ARTICLE IN PRESS

Transportation Research Part A xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

A latent class model with attribute cut-offs to analyze modal choice for freight transport

Concepción Román [a,*], Ana Isabel Arencibia [b], María Feo-Valero [c]

[a] Institute of Tourism and Sustainable Economic Development, University of Las Palmas GC, Facultad de Economía Empresa y Turismo, módulo D, Campus de Tafira, 35017 Las Palmas, G.C., Spain

[b] Institute of Tourism and Sustainable Economic Development, University of Las Palmas de Gran Canaria, Spain

[c] Instituto de Economía Internacional, Department of Economics, University Jaume I de Castellón, Spain

ARTICLE INFO

Article history:
Received 19 October 2015
Received in revised form 23 June 2016
Accepted 24 October 2016
Available online xxxx

Keywords:
Freight transport
Cut-offs
Latent class model
Choice experiments
Discrete choice models

ABSTRACT

We use stated preference data to analyze modal choice for freight transport when information about attribute cut-offs is introduced into the utility specification. Different choice models are estimated to account for the negative effect produced when these threshold values are violated. In order to better understand the heterogeneity in shippers' preferences, a latent class model incorporating cut-offs penalties is estimated. Our results provide evidence of the existence of differentiated classes of individuals regarding both the perception of the main attributes affecting modal choice in the corridor under analysis and penalties imposed when these attributes do not meet acceptable levels of service.

1. Introduction

Freight transport is a key element in the economic development of a country, having a direct effect on both the availability of goods and on the prices at which they are sold in the market. The complexity and heterogeneity involved in the carriage of goods makes it much more difficult to analyze than passenger transport. This is one of the main reasons why the freight transport sector has been studied much less extensively, and why many of the studies conducted have focused their analysis on the quantities actually demanded rather than examining the factors that determine the demand. In addition, the limited availability of appropriate data together with the negative externalities associated with the transport of goods, particularly for road transport, underline how difficult this task is. However, freight transport is of obvious importance within the productive structure of a country, both in terms of the logistics process and for economic development in general; the intensification of the internationalization of the economy makes it more important still.

Goods are not normally consumed in the same place as they are produced, thus requiring a process in which multiple interrelated decisions must be made in order to ensure that these goods reach their destination. The choice of the appropriate mode of transport is perhaps the most important factor in these decisions. Thus, a sound knowledge of the factors affecting this choice is essential for the evaluation of freight transport policies.

Discrete choice models are currently considered to be the most suitable methodology for analyzing consumer preferences, especially after the progress made in the past 20 years in terms of the availability of new software packages that allow increasing model complexity. As recognized by Ben-Akiva et al. (2008), there has been a substantial increase in the use of

* Corresponding author.
E-mail addresses: croman@daea.ulpgc.es (C. Román), aarencibia@acciones.ulpgc.es (A.I. Arencibia), maria.feo@eco.uji.es (M. Feo-Valero).

http://dx.doi.org/10.1016/j.tra.2016.10.020
0965-8564/

Please cite this article in press as: Román, C., et al. A latent class model with attribute cut-offs to analyze modal choice for freight transport. Transport. Res. Part A (2016), http://dx.doi.org/10.1016/j.tra.2016.10.020

behavioral models to analyze demand for freight transport over the last decade. Among the most relevant contributions to this field we can cite the work by Bergantino et al. (2013), Masiero and Hensher (2012), Arunotayanun and Polak (2011), Feo-Valero et al. (2011), Rich et al. (2009), Polak and Arunotayanun (2009), Bergantino and Bolis (2008), Beuthe and Bouffioux (2008), Brooks and Trifts (2008), de Jong and Ben-Akiva (2007), Danielis and Marcucci (2007) and Marcucci and Scaccia (2004), among others.

Within this body of literature, and in a broader sense, the analysis of taste heterogeneity has long been the focus for research in many different fields. When dealing with discrete choice data, this problem can be approached from different perspectives. With simple models, such as the Multinomial Logit (MNL), where parameters are constrained to be fixed across individuals, only systematic taste heterogeneity can be addressed by considering interactions of modal attributes with socioeconomic variables or covariates. Models in the Mixed Logit (ML) family can accommodate a huge variety of taste heterogeneity by allowing for the specification of continuous random distributions for model parameters. However, the downside here is that the researcher has to make assumptions about the parameter distributions. In contrast, the Latent Class Model (LCM) proposed by Greene and Hensher (2003) for the analysis of discrete choice data represents a semi-parametric version of the MNL that resembles the ML model in the sense that heterogeneity in the population is represented by parameters following a discrete distribution identical across individuals. As no previous assumption about the distribution of parameters is required this represents an advantage with respect to the ML model.

The LCM is based on the idea that individuals belong to different segments or classes with identical preferences within classes. It estimates the class membership probability together with the class preference parameters. The definition of the appropriate number of classes is, however, a critical issue when using this model, as it must be exogenously determined by the analyst.

Although Greene and Hensher (2003) recognize that the different models have both advantages and disadvantages, since the publication of their seminal paper, the use of LCMs has become a well-established methodology for analyzing non-observable heterogeneity, providing intuitive interpretations for practitioners and policymakers while at the same time being easy to estimate. In this regard, we can cite the works of Wen and Lai (2010) and Shen (2009) among other significant contributions; the first, in the context of choosing an airline and the second regarding the choice of a mode of transport. In terms of elasticity, both studies find that the use of the MNL model, where a class division is not set, results in a tendency to either over- or underestimate the sensitivity of service attributes. In particular, in the field of freight transport we can cite the works by Massiani et al. (2007), Greene and Hensher (2013) and Feng et al. (2013), which analyze the existence of heterogeneous segments through the estimation of an LCM in a framework of discrete choice experiments. The estimation results of this model are compared with those of an MNL and an ML model. In all these cases, two is the optimal number of classes, and all the authors come to the same conclusion: the LCM provides a better fit than the MNL and the ML, and also outperforms the other two models in terms of significance and predictive capability. In general, we can observe an increased use of LC models in the stated choice framework (e.g. Grisolía and Willis, 2012; Amador et al., 2014). Baerenklau (2010) even employs this type of model in an aggregate count data context. The latent class model is thus shown to be an effective tool for further analyzing the preferences of individuals, to check for attribute non-attendance and to determine ranges of willingness to pay (WTP). This enables the advantages LCM offers to be employed in order to set marketing, policy and managerial implications. Thus, an ever-growing number of authors use this type of model in different sectors (see e.g. Bhatnagar and Ghose, 2004; Scarpa and Thiene, 2005; Pulido-Fernández and Sánchez-Rivero, 2010). However, Provencher and Moore (2006) argue that, given that both the LCM and the ML models offer advantages, the choice of one model over another depends on the researcher's judgment as to the correlation of preference parameters.

Another critical issue in the analysis of consumer preferences is the incorporation of attribute cut-offs or threshold values in the decision-making process. For decades, the utility maximization framework, where decision makers are assumed to be fully rational and fully informed, has been successfully applied in many different fields to model consumers' choices. In this regard, the random utility model with linear-in-the-parameter specification for systematic utility, implicitly assumes compensatory behavior by decision makers, i.e. the individuals consider trade-offs among the attributes when evaluating the utility of the different alternatives. However, some other social scientists have pointed out that individuals may have limited information processing capability (e.g. Simon, 1955; Tversky and Kahneman, 1974) when they are trying to choose the best option given their constraints, recognizing that this framework represents only one of the many decision rules used by individuals (Payne et al., 1993). In this sense, individuals may exhibit non-compensatory behavior, such as in the process of elimination by aspects (EBA) proposed in Tversky (1972), where certain alternatives can be removed from the choice set if their attributes do not meet certain threshold values. This has led to more sophisticated models where decisions are modeled in two stages (Manski, 1977). First, alternatives in the choice set are determined by a non-compensatory process, such as the EBA, and secondly, the remaining ones are evaluated using a compensatory decision rule. Although this two-stage formulation seems more appropriate, its implementation is still relatively complex, as the number of choice sets increases exponentially with the number of alternatives considered.

Based on the ideas of Manski (1977) and Williams and Ortuzar (1982), Cantillo et al. (2005) formulate a hybrid semi-compensatory two-stage model incorporating endogenous thresholds for the acceptance of the attributes in the choice process. Using synthetic and real data, these authors conclude that models incorporating non-compensatory behavior outperform traditional specifications such as the MNL model.

As pointed out by Swait (2001), many of the applications incorporating cut-offs into the decision-making analysis are based on heuristics, where these threshold values are viewed as "hard" constraints imposed on the attributes. This means

Please cite this article in press as: Román, C., et al. A latent class model with attribute cut-offs to analyze modal choice for freight transport. Transport. Res. Part A (2016), http://dx.doi.org/10.1016/j.tra.2016.10.020