Using semi-open questions to integrate perceptions in choice models

Aurélie Glerum [a,*], Bilge Atasoy [b], Michel Bierlaire [a]

[a] École Polytechnique Fédérale de Lausanne (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory (TRANSP-OR), Lausanne, Switzerland
[b] Massachusetts Institute of Technology (MIT), Department of Civil and Environmental Engineering, Cambridge, MA, USA

ARTICLE INFO

ABSTRACT

This research investigates the measurement of perceptions by means of adjectives freely reported by respondents in semi-open questions. It involved the use of semi-open responses of 1763 Swiss individuals to develop indicators for a latent variable representing the perception of comfort of public transportation. The indicators are then incorporated into a discrete choice model of revealed mode choices.

Perceptions are assumed to impact choice significantly and this research aims at capturing their complexity using adjectives and integrating them into the hybrid choice modeling framework.

We exploit a quantification of the adjectives performed by external evaluators. Given the subjectivity that is involved, we analyze the sensitivity of the results across evaluators who rated the adjectives. We observe that the aggregate indicators of demand, such as market shares, elasticities and values of time, are rather robust across evaluators. This is not the case for the disaggregate indicators that may vary substantially across evaluators.

1. Introduction

Recently the discrete choice modeling (DCM) literature has been influenced by the development of new models, namely the *hybrid choice models* (*HCM*). Among other properties, such models allow for the integration of psychological constructs such as perceptions as explanatory variables of the choice. The recent motivation for the integration of perceptional aspects into discrete choice models has raised a few issues regarding their measurement. First, it is essential to develop an adequate way to quantify measures of the perceptions and integrate them into a discrete choice model. Second, the subjectivity inherent to the quantification procedure must be assessed. The present research aims at addressing these issues.

To evaluate psychological constructs such as perceptions, attitudes or lifestyles, survey techniques have been developed by social scientists and involve the collection of *psychometrics* (Thorndike, 1920; Likert, 1932; Bearden and Netemeyer, 1999). They usually consist of proposing statements where respondents have to indicate a rating on a five-point scale. For example, Vredin Johansson et al. (2006) collect ratings of the importance of some perceptional questions related to comfort, convenience and flexibility of transport modes. Abou-Zeid et al. (2012) ask respondents to indicate their satisfaction when they commute either by public transportation or by car. Atasoy et al. (2013) request individuals to rate their agreement on a list of statements related to environmental concern, mobility, residential choice or lifestyle. The survey methods used for these studies mostly rely on closed questions.

---

* Corresponding author.
  E-mail addresses: aurelie.glerum@epfl.ch (A. Glerum), batasoy@mit.edu (B. Atasoy), michel.bierlaire@epfl.ch (M. Bierlaire).

The main advantage of such questions is that they are easy to collect, code and integrate into existing model frameworks. However, social scientists emphasize on the importance of exploiting the information contained in responses to *open questions*. Such questions are often under-exploited or only used as a complement to closed questions (Looker et al., 1989). Mossholder et al. (1995) point out the importance of such questions to retrieve constructs that can be recognized in the semantic content of language, such as affection or emotions. In particular, Beirao and Cabral (2007) use open questions to measure attitudes towards public transportation and car. More importantly, it is recognized that such questions reflect better respondents' conception of a construct (Potkay and Allen, 1973; Mossholder et al., 1995; Kaufmann et al., 2001), while closed questions are the result of the survey designer's representation of it. The diversity of answers in open questions is also wider than in closed questions (Schuman and Presser, 1996).

Open questions can be structured in various ways. In this research we focus on a subset of them, which we denote by *semi-open questions*, and where we ask the respondents to report several adjectives describing best a variable of interest (e.g. a transportation mode). This data collection technique has been used in psychology to obtain a representation of a person (Potkay and Allen, 1973) and more recently in social sciences to obtain representations of transportation modes (Kaufmann et al., 2001, 2010). Though this data collection method and the well-known *word association* technique may show some common features, we distinguish them. In word association surveys, individuals are required to report the 'first single word which comes to mind' (Cramer, 1968) when they are exposed to a stimulus word. In semi-open questions, respondents are required to reply to a more precise question and give a more structured answer.

The use of semi-open questions to measure psychological constructs such as perceptions requires the development of an appropriate modeling framework. For behavioral researchers, the heterogeneity in individuals' behavior has always been an interesting field and several methodologies are developed in order to account for this heterogeneity. In transportation research, mixed logit models are used to capture the random taste heterogeneity across consumers (McFadden and Train, 2000; Hess and Train, 2011). Mixed logit models are flexible in terms of the underlying distribution assumptions for the unobserved heterogeneity (Greene and Hensher, 2003). On the other hand, many researchers work with latent variables (LV)/latent classes (LC) in order to explain the heterogeneity with structural equation models. These structural equation models are built with the characteristics of individuals in order to obtain a systematic representation of the heterogeneity. The latent constructs are integrated into choice models using the HCM framework (Walker, 2001; Walker and Ben-Akiva, 2002; Ben-Akiva et al., 2002). In the literature, the effect on choice of LVs measured by closed questions has been evaluated for various applications. In transportation mode choice contexts, Espino et al. (2006) assess the effect of the LV 'comfort' on the choice between bus and car. Vredin Johansson et al. (2006) also analyze the impact of comfort on mode preferences as well as the one of flexibility and care for the environment. Daly et al. (2012) model the effect of security on rail travel behavior.

In this paper, we consider the HCM framework and illustrate our methodology by modeling the impact of the LV 'perception of comfort of public transportation' on choice. Since we are using responses to semi-open questions as measurements of the LV, a method to quantify the responses to the semi-open questions is required. We propose a method which consists of asking a set of individuals (the 'evaluators') to rate the adjectives on a scale relative to the LV, where a positive number indicates a positive representation of the LV and a negative number implies a negative representation of it. The use of individuals to evaluate qualitative concepts is common practice in computer science studies (Amazon's Mechanical Turk, 2005-2012; Franklin et al., 2011) and research on natural language processing (Snow et al., 2008). Humans happen to be more effective than computers for some specific tasks, such as evaluating a language (Snow et al., 2008; Sorci et al., 2010; Robin et al., 2011; Venetis et al., 2012). Moreover, in medical and social sciences scores from a number of raters are used for decision-making. In such domains it is critical to have an agreement between the raters in order to be fair to the subjects. Therefore there is a well-studied literature on how to minimize the impact of the rater factor on the data quality. We refer to the review by Banerjee et al. (1999) for the analysis of agreement between raters which is called as *inter-rater reliability*. However in our case, we are interested in the heterogeneity of raters and we analyze the impact of different ratings on the results. Therefore we do not seek for inter-rater reliability but rather analyze the differences between ratings.

To assess the sensitivity of the market shares of the different transportation modes to variations of exogenous factors, different demand indicators need to be derived. In a transportation mode choice context, important demand indicators include *willingness-to-pay* (*WTP*) indicators, such as the *value of time* (VOT), and *elasticities*. In the literature, it has been shown that HCMs allow for a more complex representation of the demand. For example, Abou-Zeid et al. (2010) capture the heterogeneity in the VOT within the population by considering an interaction term between travel time and the LV. Bolduc and Daziano (2011) show that several other WTP indicators can be derived to valuate the demand in a vehicle choice context. Yáñez et al. (2010) analyze the effect of changes in the LV on individuals' choices. The variations are indeed triggered by changes in an explanatory variable of the structural model of the LV. Building upon the ideas of Abou-Zeid et al. (2010) and Yáñez et al. (2010), we respectively derive a series of demand indicators and analyze the effect of a change in an explanatory variable of comfort of public transportation on the choice. The novelty of our approach consists of considering the distribution of the indicators across evaluators and providing statistics on it. We hence aim at demonstrating the robustness of the method in terms of model application, when the ratings of different evaluators are considered.

In this paper, we achieve a quantification of data from adjectives and introduce this information as measurement of an LV in an HCM. We moreover propose an analysis of the sensitivity of the demand indicators to the evaluators' ratings, which leads to the empirical evidence that disaggregate indicators of demand vary more importantly across evaluators than aggregate indicators.