *status quo* level for the policy relevant attribute variation and, at the same time, assume that there will be a differentiated reaction also within the transport provider category.

In conclusion the paper underlines the need for a sophisticated *ex-ante* policy analysis if the correct policy outcomes are to be estimated with an adequate level of accuracy.

Keywords: Urban freight distribution, discrete mixture heterogeneity, non-linear attribute effects, Limited Traffic Zone, transport providers

¹ Corresponding Author: valerio.gatta@uniroma3.it

1. Introduction

Urban freight distribution (UFT) is progressively acquiring a more prominent position in policy makers' agendas throughout the world. The temporal and spatial overlapping of freight distribution operations with peaking passenger demand is duly recognized and the relevance of the problem for modern cities is adequately acknowledged. The EU is dedicating greater attention to the problem also by funding specific research projects (BESTUFS I, 2000; BESTUFS II, 2004; CUPID, 2000; EFRUD, 2009; EUTP II, 2000; MOST, 2000; PROGRESS, 2000; OSSA, 2000; REVEAL, 1999; STRAIGHTSOL, 2012; SULOGRA, 2000). The different and often contrasting interests of the various stakeholders involved need to be duly considered and the interaction among them explicitly modeled. Data needs are often superior to their availability (Samimi et al., 2010). In the EU there is no systematic and specific on-going urban freight activity survey. Ruesch and Glücker (2001), after inquiring 43 medium sized cities in Europe, discovered that 58% of them were not collecting data on UFT. Policy makers should be able to evaluate *ex-ante* the effects of the policy they intend to implement. The analysis of specific agents' preferences for UFT policies has been structurally under-researched. This is especially due to the lack of appropriate data, notwithstanding some noticeable exceptions (Marcucci and Gatta, 2013a; Marcucci and Gatta, 2013b), even if it has been, for long, suggested to study specific actors in UFT operations (Ogden, 1992).

The contribution of this paper is three-fold.

The acquisition of data concerning specific transport providers' preferences concerning possible intervention policies that could be implemented by local policy makers in the LTZ in Rome represents a first contribution. Other papers have documented the articulated and thorough selection process of the policy attributes finally considered in this paper (see Stathopoulos et al., 2011; 2012).

Furthermore, the paper tests the potential non-linear effects of attributes' levels variations, usually assumed homogeneous, which could have strong policy implications given the distortions the assumption might cause (see, for example, Marcucci and Gatta, 2013a).

Finally, the paper tests for the presence of discrete mixture heterogeneity and discusses its implications for willingness to pay (WTP) measures.

The conclusions reached underline the need of greater accuracy and sophistication in defining UFT policies.

The paper is structured as follows: section 2 reports a short literature review that helps contextualizing the innovative contributions of the paper. The methodology is discussed in section 3 while section 4 provides a succinct data description. The results are illustrated in section 5 together with policy implications. Section 6 concludes and illustrates future research endeavors.

2. Literature review

This section reports a brief literature review of the evolution of the logistic sector and its repercussions on UFT as well as on the most commonly used modeling approaches in this sector.

2.1. Logistic definition, evolution and supply chain management

The concept of logistics originates from military operations and there is no common and generally accepted definition (Ballou, 1999; Logistics World, 2007; Rushton et al., 2000). The growing number of interactions among supply chain agents, the continuous development of technology, the enlargement of firms' structures and the increasing role of tailor-made services offered have all induced a progressive sophistication in business logistics. This implies a greater difficulty in precisely defining the exact boundaries of this industrial sector. Key logistics components usually include: inventory, information, warehousing, material handling, packaging, network design, customer service and transport (Bowersox and Closs, 1996). In the '60s the formulation and planning of the various logistics components were *de facto* completely neglected. Greater attention to this sector aroused in the '80s when the distribution costs became an issue. Different measures to reduce inventory time were adopted in the early '90s and the use of third-party companies blossomed. In the late '90s it became clear that, even for a single transaction, more than one actor could be involved in product distribution. The traditional logistic concept had to be extended and the whole logistical system had to be interpreted as a network of organizations co-operating through/across logistic chains pertaining to different processes and activities. The concept of supply chain management (SCM) had come to light. This new conceptualization extended the boundary of freight transport systems now covering different business sectors including manufacturing, warehousing, marketing, sales, and information technology (Christopher, 1998). Logistics and SCM have progressively acquired a greater role in both private and public sectors. An in depth analysis of the interactions between the agents involved is fundamental to gain a clear insight of the functioning of the current transport system and to evaluate ex-ante the possible effects of intervention policies (Arunotayanun, 2009). The most important supply chain members are: 1) manufacturers, 2) shippers, 3) carriers, 4) freight forwarders, 5) third-party logistics providers, 6) customers, 7) suppliers. In present supply chains, even for a single transaction, it is quite customary to have more than two companies involved and the functional relationship can become quite complex especially when a third party logistic provider is involved. SCM is multifaceted task and, therefore, it is crucial to identify: 1) key and supporting members, 2) processes assigned to each member, 3) the level of integration and management characterizing each process link (Lambert, 2001). This paper contributes to this articulated field of research (Russo and Comi, 2011; Filippi et al., 2010) by providing and in-depth analysis of transport providers' preferences, a key actor in the SCM context, for UFT policies².

2.2. UFT modeling approaches

UFT literature analysis reveals a substantial heterogeneity in the approaches adopted (Marcucci and Gatta, 2013; Gentile and Vigo, 2007, Comi et al., 2012). The main and most evident distinction among the various approaches adopted relates to the public or private perspective considered. In fact, the public approach has mainly focused on the definition and implementation of policies aimed at reducing the negative external effects UFT imposes on

² Transport providers, in fact, play a crucial role in a widely adopted intervention policy in Europe. The success or failure of Urban Freight Distribution initiatives is usually linked to the reaction of transport providers to the implementation characteristics' of this policy (see, for instance, Marcucci and Danielis, 2008; Paglione and Gatta 2007; Danielis and Marcucci, 2007; Marcucci et al., 2007).

cities. The private approach, instead, mainly focuses on enhancing the efficiency of business operations (Corò and Marcucci 2001; Marcucci and D'Agostino, 2003). Recent modeling advances attempt to develop hybrid models. These attempts try to include supply chain elements and considerations in public decision-making models. The intent is providing a balanced account of both public and private objectives. Preferences towards policy interventions are elicited in a transparent way and explicitly considered (Roorda, 2010). Stakeholders' decision making would benefit from appropriate evaluative systems based on good quality modeling.

Not only there is heterogeneity in UFT modeling approaches but there is also heterogeneity in the classification adopted in review papers (e.g. Boerkamps and Binsbergen, 1999; Regan and Garrido, 2001; Taniguchi et al., 2003; Groothedde, 2005; Yang et al., 2010). The criteria usually adopted for categorizing the approaches prevalently rely on the modeling principles used. In other words, modeling approaches are partitioned according to aggregate versus disaggregate models, simulation of commodity versus vehicle movements, systemic and operational models (Paglione, 2006).

The approach adopted in this paper refers to a well-established tradition in behavioral and disaggregate freight modeling. Winston (1983; 1985) presents an overview of different freight models and describes the early evolution from aggregate to disaggregate models. The first group of models were used to forecast the behavior of an entire transport system while the second was used to predict the behavior of individual agents within a given transport system. The use of disaggregate models quickly developed thanks to their theoretical and empirical advantages. In fact, disaggregate models rest on individual behavior analysis, provide a richer model specification capable of capturing important decision-makers' characteristics and warrant a better understanding of the effects of UFT policies. This wealth of benefits comes at the cost of relevant data needs and high accuracy in defining the relevant attributes to be considered. Previous works testify how important and demanding the data acquisition process can be (Stathopoulos et al., 2011; 2012).

This paper adopts a disaggregate behavioral approach at the agent-type level since it is the most appropriate and neglects the interaction among the various agent-types collaborating/competing along the supply chain. The empirical analysis of the interaction effects is postponed to a subsequent paper specifically targeted to this issue.

3. Methodology

This section describes the methodology employed. In particular: 3.1 deals with random utility maximization (RUM) in general and multinomial logit (MNL) and latent class (LC) in particular; 3.2 illustrates Classification and Regression Trees (CART) and its application to LC characterization; 3.3 discusses the main characteristics and properties of the Krinsky-Robb (KR) method to evaluate WTP standard deviation measures; 3.4 reports the essentials of the Wilcoxon test.

3.1. RUM, MNL and LC

Discrete choice modeling when adopting a RUM principle assumes that an agent's (i) indirect – latent – utility function (U) for a choice alternative (j) is composed of a systematic or observable part (V) and an unobservable and random one (ϵ). One can write the indirect utility function of agent i for alternative j as follows:

$$U_i(j) = V_{ij} + \varepsilon_{ij} \quad (1)$$

where $V_{ij} = \beta'x_{ij}$

McFadden developed the MNL model in the early '70s (McFadden, 1974). MNL is characterized by interesting and much appreciated advantages (*e.g.* closed form, ease of interpretation, etc.) as well as by relevant drawbacks linked to the assumption of, across respondents, preference homogeneity. The estimated parameters represent the marginal utility of each attribute variation (even if confounded for the scale) assuming all agents have an equal taste for a given attribute (Marcucci, 2005).

Research, induced by the concerns for MNL drawbacks, has developed more flexible and sophisticated ways to treat preference heterogeneity. LC is one of the results of these efforts. LC incorporates preference heterogeneity via the systematic component of utility (Marcucci and Gatta, 2012). It investigates heterogeneity assuming a discrete mixing distribution of preference parameters where a small number of mass points (c) are interpreted as different groups/segments of agents (Kamakura and Russell, 1989; Boxall and Adamowicz, 2002).

The probability of choosing alternative j for agent i at time t is the expected value, over classes, of the choice probability within each class and is analytically reported below.

$$\text{Prob}(y_{it} = j) = P(j|c)P(c) = \frac{\exp(\beta'_c x_{jit})}{\sum_{q=1}^J \exp(\beta'_c x_{qit})} \sum_{c=1}^C \left(\frac{\exp(\varphi'_c k_i)}{\sum_{c=1}^C \exp(\varphi'_c k_i)} \right) \quad (2)$$

The class probabilities can also be functions of socio-economic variables (k_i).

3.2. CART

CART is a classification method that, in general, uses historical data to construct so-called decision trees, which are subsequently used to classify new data. In order to use CART one needs to know the number of classes *a priori* as it is in our case. The methodology was developed in the '80s (Breiman et al., 1984). The decision trees are developed via a learning sample that is a set of historical data with pre-assigned classes. Decision trees are represented by a set of questions, which split the learning sample into progressively smaller parts. The CART algorithm will search among all variables investigated the variable producing the best split of the data into two parts with maximum internal homogeneity. The process is iterated for each of the resulting data fragments. CART can handle both numerical and categorical variables, is robust to outliers (the splitting algorithm usually isolate outliers in individual nodes), and produces classification or regression structures that are invariant with respect to monotone transformations of independent variables. CART consists of three parts: 1) creation of the maximum tree, 2) choice of the right tree size, 3) classification of new data using the constructed tree.

The construction of the maximum tree is the most time consuming and daunting task. In fact, building the maximum tree implies splitting the learning sample until the terminal nodes contain observations of only one class. Splitting algorithms are different for classification and regression trees. Only the construction of classification trees is discussed since this alone will subsequently be used in the paper.

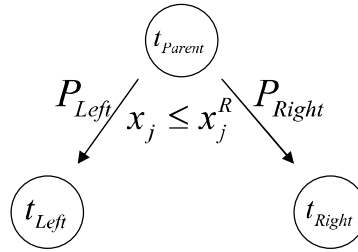
Classification trees are used when for each observation of the learning sample the number of classes is known in advance. Classes in the learning sample may be provided by the user or calculated in accordance with some exogenous rule. The classification tree is built in accordance to a *splitting rule*. The Gini splitting rule is the most frequently used but the Twoing one is also employed. The results reported are based on the Gini splitting function but we also tested Twoing's rule without detecting any major differences.

The Gini splitting rule uses the following impurity function:

$$i(t) = \sum_{k \neq l} p(k|t) p(l|t) \quad (3)$$

where k and l index the class; $p(k|t)$ represents the conditional probability of class k at node t . Let t_p be a parent node; t_l and t_r are respectively left and right child nodes of parent node t_p .

Figure 1 – Representation of parent and child nodes in CART



Maximum homogeneity of child nodes is defined by the impurity function, whose variations are defined as follows:

$$\Delta i(t) = i(t_p) - P_l i(t_l) - P_r i(t_r) \quad (4)$$

where P_l and P_r are the probabilities of left and right node. The maximization problem to be solved, at each node, is the following:

$$\arg \max_{x_i \leq x_j^R, j=1, \dots, M} [i(t_p) - P_l i(t_l) - P_r i(t_r)] \quad (5)$$

Applying the Gini rule to the maximization problem we get equation that measures the change of the impurity measure:

$$\Delta i(t) = \sum_{k=1}^K p^2(k|t_p) + P_l \sum_{k=1}^K p^2(k|t_l) + P_r \sum_{k=1}^K p^2(k|t_r) \quad (6)$$

Therefore the Gini algorithm solves the following problem:

$$\arg \max_{x_i \leq x_j^R, j=1, \dots, M} [-\sum_{k=1}^K p^2(k|t_p) + P_l \sum_{k=1}^K p^2(k|t_l) + P_r \sum_{k=1}^K p^2(k|t_r)] \quad (7)$$

The Gini algorithm searches within the learning sample for the largest class, isolate it from the rest and works well for noisy data. The Twoing splitting algorithm searches for two classes that will make up together more than 50% of the data.

3.3. KR

The KR method (Krinsky and Robb, 1986; 1990), also known as the parametric bootstrap (Efron and Tibshirani, 1993), adopts a simulative approach to the estimation of confidence intervals for WTP measures. In fact, estimates are produced by taking a large number of draws from a multivariate normal distribution with means calculated by the estimated coefficients and covariance given by the estimated covariance matrix of the coefficients (Hole, 2007). Subsequently r simulated WTP values, calculated by taking r draws from the joint distribution of the coefficients, are used to calculate the percentiles of the simulated distribution reproducing the desired level of confidence. If, for instance, 1,000 simulated WTP values are estimated, the lower and upper limits of a 95% confidence interval, also called percentile intervals (Mooney and Duval, 1993; Efron and Tibshirani, 1993), are respectively given by the 25th and 975th sorted WTP estimates.

On the assumption that WTP is symmetrically distributed, similarly to the delta method³, the confidence interval could also be derived by using the draws to calculate the variance of WTP and then plugging it into the standard equation to measure the confidence interval. This approach, however, would not exploit the relaxation of the normal distribution assumption that the KR method circumvents and should, therefore, be avoided. The only assumption required, in fact, is that the coefficients are jointly normally distributed, which is a realistic assumption with a relatively large sample. The KR method should, in principle, produce more accurate confidence intervals with non-symmetrically distributed WTP measures.

3.4. Wilcoxon test

The Wilcoxon test also known as rank sum test is used to test for a difference between two samples and represents the nonparametric counterpart to the two-sample Z or t tests (Wilcoxon, 1945). Instead of comparing two population means, the test compares two population medians. The Wilcoxon rank sum test of $H_0: \eta_1 = \eta_2$ (medians) looks at how the ranks of the combined samples are grouped into the separate samples, to see if one sample gets more than its share of larger or smaller ranks. The test statistic is developed assuming that $n_1 \leq n_2$ and letting $R_1 = \text{rank of } X_1 \text{ among } \{X_1, X_2, \dots, X_{n_1}, Y_1, Y_2, \dots, Y_{n_2}\}$, $R_2 = \text{rank of } X_2 \text{ among } \{X_1, X_2, \dots, X_{n_1}, Y_1, Y_2, \dots, Y_{n_2}\}, \dots, R_{n_1} = \text{rank of } X_{n_1} \text{ among } \{X_1, X_2, \dots, X_{n_1}, Y_1, Y_2, \dots, Y_{n_2}\}$. The largest gets rank $n_1 + n_2$ while the smallest gets rank 1. Then the Wilcoxon rank sum test statistic is the sum of the ranks for the smaller sample:

$$W = \sum_{i=1}^{n_1} R_i \quad (8)$$

The H_0 -distribution of W is found by using the exact distribution of W . The exact mean and variance of W are, respectively:

$$E_{H_0}(W) = \frac{1}{2}n_1(n_1 + n_2 + 1); \quad Var_{H_0}(W) = \frac{1}{12}n_1n_2(n_1 + n_2 + 1) \quad (9)$$

The normal approximation to distribution of W is:

$$N\left(\frac{1}{2}n_1(n_1 + n_2 + 1), \frac{1}{12}n_1n_2(n_1 + n_2 + 1)\right) \quad (10)$$

The p -value can be calculated using either the exact distribution of W or the normal approximation to the distribution of W . As usually, one should interpret lower p -value as stronger evidence against H_0 .

³ For a description of the Delta Method please refer to Oehlert (1992).

4. Survey instrument and data description

The data used in this paper were acquired in Rome's LTZ between March and December 2009 (VRF, 2009). The LTZ in the city center of Rome, first implemented in the late '80s, encompasses a 5km² area originally banned to non-residents' vehicles only. Nowadays, Euro1 and more fuel-efficient vehicles only can enter the LTZ. Residents are granted free access while others (e.g. transport providers) pay an entrance fee. Enforcement relies on cameras and optical character recognition software; the system operates diurnally with a yearly entrance fee of 565€ per number plate.

An extensive list of prohibitions generically applies to agents while a wide ranging of *ad hoc* exemptions pertains to third party freight operators. The regulation, after detailed examination and careful interpretation of the exemptions conceded, is mostly aimed at dissuading own-account operators' entrance.

The questionnaire was progressively developed starting from attribute definition. A focus group with relevant stakeholders was conducted to define the most relevant problems, determine appropriate policy intervention and forecast likely stakeholders' reactions (Stathopoulos et al., 2011). The experimental design was developed using a d-efficient approach using Ngene 1.1 (www-choice-metrics.com). We opted for a stated ranking exercise (SRE) rather than stated choice since this response format was considered most appropriate given the aim of the research and study context (Marcucci et al., 2012). In fact, we were interested in unveiling agents' preferences concerning UFT policies which cannot *de facto* be "chosen" thus inducing us to ask interviewees to rank the policies considered.

The alternatives used in the SRE are described by a set of attributes taking several levels. The attributes selection process was based on: 1) literature survey; 2) previous UFT studies performed in Rome; 3) focus groups with experts. A literature review permitted the identification of a set of eligible attributes that embodied potentially conflicting policy instruments⁴.

The attribute selection process was based on previous UFT studies carried out in Rome (STA, 2001; Filippi and Campagna, 2008) and focus groups with experts/stakeholders⁵. The attributes selected were considered highly important by at least two of the stakeholders categories contacted (Stathopoulos et al., 2011) and were also validated via a pre-test with real operators. Table 1 reports attributes, levels, and ranges. Attributes are characterized by, at least, three levels thus allowing the test of non-linear effects. All the attributes used are contemplated both as possible levers of intervention by decision-makers and perceived as apt measures by stakeholders.

Table 1 - Attribute levels and ranges used in the SRE

Attribute	Number of levels	Level and range of attribute - (<i>Status Quo</i> underscored)
Loading/unloading bays:	3	<u>400</u> , 800, 1200
Probability of free l/u bays:	3	<u>10%</u> , 20%, 30%
Fees:	5	200€, 400€, <u>600€</u> , 800€, 1000€

⁴ Night-time deliveries were excluded since they are considered efficiency enhancing by carriers but are opposed by retailers that see them as a mere increase in costs.

⁵ Expert surveys focused on the definition of the most appropriate policies to mitigate the identified UFT problems (Stathopoulos et al., 2011).

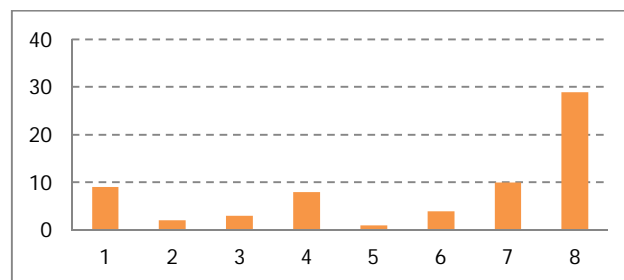
A SRE is adopted to test the effects of currently unavailable options. The alternatives presented to respondents include two policy options plus the *status quo* alternative. Table 2 reports a SRE sample task.

Table 2 - Example of a ranking task

	Policy 1	Policy 2	Status Quo
Loading/Unloading bays	400	800	400
Probability to find L/U bays free	20%	10%	10%
Entrance fee	1000 €	200 €	600 €
Policy ranking	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

In total, 252 interviews were completed and 229 used⁶. We sampled 66 transport providers scattered in 8 main macro-freight sectors (see Figure 2), namely: 1) *food* (fresh, canned, drinks, tobacco, bars, hotels and restaurants); 2) *personal and house hygiene* (detergents, pharmaceuticals, cosmetics, perfumes, watches, barbers, etc.); 3) *stationery* (e.g. paper, newspapers, toys, books, CDs etc.); 4) *house accessories* (e.g. dish washers, computers, telephones, metal products etc.); 5) *services* (e.g. laundry, flowers, live animals, accessories and animal food, etc.); 6) *clothing* (cloth, leather, etc.); 7) *construction* (e.g. cement, scaffold, chemical products, etc.); 8) *cargo* (general cargo).

Figure 2 – Transport provider agent distribution by main freight sector



5. Econometric results and policy implications

This section reports the results of the models estimated for transport providers, reviews the WTP derived and discusses their policy implications.

⁶ Pilot interviews were not used for estimation purposes.

5.1. Model results

We preliminary tested the potentially non-linear effects of the variations of the attributes considered⁷. The first model estimated (M1) adopts a MNL specification of the utility function; all attributes are linear and normalized. The results for (M1), reported in Table 3, are in line with expectations both for variables' signs and for their relative weight. In fact, EF, with a negative sign, is the variable with the highest explanatory power while both LUB and PLUBF, with approximately half EF weight, have a positive sign and an almost equal impact on utility. It is important to underline the aversion towards the *status quo* situation testified by the positive sign of the two alternative specific constants (ASCs) for the unlabeled alternatives which cannot be considered different.

Table 3 – Model estimates for MNL with linear effects

VARIABLE	COEFFICIENT	ST.ERR.	T-STAT
LUB	0.558	0.061	9.16
PLUBF	0.435	0.069	6.31
EF	-1.170	0.069	-16.85
ASC_Alt1	0.686	0.173	3.97
ASC_Alt2	0.709	0.159	4.46

The second model (M2), reported in Table 4, investigates potential non-linear effects by effects-coding the explanatory variables. The test could be performed since the original choice design foresaw three levels for LUB and PLUBF and five for EF. The log-likelihood ratio test compared the model fit of the two MNL models (restricted vs. unrestricted). (M2) fitted the data substantially better than (M1) signaling the presence and statistical relevance of *non-linear effects* when moving from one attribute level to another⁸. Equally scaled variations in attribute levels have, thus, heterogeneous effects on the ranking probability of a given policy change. (M2), the best fitting parsimonious model, uses effects coding for LUB and EF while keeping PLUBF linear and normalized⁹. In fact, a MNL model, not reported for reasons of brevity, where all the variables were effects coded, shows that PLUBF2 has a coefficient close to zero (0.026) implying a linear effect of the attribute¹⁰. A preliminary graphical analysis suggests the presence of non-linear effects for both LUB and EF only (see Figure 3).

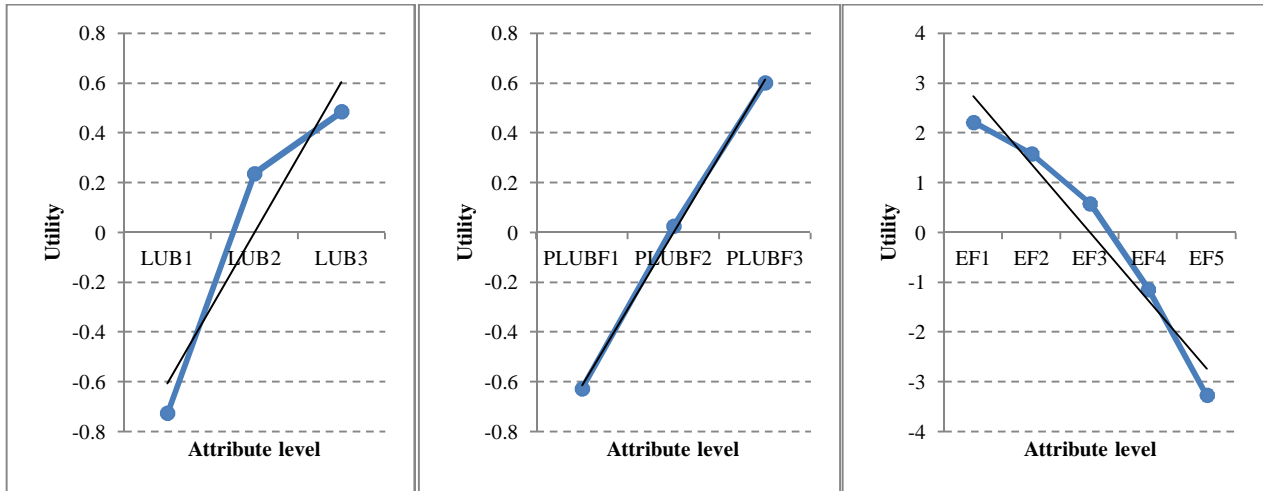
⁷ The attributes finally included in the experimental design developed for transport providers were coded as follows: LUB = loading and unloading bay; PLUBF = probability of finding a load and unloading bay free; EF = entrance fee. The numbers to the right of the acronym of the attribute relate to the specific level of the variable considered (e.g. LUB1 = first level of the loading and unloading bay attribute).

⁸ (M2) can be considered a better fitting model in comparison to (M1) on the basis of a LL ratio test. In fact (M1) has an adjusted rho-sq of 0.251 and (M2) has an adjusted rho-sq of 0.281.

⁹ We also estimated an Error Component model that produced no interesting results. No random effects linked to the alternatives were detected. The result can intuitively be attributed to the non-labeled nature of the ranking exercise performed.

¹⁰ This result should not be interpreted as non-relevant for the impact this level of the variable has on the utility of the decision-maker. In fact, it would be erroneous to affirm that moving from a base of 10 probability points of finding a loading and unloading bay free to a probability of 20 points has no effect on the utility of the decision-maker. In this case the fact that PLUBF1 = -0.627 implies that the impact of moving from the first level to the second can be calculated as follows $PLUBF1 \rightarrow 2 = 0.026 - (-0.627) = 0.653$. A further check of the correct interpretation of the effects coded attribute can be performed by re-estimating the model after having dummy coded the variables so to discover that the PLUBF2 coefficient is, in that case, statistically significant.

Figure 3 – Part-worth utilities for different attributes



Subsequent tests performed by re-estimating the same model adopting a selective linear normalization of the variables confirms the intuitions induced by direct visual inspection.

Table 4 – Model estimates for MNL with non-linear effects

VARIABLE	COEFFICIENT	ST.ERR.	T-STAT
LUB2	0.238	0.077	3.07
LUB3	0.491	0.077	6.40
PLUBF	0.613	0.076	8.05
EF1	2.220	0.146	15.19
EF2	1.586	0.116	13.62
EF4	-1.131	0.109	-10.40
EF5	-3.260	0.224	-14.56
ASC_Alt1	1.041	0.212	4.92
ASC_Alt2	0.972	0.184	5.28

The signs of the attributes are in line with expectations and the coefficients of the attribute levels considered are statistically significant. EF, given the range considered, is the attribute with the greatest impact. The previously detected aversion to the *status quo* is confirmed while the other two hypothetical policies considered have no distinct impact *per se*. Effects coding the variables facilitates interpretation. In fact, the constant term can only reflect the utility associated with the base case alternative thus avoiding any possible misinterpretation.

Furthermore, the paper investigates the presence of heterogeneity in preferences via a discrete mixture model that relaxes the restrictive independence of irrelevant alternatives

(IIA) assumption characterizing the stochastic part of utility in a MNL specification¹¹. To facilitate model comparison, using (M1) and (M2) as a reference, two LC models ((M3) and (M4)) are estimated coding the variables in the same way as in the MNL specification. Both LC models were specified and estimated using two classes notwithstanding only 66 transport providers were interviewed. The models suggest the presence of two separate latent classes of transport providers each characterized by rather differentiated preferences. Furthermore, it is important to underscore that the results obtained are quite different from the MNL models.

When adopting a discrete mixture approach for detecting preference heterogeneity one crucial issue relates to the number of classes to be estimated (Bujosa et al., 2010). The conventional specification tests cannot be used to determine the optimal number of classes. In particular, the likelihood ratio or Wald tests cannot be employed. In fact, within this context, these tests do not satisfy the regularity conditions for a limiting chi-square distribution under the null hypothesis adopted. The optimal number of classes to be used should thus be determined on the basis of some information criteria statistics such as those developed by Hurvich and Tsai (1989). The entropy index has also been commonly employed when choosing the number of segments to be estimated (Thacher et al., 2005). Notwithstanding the relevance and pertinence of the considerations reported other issues to be considered relate to the: 1) significance of the estimated parameters, 2) plausibility of model results (e.g. signs and magnitudes of the parameters) and 3) *a priori* information concerning existent groups. The statistical analysis of our data shows that transport providers can be split in two groups with a clearly differentiated behavioral profile. If more than two classes are included in the model the intelligibility of the results is reduced, the parameters cannot easily be associated to a specific behavior and the additional groups represent only a small portion of the total respondents.

Table 5 – Model estimates for LC with linear effects

	CLASS 1		CLASS 2	
VARIABLE	COEFFICIENT	T-STAT	COEFFICIENT	T-STAT
LUB	0.545	4.90	0.920	13.67
PLUBF	0.203	1.54	0.966	10.67
EF	-2.271	-13.55	-0.747	-10.55
ASC_Alt1	0.839	2.94	0.847	3.30
ASC_Alt2	0.740	2.99	0.869	3.63

Looking at the results of (M3) and (M4) similar conclusions can be drawn (see Table 5 and 6). EF has, in general, a marked impact. Class 1 (C1) comprises more price-sensitive transport providers while class 2 (C2) includes those agents more interested in bays based policies implemented either via construction or by increasing the probability of finding them free. Moreover, the estimated latent class probabilities are almost equal in both models. As for the MNL specification, the *status quo* is perceived negatively.

¹¹ Heterogeneity in preferences is important in itself. One should, however, recall that the hypotheses and procedures used to search for it have also an impact on the final results. See Marcucci and Gatta (2012) on this point.

Table 6 – Model estimates for LC with non-linear effects

VARIABLE	CLASS 1		CLASS 2	
	COEFFICIENT	T-STAT	COEFFICIENT	T-STAT
LUB2	0.236	1.63	0.339	3.64
LUB3	0.405	3.03	0.982	10.37
PLUBF	0.480	3.61	1.277	10.23
EF1	4.105	12.60	1.724	9.43
EF2	3.003	11.53	1.138	8.92
EF4	-2.204	-8.56	-0.363	-2.71
EF5	-6.079	-9.39	-2.818	-12.57
ASC_Alt1	1.534	4.15	0.735	2.24
ASC_Alt2	1.359	4.20	0.770	2.61

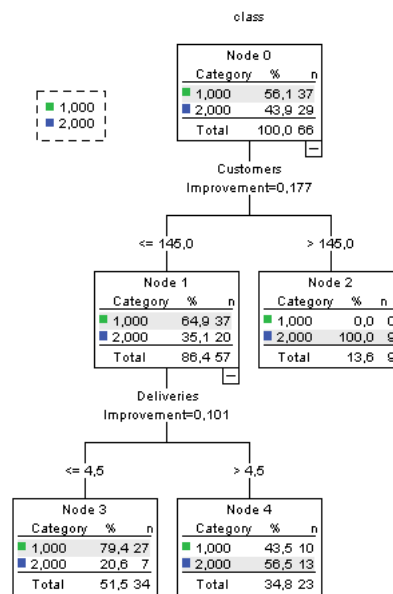
Furthermore, we opted for a CART model to aid analysis after testing, without success, the use of various socio-economic variables to further specify class probability and facilitate the interpretation of results from a policy perspective. CART analysis, in fact, helps allocating the observations to subsequently more homogeneous groups making use of different exogenous variables (*i.e.* socio-economic). It is important to underline the conceptual difference in the use made of the exogenous variables in the two different approaches. In fact, the introduction of socio-economic variables in the LC model has an *inferential* purpose aimed at explaining the class membership of the companies in the sample. In the case of CART, socio-economic variables are used with a *descriptive* intent and represent, by definition, the discriminating elements when separating the observations in the two groups. Furthermore, in our opinion, the inability of the socio-economic variables in explaining class probability is mostly due to the limited number of observations rather than to the intrinsic capability of the variables in explaining class membership. As a concluding note, justifying and motivating our methodological choice in using CART, we underline that the multivariate method proposed represents an improvement with respect to the commonly used univariate approach.

The largest conditional probability was used to assign a given agent to a specific latent class. The 80%-20% class probability combination was used to assign observations to classes. Only two were the controversial cases encountered.

The socio-economic variables used for segmentation are: 1) number of customers in the LTZ (*Customers*), 2) number of daily deliveries (*Deliveries*), 3) number of employees (*Employees*), 4) number of vehicles used for deliveries within the LTZ (*Vehicles-LTZ*), 5) freight sector (*Sector*), 6) total number of vehicles (*Vehicles-TOT*), 7) ratio between the number of customers served within the LTZ and the number of vehicles used for deliveries within the LTZ (*Customers/Deliveries*), 8) Vehicle emission standard (*Standard*).

The CART model, reported in Figure 4, stops the tree building process when one of the following cases applies: 1) there is only one observation in each of the child nodes; 2) all observations within each child node have the same distribution of the predictor variable; or 3) the improvement (*i.e.* decrease of impurity) is smaller than a 0.05 threshold. Initially we have 56% of the observations falling in C1 and 44% in C2.

Figure 4 – CART model



The variables that proved most important in segmenting the sample were (*Customers*) and (*Deliveries*). More in detail: all the transport providers with more than 145 customers (100%) belong to C2 while those with less than 145 customers belong in 65% of the case to C1 and 35% to C2. For the group of transport providers with less than 145 customers the variable most capable of providing a relevant subdivision is (*Deliveries*). In fact, when there is a high number of deliveries (e.g. more than 4.5 deliveries per day) the original percentage of transport providers belonging to C1 and C2 is reversed since we now have 56% belonging to C2 and 44% in C1. On the other hand, with less than 4.5 deliveries per day the original percentage of observations split between C1 and C2 is further polarized with 79% belonging to C1 and 21% to C2.

In conclusion, one can state that those transport providers belonging to C2 either have more than 145 customers served within the LTZ or, in case of a lower amount of customers, perform more than 4.5 deliveries per day. Those transport providers belonging to C1 either have less than 145 customers or perform less than 4.5 deliveries per day. This information is very useful in commenting the results of the LC model. In fact, C1 is the price sensitive class whereas C2 is interested in bays based policies. It is reasonable to assume that a transport provider frequently travelling to the LTZ and serving more customers is naturally interested more in the number of LUB and PLUBF whereas a provider seldom serving the LTZ and with a limited number of customers is more interested in the EF level since, in this last case, it cannot easily amortize the cost incurred for the entrance permit.

5.2. WTP measures

WTP are used to analyze the impact of different estimation methods and measure the potential biases. This also circumvents scale problems that would, otherwise, fraud the comparison.

We calculated, using the Krinsky and Robb method, the confidence interval for the WTP measures estimated. In particular, we generated 10,000 pseudo-random draws from the unconditional distribution of the estimated parameters and subsequently calculated the simulated estimates for each draw.

Table 7 reports the WTP measures estimated based on the four models previously presented.

Table 7 – WTP estimation for the different models considered

Model	M1	M2	M3		M4	
Policy			<i>Class 1</i>	<i>Class 2</i>	<i>Class 1</i>	<i>Class 2</i>
Additional 400 LUB	95 (77-114)	113 (83-147)	48 (30-66)	246 (218-282)	52 (26-79)	488 (394-649)
Additional 800 LUB	191 (155-227)	142 (112-179)	96 (61-132)	492 (435-566)	62 (40-85)	676 (517-942)
Additional 20 PLUBF	149 (108-188)	143 (112-178)	36 (2-79)	517 (454-589)	57 (27-86)	748 (623-969)

The comparison among the different WTP measures is performed to underline the differences deriving from the consideration of: 1) non-linear effects, 2) heterogeneity, 3) joint accounting of non-linear effects and heterogeneity. The results, in our case, suggest the relevance of the methodological issues raised.

Using the Wilcoxon rank sum test, we find that the distributions intra-model (across classes) and intra-class (across models) are statistically distinct.

5.2.1. Non-linear effects

The impact of explicitly accounting for non-linear effects can be determined comparing WTP measures derived from (M1) vs. (M2) and (M3) vs. (M4) for C1 and C2 respectively.

With reference to the comparison between (M1) and (M2) one notices that while there is not a substantial difference for the WTP for 20% additional PLUBF for both models (149€ (M1) and 143€ (M2)) in the case of LUB, instead, the effects of both increments (+400 and +800) on final results are substantially different for (M1) and (M2). In fact, in this last case for +400 LUB we have a WTP of 95€ (M1) while 113€ for (M2), while for +800 LUB we have a WTP of 191€ (M1)¹² while 142€ for (M2). The difference between the measures derived from (M1) and (M2) is of around 20% for the first increment (+400) and 35% for the second (+800). It is interesting to note that, due to the erroneously assumed linear impact of the LUB increments, in the case non-linear effects are tested we avoid a progressively larger over-estimation of the impact the further away we move from the status quo level.

¹² With respect to the WTP estimates for +400 and +800 using (M1) we have respectively a WTP of 95€ and 191€ which are, considering rounding approximations, one the double of the other just because the variation of the LUB is one the double of the other and the coefficient estimating the variable's effect on choice is the same derived from the MNL model.

Comparing, respectively, the WTP of C1 and C2 in (M3) and (M4) one notices in the case of LUB for C1 that the magnitude is smaller for C1 in (M3) (48€) with respect to C1 in (M4) (52€) in the case of the first increment (+400 LUB) but this is reverted for the second (+800 LUB) where C1 in (M3) (96€) with respect to C1 in (M4) (62€). The difference with the two respective measures is 8% in the first case and 55% in the second. Not only there is an inversion of the relevance of the attribute but also the difference is substantially different in the two cases. These considerations do not apply to C2 estimates in (M3) and (M4). In this case, in fact, there is no inversion of the order of magnitude when moving from the first to the second increment since C2 in (M3) is smaller in both cases with respect to C2 in (M4). The difference between the two WTP estimates, considering the different models is 50% for +400LUB and 27% for +800 LUB. As it is for PLUBF the difference for both C1 and C2 calculated respectively with (M3) and (M4) is 36% in the first case and 30 % in the second denoting an almost equal level of mistake for both classes when non-linear effects are appropriately accounted for.

5.2.2. Heterogeneity

Heterogeneity is investigated in this paper by adopting a discrete mixture assumption. In order to verify the impact the homogenous assumption of the standard MNL model concerning the impact similar variations of a given attribute have on the sampled respondents' choices it is appropriate to compare the results obtained when estimating (M1) with those of (M3)C1 and (M3)C2 and those of (M2) with (M4)C1 and (M4)C2. The comparisons proposed, in fact, sterilize the relative impact of the non-linear effects.

When comparing the results of (M1) with those of (M3)C1 and (M3)C2 for all the three WTP measures considered it is evident the (M1) estimate is always, as expected, an average of two substantially different values (C1 and C2). In particular the differences between (M1) and (M3)C1 and (M3)C2 results, for +400 LUB, + 800 LUB and + 20% PLUBF, are respectively +47€ and -151€; +95€ and -301€; +113€ and -368€. It is evident that (M1) results represent a very rough approximation of the effective WTP measures of two completely different classes.

When comparing the results of (M2) with those of (M4)C1 and (M4)C2 for all the three WTP measures considered we have that the (M2) estimates are always an average of two even more different classes (C1 and C2). The differences between (M2) and (M4)C1 and (M4)C2 results, for +400 LUB, + 800 LUB and + 20% PLUBF, in fact, are respectively +61€ and -375€; +80€ and -534€; +86€ and -605€. The magnitude of the errors for not accounting for heterogeneity is greater when comparing models that have both accounted for non-linear effects.

5.2.3. Joint accounting of non-linear effects and heterogeneity

In order to compare the impact of the joint accounting of non-linear effects and heterogeneity in preferences on WTP measures one should compare the results derived from (M1) with those of (M4) for both C1 and C2.

Comparing (M1) with (M4)C1 and (M4)C2 respectively for +400 LUB, +800 LUB and +20% PLUBF we have that the MNL estimates overestimate form C1 and underestimate for C2 by substantial amounts and, in particular we have +43€ and -393€; +129€ and -485€; +92€ and -599€.

5.3. Policy implications

The results are extremely important from a policy perspective. This is especially so since the allocation of transport providers to the two latent classes shows that the numerosity of the two groups is almost equal and the effects of heterogeneity would, therefore, be overall pronounced. Furthermore, the non-linear effects detected in our sample confirm previous research in this field (Rotaris et al., 2012) and empirically corroborate the loss aversion hypothesis first put forward by Kahneman and Tversky (1979) and the declining impact of progressively large increases of the levels of positively impacting attributes. In other words our results suggest that the linearity in the attributes assumption should be rejected and the marginal impact on utility of the LUB attribute is not constant¹³.

The WTP estimates for qualitative attributes obtained with the linear model are similar to those estimated by the same Authors (Marcucci and Gatta, 2013) in a previous work for retailers in the same LTZ in Rome. This preliminary result should be tested checking for further heterogeneity and non-linear effects also for retailers. The fact that MNL results provide similar evaluations should be tested via a more sophisticated modeling approach of heterogeneity and non-linear effects rather than accepted as solid evidence. This, in fact, constitutes the specific objective of a new paper the Authors are presently working on.

The investigation of preferences reported in this paper testify the need for a deeper understanding of the multi-faceted world of urban freight transport if well-tailored and effective policies are to be implemented by local public administrators.

5.4. Caveats and future research

Before concluding it is important to underline also three weaknesses of our work: (a) the specificity and small dimension of the sample limits the transferability of our results to other contexts and cities; (b) the analysis focuses only on deliveries performed by manufacturers residing outside of city boundaries using third party transport providers whereas we have evidence that, in general (Patier and Routhier, 2009) and especially for Rome (Filippi and Campagna, 2008), both own-account transportation and direct deliveries by manufacturers using own vehicles (e.g. internalizing the freight distribution function) play an important role. (c) there is no explicit consideration of the specific distribution channel freight is using while we have recently acquired evidence (Danielis et al., 2010) that this is an important element in explaining different logistic constraints requiring different transport services characteristics that could justify also heterogeneous WTP measures for the policies implemented. The analysis reported in this paper can be improved in the future in a number of ways. The most important, in our judgment, are to: (a) test attribute levels with a varying interval range in order to further test for non-linearity; (b) control for heteroschedasticity by selecting a larger homogenous sample and collecting all relevant information that allows a richer specification of the model; (c) explore other ways of investigating heterogeneity in preferences.

6. Conclusions

This paper jointly tests for non-linear effects and discrete mixture heterogeneity analysis in the case of urban freight transport policy. The paper adopts an agent-specific perspective. In

¹³ Non-linear effects can be also handled via self-stated attribute cutoff (see, for example, Marcucci and Gatta, 2011; Gatta, 2006 for a detailed description and application).

particular, it investigates, via a SRE, whose results are exploded into choice decisions, the preferences for alternative policy options.

The paper investigates the effects on choices determined by variations of a specific selected set of attributes considered relevant by transport providers in the LTZ in Rome's city center. The attributes considered are LUB, PLUBF and EV. The paper: (1) tests *for non-linear effects* in attribute level variations and evaluate their impact on WTP measures, (2) investigates the presence of *heterogeneity among transport providers* by estimating a latent class model, (3) allows, via comparable WTP measures, the analysis of the joint accounting of non-linearity and heterogeneity.

The contributions to the literature relate to: (1) acquisition of a new and relevant *data set* collected during a research project conducted for Volvo Research Foundation, (2) development of an agent-specific approach to the evaluation of WTP measures for urban freight policies, (3) determination of the specific WTP measures for the various policy options considered under the assumptions tested, (4) suggestion of relevant future research paths to be explored.

The results obtained testify the need for a sophisticated treatment of the policy options that might be implemented at a local level. Both heterogeneity and non-linear effects give the WTP measures that they generate are issues that cannot be overlooked if bitter surprises are to be avoided.

The paper clearly indicates that a deeper knowledge of the sophistications and subtleties that influence the results a policy intervention can determine in the extremely complex and interrelated environment of urban freight distribution is badly needed.

Acknowledgments

The Authors would like to thank Volvo Research Foundation for providing financial support for the "Innovative solutions to freight distribution in the complex large urban area of Rome" research project and the Italian Ministry of Education, University and Research for PRIN funding of the "Methods and models for estimating the effectiveness of strategies for urban distribution of goods" research project.

7. References

- Arunotayanun K., (2009), "Modelling freight supplier behaviour and response", PhD thesis, Centre for Transport Studies, Imperial College of London.
- Ballou R.H., (1999), "Business Logistics Management: Planning, Organizing, and Controlling the Supply Chain", 4th edn., Prentice-Hall International, London.
- BESTUFS I, (2000), "Best Urban Freight Solution", <http://www.bestufs.net/>
- BESTUFS II, (2004), "Best Urban Freight Solution", <http://www.bestufs.net/>
- Boerkamps J., Binsbergen A. Van, (1999), "GoodTrip - A New Approach for Modeling and Evaluation of Urban Goods Distribution Urban Transport Systems", 2nd KFB-Research Conference, Lund, 1999.

Bowersox D.J., Closs D.J., (1996), "Logistical Management: The Integrated Supply Chain Process", McGraw-Hill Companies, New York.

Boxall P.C., Adamowicz W., (2002), "Understanding heterogeneous preferences in random utility models: A latent class approach", *Environmental and Resource Economics*, 23, pp. 421–446.

Breiman L., Friedman J.H., Olshen R.A., Stone C.J., (1984), "Classification and regression trees", Belmont: Wadsworth.

Bujosa A., Riera A., Hicks R., (2010), "Combining discrete and continuous representation of preference heterogeneity: a latent class approach", *Environment and Resource Economics*, 47, pp. 477–93.

Christopher M., (1998), "Logistics and supply chain management: strategies for reducing cost and improving service", Pearson Education, London.

Comi A., Delle Site P., Filippi F., Nuzzolo A., (2012), "Urban Freight Transport Demand Modelling: a State of the Art", in *European Transport/Trasporti Europei 2012* (51), ISTIEE, Trieste, Italy.

Corò G., Marcucci E., (2001), "Le politiche per la logistica." In: *Quaderni della Mobilità, Proceeding della Conferenza Politiche economiche per il sistema dei trasporti e della logistica*, Regione Marche. Ancona, 7 luglio 2001, vol. 1, p. 41-50.

CUPID, (2000), "Co-ordinating Urban Pricing Integrated Demonstrations", <http://www.transport-pricing.net/cupid.html>

Danielis R., Marcucci E., (2007), "L'accettabilità del centro di distribuzione urbana delle merci. Un'analisi sulle preferenze dei negozianti nella città di Fano", working paper SIET.

Danielis R., Rotaris L., Marcucci E., (2010), "Urban freight policies and distribution channels: a discussion based on evidence from Italian cities", *European Transport/Trasporti Europei*, 46, pp. 114 - 146.

Efron B., Tibshirani R., (1993), "An Introduction to the Bootstrap", Chapman and Hall: New York.

EFRUD, (2009), "Emission Free Refrigerated Urban Distribution", <http://www.efrud.info/>

EUTP II, (2000), "Thematic Network on Freight Transfer Points and Terminals", http://www.transport-research.info/web/projects/project_details.cfm?ID=5700

Filippi F., Campagna A., (2008), "Indagine sulla distribuzione delle merci a Roma, nell'ambito dello Studio di settore della mobilità delle merci a Roma", CTL: Centre for studies on Transport and Logistics and ATAC: Agency of mobility of Rome's municipality.

Filippi F., Nuzzolo A., Comi A., Delle Site P., (2010), "Ex-ante assessment of urban freight transport policies", in *Procedia - Social and Behavioral Sciences* 2 (3), E. Taniguchi and R. G. Thompson (eds.), DOI: 10.1016/j.sbspro.2010.04.042, Elsevier Ltd, 6332-6342.

Gatta V., (2006), *Valutare la Qualità dei Servizi. Un nuovo approccio basato sulla Conjoint Analysis*, *Statistica*, LXVI, 1, pp. 85-113.

Gentile G., Vigo D., (2007), "Movement Generation and Trip Distribution for Freight Demand Modelling Applied to City Logistics", Technical report, Bologna University, Italy.

Groothedde B., (2005), "Collaborative Logistics and Transportation Networks: a modeling approach to hub network design", PhD dissertation, Delft University of Technology, Trail Research School, Delft, The Netherlands.

Hole A. R., (2007), "A comparison of approaches to estimating confidence intervals for willingness to pay measures", *Health Economics*, 16(8), pp. 827-840.

Hurvich C.M., Tsai C.L., (1989), "Regression and time series model selection in small samples", *Biometrika*, 76(2), pp. 297-307.

Kahneman D., Tversky A., (1979), "Prospect theory: An analysis of decision under risk", *Econometrica*, 47, pp. 263-291.

Kamakura W., Russell G., (1989), "A probabilistic choice model for market segmentation and elasticity structure", *Journal of Marketing Research*, 26, pp. 379-390.

Krinsky I., Robb A.L., (1986), "On Approximating the Statistical Properties of Elasticities", *The Review of Economics and Statistics*, 68(4), pp. 715-719.

Krinsky I., Robb A.L., (1990), "On Approximating the Statistical Properties of Elasticities: A Correction", *The Review of Economics and Statistics*, 72(1), pp. 189-190.

Lambert D. M., (2001), "The Supply Chain Management and Logistics Controversy," in *Handbook of Logistics and Supply-Chain Management*, vol. 2, A. M. Brewer, K. J. Button, and D. A. Hensher, eds., Elsevier Science Ltd., Oxford, pp. 99-126.

Logistics World, (2007), "What is Logistics?", <http://www.logisticsworld.com/logistics.htm>, Access date: 15 November 2007.

Marcucci E., (2005), "I modelli a scelta discreta per l'analisi dei trasporti. Teoria, metodi ed applicazioni", (a cura di), Carocci, Roma.

Marcucci E., D'Agostino Z., (2003), "Centro e periferia nell'attuale evoluzione del sistema italiano della logistica e dei trasporti: alcune evidenze empiriche dal Nordest", in "Atti della VI riunione scientifica della SIET, Trasporti e politiche economiche", Palermo, 13 e 14 novembre, Numero speciale, *Annali della Facoltà di Economia e Commercio di Palermo*, pp. 253-272.

Marcucci E., Danielis R., (2008), "The Potential Demand for a Urban Freight Consolidation Centre", *Transportation*, vol. 35 (2), pp. 269-289.

Marcucci E., Danielis R., Paglione G., Gatta V., (2007), "Centri urbani di distribuzione delle merci e politiche del traffico: una valutazione empirica tramite le preferenze dichiarate", *Atti del 7° Congresso CIRIAF*, pp. 373-378.

Marcucci E., Gatta V., (2011), Regional airport choice: Consumer behaviour and policy implications, *Journal of Transport Geography*, 19(1), pp. 70-84.

Marcucci E., Gatta V., (2012), "Dissecting Preference Heterogeneity in Consumer Stated Choices", *Transport Research Part E*, 48(1), pp. 331-339.

Marcucci E., Gatta V., (2013a), "Behavioral modeling of urban freight transport. Testing non-linear policy effects for retailers", in Gonzalez-Feliu J. et al. (eds.), *Sustainable Urban Logistics: Concepts, Methods and Information Systems*, Springer.

Marcucci E., Gatta V., (2013b), "Intra-agent heterogeneity in urban freight distribution: the case of own-account operators", *International Journal of Transport Economics*.

Marcucci E., Stathopoulos A., Gatta V., Valeri E., (2012), "A Stated Ranking Experiment to Study Policy Acceptance: The Case of Freight Operators in Rome's LTZ", *Italian Journal of Regional Science*, 11(3), pp. 11-30.

Mc Fadden D., (1974), "Conditional logit analysis of qualitative choice behavior", in Zarembka P., *Frontiers in econometrics*, New York: Academic Press, pp. 105-142.

Mooney C.Z., Duval R.D., (1993), "Bootstrapping: A Nonparametric Approach to Statistical Inference", *Sage University Paper Series on Quantitative Applications in the Social Sciences*, series no. 07-095, Sage: Newbury Park.

MOST, (2000), "MObility management STRategies for the next decades", <http://www.max-success.eu/mo.st/>

Oehlert G.W., (1992), "A Note on the Delta Method", *The American Statistician*, 46(1).

Ogden K.W., (1992), "Urban goods movement: a guide to policy and planning", Ashgate, Aldershot, UK.

OSSA, (2000), "Open framework for Simulation of transport Strategies and Assessment", http://www.transport-research.info/web/projects/project_details.cfm?ID=4169

Paglione G., (2006), "City logistics: the need for a behavioral model", paper presented at Società Italiana degli Economisti dei Trasporti - VIII Riunione Scientifica. Trieste, Italy.

Paglione G., Gatta V., (2007), "Le preferenze degli attori economici verso l'installazione di un Centro di Distribuzione Urbana (CDU): il caso di Pesaro", IX Riunione Scientifica SIET, http://www.sietitalia.org/siet9/papers/Paglione-Gatta_SIET%202007.pdf

Patier D., Routhier J.L., (2009), "How to Improve the Capture of Urban Goods Movement Data?", in Bonnel P., Lee-Gosselin M., Zmud J., Madre J.L., (eds), "Transport Survey Methods: Keeping up with a Changing World", Emerald Group Publishing Limited.

PROGRESS, (2000), "PRicing REgimes fOr inteGrated SuStainable mobility", http://cordis.europa.eu/search/index.cfm?fuseaction=proj.document&PJ_RCN=4697528

Regan A.C., Garrido R., (2001), "Freight demand and shipper behavior modeling: state of the art, directions for the future", in *The Leading Edge of Travel Behavior Research*, eds D. Hensher and J. King, pp. 185-216, Pergamon Press, Oxford.

REVEAL, (1999), "Open framework for Simulation of transport Strategies and Assessment", http://www.transport-research.info/web/projects/project_details.cfm?ID=13723

Roorda M.J., Cavalcante R., McCabe S., Kwan H., (2010), "A conceptual framework for agent-based modelling of logistics services", *Transportation Research Part E*, 46, pp. 18-31.

Rotaris L., Danielis R., Marcucci E., Sarman I., (2012), "Testing for nonlinearity in the choice of a freight transport service", in Marcucci E., Puckett S., (Guest Editors), *Special issue on:*

"Freight transport analysis: new trends and methodologies", *European Transport/Trasporti Europei*, 50, pp. 1-22.

Ruesch M., Glücker C., (2001), *BEST Urban Freight Solutions*. Report of Deliverable 2.1 of Best Practice Handbook Year I. Project funded by the European Commission on the 5th RTD Framework Programme.

Rushton A., Oxley J., Croucher P., (2000), *"Handbook of Logistics and Distribution Management"*, 2 edn., Kogan Page, London.

Russo F., Comi A., (2011), "Measures for sustainable freight transportation at urban scale: expected goals and tested results in Europe", in *Journal of Urban Planning and Development* 137 (2), DOI: 10.1061/(ASCE)UP.1943-5444.0000052, American Society of Civil Engineers (ASCE), 142-152.

Samimi A., Mohammadian A., Kawamura K., (2009), "Integrating supply chain management concept in a goods movement microsimulation", *IEEE International Conference on Service Operations, Logistics and Informatics*, Chicago.

STA, (2001), "Quadro conoscitivo del problema della distribuzione delle merci nel Centro Storico di Roma", STA, Rome.

Stathopoulos A., Valeri E., Marcucci E., Gatta V., Nuzzolo A., Comi A., (2011), "Urban Freight Policy Innovation for Rome's LTZ: A Stakeholder Prospective", in *City Distribution And Urban Freight Transport: Multiple Perspectives*, Macharis C., Melo S., (eds.), Edward Elgar Publishing Limited, Cheltenham, pp. 75-100.

Stathopoulos A., Valeri E., Marcucci E., (2012), "Stakeholder reactions to urban freight policy innovation", *Journal of Transport Geography*, 22, pp. 34-45.

STRAIGHTSOL, (2012), "Strategies and measures for smarter urban freight solution", <http://www.strightsol.eu/>

SULOGRA (2000), "Effects on Transport of Trends in Logistics and Supply Chain Management", http://www.transport-research.info/web/projects/project_details.cfm?ID=6152

Taniguchi E., Thompson R., Yamada T., (2003), "Predicting the effects of city logistics schemes", *Transport Reviews*, 23 (4), pp. 489-515.

Thacher J.A., Morey E., Craighead W.E., (2005), "Using patient characteristics and attitudinal data to identify depression treatment preference groups: a latent-class model", *Depress Anxiety*, 21(2), pp. 47-54.

VRF, (2009), "Innovative solutions to freight distribution in the complex large urban area of Rome", Volvo Research and Educational Foundations and The Italian Centre of Excellence CTL – Centre for Transport and Logistics, 2008-2009.

Wilcoxon F., (1945), "Individual Comparisons by Ranking Methods", *Biometrics Bulletin*, 1, pp. 80-83.

Winston C., (1983), "The Demand for Freight Transportation: Models and Applications", *Transportation Research A*, 17A, pp. 419-27.

Winston C., (1985), "Conceptual Developments in the Economics of Transportation: an Interpretative Survey", *Journal of Economic Literature*, 23, pp. 57-94.

Yang C., Regan A., Son Y., (2010), "Another view of freight forecasting modeling trends", KSCE Journal of Civil Engineering, 14 (2), pp. 237-242.