TRANSP-OR

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Correcting for endogeneity using the multiple indicator solution

Anna Fernández Antolín * Angelo Guevara-Cue †
Matthieu de Lapparent * Michel Bierlaire *

April 5, 2016

Report TRANSP-OR 160405
Transport and Mobility Laboratory
Ecole Polytechnique Fédérale de Lausanne
`transp-or.epfl.ch`

*Transp-OR, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, anna.fernandezantolin@epfl.ch, michel.bierlaire@epfl.ch, matthieu.delapparent@epfl.ch

†Universidad de Chile, aguevara@uandes.cl

Abstract

In this paper we extend the Multiple Indicator Solution (MIS), used to correct for endogeneity, so that it can also be used when there are interactions between observed and unobserved factors in the specification of the utility function. We show the theoretical derivation and illustrate it with a case study. Policy indicators, such as time elasticity and value of time are derived, and the results are compared with a logit model and with an Integrated Choice and Latent Variable (ICLV) model. The contribution is twofold: this is the first application of the MIS methodology with revealed preference data, and the MIS has been adapted to account for interactions between observed and unobserved attributes.

Keywords: Discrete choice models, mode choice, value of time, endogeneity, multiple indicator solution, latent variable, revealed preference.

1 Introduction

Endogeneity is an issue that often arises in demand modeling. One of the assumptions to derive random utility models such as logit, probit, nested logit and cross nested logit is that the deterministic part of the utility function is independent from unobserved factors. If this assumption is violated, it may result in inconsistent estimates of the parameters. This is what is known as endogeneity. As Guevara-Cue (2010) describes, it can have three main causes: (i) errors in the measurements of the variables, (ii) simultaneous determination and (iii) omitted variables.

The first cause is very intuitive: if there are systematic errors in the measurements, these propagate to the error term, which is then correlated with the wrongly measured variable. An example of the second cause in the context of transportation can be found in the simultaneous modeling of mode and housing choice. People with a tendency to travel by public transportation locate closer to stations, thus making their travel times shorter. The residential location choice is affected by the mode choice, but at the same time the mode choice is affected by the residential location choice. This is known as simultaneous determination, and assuming that one is an exogenous explanatory variable of the other is wrong.

An example of the third cause can also be found in transportation, when an unobserved variable - such as comfort - is not included in the model. In a mode choice between public transportation and private modes, assume that there is an observed attribute (travel time, travel cost) that is correlated with an unobserved attribute (perception of comfort). If comfort is omitted, we may obtain biased estimates for the parameters associated with time and/or cost. This can be seen intuitively as follows: if people are traveling at peak hours when public transportation is very congested, the disutility towards public transportation caused by discomfort is captured by the travel time parameter. It results in a downwards-estimated parameter for travel time, since it captures both the disutility towards public transportation caused by travel time and the disutility caused by discomfort. In a similar way, transportation systems that are more expensive because they are more comfortable - like traveling in the first class in a train - have an upwards estimated parameter related to cost. This parameter is capturing on the one hand the disutility for high prices, but on the other hand the fact that travelers are willing to pay higher prices to travel in a more comfortable way. It can even result in positive estimates for parameters related to cost. This results, of course, in wrong willingness to pay estimates.

The problem may appear as well when a latent variable is omitted. There is evidence in the literature that *car lovers* have a different value of time for private motorized modes compared to other individuals who don't have this preference (Atasoy *et al.* , 2013). If this is not explicitly modeled, the estimator of value of time may not be consistent. In terms of the specification of the utility, there is evidence in the literature that suggests to use the interaction of *car lovingness* and cost to address heterogeneity of taste (Abou-Zeid *et al.* , 2010).

As discussed above, endogeneity can yield to biased and inconsistent estimates. However, it is rarely assessed and corrected for in practical applications. This is due to the fact that although several methods to correct for it exist (BLP, control function...), they rely on instruments, that are not straightforward to identify in practice. A complete review of these methods is found in Section 2. In this paper, we build on the Multiple Indicator Solution (MIS), that can be applied when there is an interaction between the unobserved factor and a measurable variable. We show that it can be generalized to models with interactions between observed and unobserved factors. Moreover, it is the first application with real data of the MIS method, that has only been tested with stated preference (SP) data (Guevara & Polanco, 2016). We apply the MIS methodology in order to get more realistic value of time (VOT) and time elasticity estimates from a mode choice revealed preference dataset in Switzerland. We show that it handles correctly the endogeneity issue by comparing it with the integrated choice and latent variable (ICLV) approach.

This paper is structured as follows: the literature review is presented in Section 2, followed by the description of the theoretical framework in Section 3. Section 4 contains the case study, along with a discussion of the results obtained. Finally, the conclusions and future work directions are discussed in Section 5.

2 Literature Review

This section is divided in two subsections: Section 2.1 is a detailed review of the different methodologies that have been proposed in the literature to address endogeneity. Section 2.2 gives some insight in the existing literature related to modeling attitudes and perceptions.

2.1 Endogeneity

Louviere *et al.* (2005) present the recent progress that has been done in the field of endogeneity in discrete choice models. However, they give a very broad definition of endogeneity and focus also on choice set formation, interactions among decision makers and models of multiple discrete/continuous choice amongst other topics. In this review, as well as during the whole paper, we are going to focus only on how to correct for endogenous explanatory variables.

A widely used methodology is the BLP (Berry *et al.* , 1995, 2004) named after its authors. This approach consists in removing the endogeneity from the non-linear choice model and dealing with it in linear regressions. This requires adding an alternative specific constant (ASC) for each product and each market. A description of the instrumental variable methodology can be found in most of the basic econometric textbooks such as Baum (2006), Lancaster (2004) or Wooldridge (2010). Guevara-Cue (2010) describes in his thesis why it is more complex to deal with endogeneity in discrete choice models compared to linear models: these corrections lead to changes in the error term which imply a change of scale in the discrete choice models.

There are many studies that use the BLP approach to deal with endogeneity in discrete choice models. To name some examples, Walker *et al.* (2011) introduce a social influence variable in a behavioral model which is endogenous, as the factors that impact the peer group also influence the decision maker and this causes correlation between the field effect variable and the error. Train & Winston (2007) use the BLP approach to correct for price endogeneity in automobile ownership choice. Crawford (2000) uses it for consumers' choice among TV options and Nevo (2001) uses it for a study of the cereal industry. It is also the approach chosen by Goolsbee & Petrin

(2004) where they examine the direct broadcast satellites as a competitor to cable TV.

A second approach in the literature is the control function methodology. The concept dates back to Hausman (1978) and Heckman (1978), although the term *control function* was introduced by Heckman & Robb Jr. (1985). Petrin & Train (2009) describe a control function approach to handle endogeneity in choice models. They apply both the control function and the BLP methodologies in a case study and find similar and more realistic demand elasticities than without correcting for endogeneity. They describe the control function methodology in detail. Guevara-Cue (2010) also uses this method to study the choice of residential location. He also shows that there is a link between the control-function methods and a latent-variable approach.

The third frequently used approach is the one that Guevara-Cue (2010) calls *the control-function method in a maximum-likelihood framework* and Train (2003) calls *maximum-likelihood method*. It is the same formulation used by Villas-Boas & Winer (1999) in brand choice models and Park & Gupta (2012). In particular, Park & Gupta (2012) propose what they describe as a "new statistical instrument-free method to tackle the endogeneity problem". They model the joint distribution of the endogenous regressor and the structural error term by a Gaussian copula and use nonparametric density estimation to construct the marginal distribution of the endogenous regressor. Also, Bayesian methods to handle endogeneity have been introduced by Yang *et al.* (2003) and Jiang *et al.* (2009).

Endogeneity can also be mitigated by the Integrated Choice and Latent Variable (ICLV) approach, where a latent factor captures an unobserved qualitative attribute. This methodology explicitly models attitudes and perceptions using psychometric data. For the estimation of the parameters, maximum likelihood techniques are used, which lead to complex multi-dimensional integrals. Thus, it is a computationally intensive method.

A more novel method used for discrete choice models is the Multiple Indicator Solution (MIS) which is described by Wooldridge (2010) in the context of linear models and generalized by Guevara & Polanco (2016) for discrete choice. As opposed to the control-function method, the MIS method does not need instrumental variables. Instead, it uses indicators to introduce a factor of correction in the choice model in order to obtain consistent estimators. Its performance is compared using Montecarlo experiments to other methodologies in Guevara (2015).

There are also other methods, but are less used because they are either very recent, such as Grange *et al.* (2015), or outperformed by the methods reviewed above. For example, the analogous to the standard 2-stage instrumental variable approach used in regression, described by Newey (1985) does not provide correct estimates of the aggregate elasticities of the models. Guevara-Cue (2010) shows it with a case study. Another method, developed by Amemiya (1978), is as efficient as the control function approach, as shown by Newey (1987), and is globally efficient under some circumstances, but is much more complex to calculate because it involves the estimation of auxiliary models.

2.2 Attitudes and perceptions

A lot of literature also exists in how attitudes, perceptions and psychological factors in general play an important role in the modeling of behavior. A non-exhaustive list of research related to this would include Ajzen (2001); Olson & Zanna (1993); Wood (2000); McFadden (1986); Ben-Akiva & Boccara (1987). In particular, there are several studies describing the role of attitudes and perceptions in mode choice, such as Koppelman & Hauser (1978); Proussaloglou & Koppelman (1989); Golob (2001); Outwater *et al.* (2003); Vredin Johansson *et al.* (2006)). Walker (2001) develops the most commonly used framework to include these in discrete choice models: the integrated choice and latent variable approach. However there had already been some developments of latent variable models prior to her work, such as Everitt (1984); Bollen (1989).

An interesting measure that can be derived from mode choice models is the value of time (VOT), that is defined as the amount of money that users are willing to pay to save one unit of travel time. In other words, it is the trade-off that users consider between the time that they spend traveling and the amount of money that they are willing to pay. The first person to introduce the concept of value of time in travel behavior was Dupuit

(1844, 1849). The VOT varies across individuals and the trips, characterized by variables such as age, gender, income, trip purpose... It can also be distributed (see, among others, Ben-Akiva *et al.* (1993); Fosgerau (2006); Hess & Axhausen (2004)).

An attitude that has been considered relevant for the estimation of the VOT is the *car loving* attitude (Abou-Zeid *et al.* , 2010; Atasoy *et al.* , 2013). *Car lovers* are defined as people that have an intrinsic preference towards car, for many reasons, including convenience, reliability, and symbol of social status. If either the time or the cost are actually interacting with the attitude, and it is omitted in the model specification, it then enters the error term, causing endogeneity.

3 Methodology

This section introduces the methodology that is used in the paper. Section 3.1 is an introduction to the Multiple Indicator Solution (MIS) method. The following sections investigate how to adapt this methodology to capture possible interactions between observed attributes and unobserved factors. Section 3.2 contains the derivation of an intuitive but not useful approach, while Section 3.3 proposes a way to overcome the limitations of the previous approach. Finally, Section 3.4 is a reminder of the Integrated Choice and Latent Variable (ICLV) framework, that is used as a benchmark for the MIS with interactions in the case study.

3.1 MIS method

The multiple indicator solution method was introduced by Wooldridge (2010) for linear models and extended to discrete choice models by Guevara & Polanco (2016). It can be summarized as follows.

Consider a setup where the choice of an alternative i by a decision-maker n depends on an economic factor t_{in}, an unobserved attribute q_{in} that is correlated to q_{in}, and on a set of other explanatory variables x_{in}. The utility function of this alternative is specified as follows

$$U_{in} = \mathrm{ASC}_i + \beta_x x_{in} + \beta_t t_{in} + \beta_q q_{in} + e_{in}, \tag{1}$$

where ASC_i, β_x β_t and β_q are parameters to estimate and e_{in} is a random error term. If the term $\beta_q q_{in}$ is omitted, it would enter the error term. Therefore, the error term would be correlated to t_{in} causing endogeneity. We assume that we have two indicators I_{1in} and I_{2in} which are related to the omitted variable q_{in}. The following relation can be defined

$$I_{1in} = \alpha_0 + \alpha_q q_{in} + e_{I_{1in}}, \tag{2}$$

$$I_{2in} = \delta_0 + \delta_q q_{in} + e_{I_{2in}}, \tag{3}$$

where

$$\mathrm{cov}(q, e_{I_1}) = \mathrm{cov}(x, e_{I_1}) = \mathrm{cov}(q, e_{I_2}) = \mathrm{cov}(x, e_{I_2}) = \mathrm{cov}(e_{I_1}, e_{I_2}) = 0,$$

$\alpha_q \neq 0$ and $\delta_q \neq 0$. x represents the vector of explanatory variables in Equation 1. From Equation (2) we obtain $q_{in} = (I_{1in} - \alpha_0 - e_{I_{1in}})/\alpha_q$. By substituting this expression in Equation (1) and denoting $\theta_q = \frac{\beta_q}{\alpha_q}$ we obtain

$$U_{in} = ASC_i + \beta_t t_{in} + \beta_x x_{in} + \theta_q I_{1in} - \theta_q \alpha_0 - \theta_q e_{I_{1in}} + e_{in}. \tag{4}$$

The above model is still endogeneous since I_{1in} is correlated with $e_{I_{1in}}$. We therefore apply the control function method (similarly as in Guevara-Cue (2010)) and use I_{2in} as an instrument for I_{1in}. This can be done because both indicators are correlated, and I_{2in} is uncorrelated with $e_{I_{1in}}$. We can therefore define the following relations

$$I_{1in} = \gamma_0 + \gamma_1 I_{2in} + \gamma_t t_{in} + \gamma_x x_{in} + \delta_{in}, \tag{5}$$

$$e_{I_{1in}} = \beta_\delta \delta_{in} + \nu_{in}, \tag{6}$$

where δ_{in} captures the part of $e_{I_{1in}}$ which is correlated with I_{1in} and ν_{in} is an exogenous error term.

Substituting Equation (6) to (4) we obtain

$$U_{in} = (ASC_i - \theta_q \alpha_0) + \beta_t t_{in} + \beta_x x_{in} + \theta_q I_{1in} - \theta_q \beta_\delta \delta_{in} - \theta_q \nu_{in} + e_{in}. \tag{7}$$

By denoting $A\tilde{S}C_i := ASC_i - \theta_q\alpha_0$, $\theta_\delta := -\theta_q\beta_\delta$ and $\tilde{e}_{in} := -\theta_q\nu_{in} + e_{in}$ we obtain

$$U_{in} = A\tilde{S}C_i + \beta_t t_{in} + \beta_x x_{in} + \theta_q I_{1in} + \theta_\delta\delta_{in} + \tilde{e}_{in}, \tag{8}$$

where there is no endogeneity anymore.

A limitation of this methodology is that the indicator I_{1in} and the residuals of the regression δ_{in} appear directly in the utility function, as seen in Equation (8). This might not be an issue when the purpose of the model is to derive trade offs such as willing to pay estimates or elasticities. However, if the purpose is to do forecasting, it would be. How to overcome this limitation is out of the scope of the paper, but a research direction would be to write a measurement equation of the indicators that depends on socioeconomic characteristics. By doing this, the indicators could be forecasted and so could be the result of the regression in Equation (5). This also applies to the following Sections 3.2 and 3.3.

3.2 MIS method and interactions: first approach

Assume now that the variable q is an interaction term $t_{in} \cdot \xi_n$, where ξ_n is a characteristic of the decision-maker. The specification of the utility function is then

$$U_{in} = \mathrm{ASC}_i + \beta_x x_{in} + \beta_t t_{in} + \beta_\xi t_{in}\xi_n + e_{in}. \tag{9}$$

Suppose again that we have two indicators I_{1in}, I_{2in} for the variable ξ_n, that is, $I_{1in} = \alpha_0 + \alpha_\xi\xi_n + e_{1in}$. If we repeat the derivation from section 3.1 we obtain

$$U_{in} = ASC_i + (\beta_t - \theta_\xi\alpha_0)t_{in} + \beta_x x_{in} + \theta_\xi t_{in} I_{1in} - \theta_\xi t_{in}\beta_\delta\delta_{in} + \theta_\xi t_{in}\nu_{in} + e_{in}, \tag{10}$$

and by denoting $\tilde{\beta}_t := \beta_t - \theta_\xi\alpha_0$ and $\theta_\delta := -\theta_\xi\beta_\delta$ we obtain

$$U_{in} = ASC_i + \tilde{\beta}_t t_{in} + \beta_x x_{in} + \theta_\xi t_{in} I_{1in} + \theta_\delta t_{in}\delta_{in} + \underbrace{\theta_\xi t_{in}\nu_{in} + e_{in}}_{\tilde{e}_{in}(t)}. \tag{11}$$

For this reason, this approach is not further investigated.

3.3 MIS method and interactions: correct approach

In order to use the MIS method in the presence of interactions between an attribute t_{in} and an unobserved factor ξ_{in}, we need to assume $t_{in} \cdot I_{1in}$ and $t_{in} \cdot I_{2in}$ to be indicators for $t_{in} \cdot \xi_n$. We define $\xi'_{in} = t_{in} \cdot \xi_n$, $I'_{1in} = t_{in} \cdot I_{1in}$ and $I'_{1in} = t_{in} \cdot I_{2in}$. The following relation can therefore be defined

$$I'_{1in} = \alpha_0 + \alpha_\xi \xi'_{in} + e_{I_{1in}}. \tag{12}$$

We can also define

$$I'_{1in} = \gamma_0 + \gamma_1 I'_{2in} + \gamma_t t_{in} + \gamma_x x_{in} + \delta_{in}, \tag{13}$$

$$e_{I_{1in}} = \beta_\delta \delta_{in} + \nu_{in}, \tag{14}$$

where δ_{in} captures the part of $e_{I_{1in}}$ which is correlated with I'_{1in} and ν_{in} is an exogenous error term.

From Equation (12) we obtain $\xi'_{in} = (I'_{1in} - \alpha_0 - e_{I_{1in}})/\alpha_\xi$. By substituting this expression in Equation (9), denoting $\theta_\xi = \frac{\beta_\xi}{\alpha_\xi}$; proceeding as in Section 3.2, denoting $A\tilde{S}C_i := ASC_i - \theta_\xi \alpha_0$, $\theta_\delta := -\theta_\xi \beta_\delta$ and $\tilde{e}_{in} := -\theta_\xi \nu_{in} + e_{in}$ we obtain

$$U_{in} = A\tilde{S}C_i + \beta_t t_{in} + \beta_x x_{in} + \theta_\xi t_{in} I_{1in} + \theta_\delta \delta_{in} + \tilde{e}_{in}, \tag{15}$$

where the endogeneity has been corrected. The model with the MIS correction is estimated in two stages. First δ_{in} is obtained by taking the residual values of Equation (13). Second, all parameters of Equation (15) are estimated by maximum likelihood. Note that using the full information maximum likelihood would render a one-stage estimation possible.

3.4 Integrated choice and latent variable (ICLV) model framework

Instead of using the MIS method to account for the omission of $t_{in}\xi_n$, the ICLV methodology can also be used. We introduce it here briefly and refer the reader to Walker (2001) for more details. Let us now consider a model with the same formulation of utility as in Equation (9).

The structural equation of the latent variable model is given as follows

$$\xi_n = \eta_o + \eta s_n + \omega_\xi, \tag{16}$$

where η_o, η are (vectors of) parameters to estimate, s_n is a vector of socio-economic characteristics of the respondent n, and ω_ξ is an error term.

The measurement model specifies the following k measurement equations

$$t_{in} I_{kin} = \alpha_k + \lambda_k \xi_n t_{in} + \omega_{I_{kin}}, \tag{17}$$

where α_k and λ_k are parameters to estimate, and $\omega_{I_{kin}}$ is a random error term. To compute the maximum likelihood function, integration over ξ is performed which makes it more computationally complex to estimate. Therefore, the identification of the parameters is not as straight forward as for the MIS method.

4 Case Study: Mode choice in Switzerland with RP data

The description of the case study is organized as follows: Section 4.1 introduces the dataset that is used, including details of the data collection and some descriptive statistics. It is followed by the model specification in Section 4.2. Finally, the results are presented in Section 4.3

4.1 Data used: collection and exploratory analysis

The dataset used for the case study was collected in Switzerland between 2009 and 2010 as part of a project to understand mode choice and to enhance combined mobility behavior. It consists of a revealed preferences (RP) survey. Details about the data collection procedure can be found in Bierlaire *et al.* (2011); Glerum *et al.* (2014), and more information about the project can be found in `http://transport.epfl.ch/optima`.

The structure of the questionnaire is as follows. There is a first part consisting of a revealed preferences survey where information on all the trips performed during one day are collected. Respondents report travel time, travel cost, socioeconomic characteristics of themselves and of their household, opinions on a list of statements, mobility habits and what is referred to in Glerum *et al.* (2014) as *semi-open questions*. In these semi-open questions, respondents are asked to provide three adjectives to describe each mode. Each observation corresponds to a round trip, not to a single trip. After removing (i) observations where the mode is not reported, (ii) observations corresponding to respondents who claim to use the car, but answer simultaneously that they do not have access to a car, (iii) those who do not answer to the opinion statement that are used for the modeling and (iv) those who do not report their income level, there is a total of 1,686 observations.

The mode alternatives are public transportation (PT), private motorized modes (PMM) (car, motorbike, etc.) and slow modes (SM) (bike, walk). PMM is also referred to as *Car*. Table 1 shows the sample market shares for each of the three considered modes. These are the results after excluding the respondents described above. Of these, only 83 had no access to car. This is taken into account for the modeling. The market shares observed in the sample are coherent with the real market shares in the population (Office fédéral de la statistique, 2012).

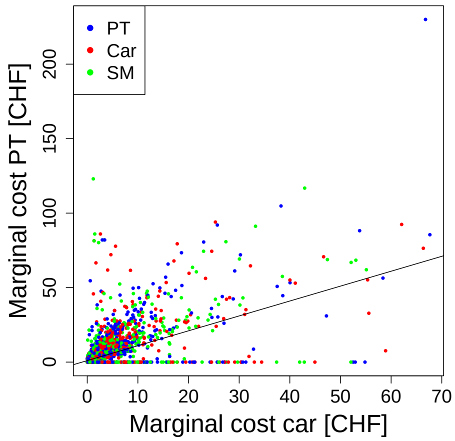
	PT	PMM	SM	Total
Number of observations	456	1,128	102	1,686
Observed market shares (%)	27	67	6	100

Table 1: Observed market shares and number of observations for each of the three alternatives in the choice set (public transportation, private motorized modes and slow modes).

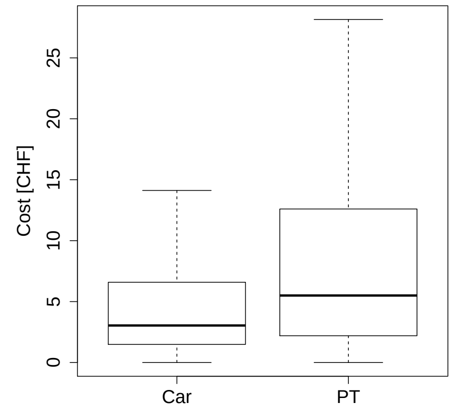
4.1.1 Travel time and travel cost

Figure 1 shows the travel time and cost both by car and public transportation for each individual. The reported travel time for the chosen mode is not used, instead, it is imputed. Details can be found in Bierlaire *et al.* (2011).

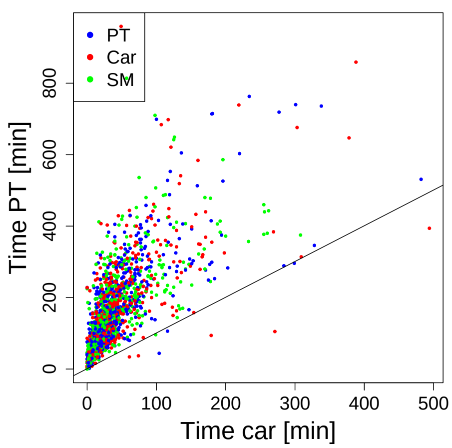
It is observed in Figures 1(b),1(d) that in general terms car is faster and cheaper than public transportation. This is confirmed by Figure 1(c) where we see that there are less than 10 observations where public transportation is faster than car. In Figure 1(a), we see that there are several respondents for which the marginal cost by public transportation is zero. This is due to the fact that respondents in the dataset can have several travel cards that makes their marginal cost null. In both figures, the black line represents the $x = y$ line.



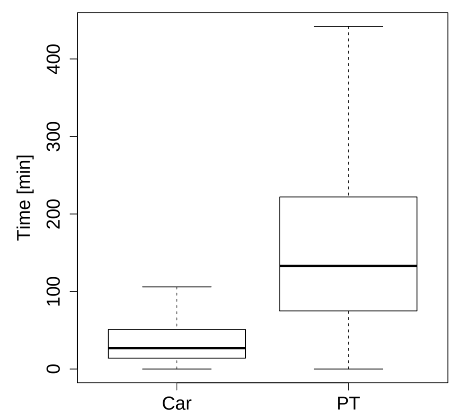
(a) Marginal cost.



(b) Boxplot of marginal cost.



(c) Travel time.



(d) Boxplot of travel time.

Figure 1: Plots and boxplots of travel time and travel cost for the different alternatives.

4.1.2 Attitudinal questions

Several attitudinal questions related to the car-loving attitude are rated in a 1 to 5 likert scale by the respondents. The statements that are used in this case study are the following

1. It is difficult to take the public transportation when I travel with my children.

2. With my car I go whenever and wherever.

As described in Section 3.3, the indicators that are considered for this case study are the product of these ratings and the travel time. The one corresponding to statement *With my car I go whenever and wherever* is referred to as *flexibility indicator* and the one related to statement *It is difficult to take the public transportation when I travel with my children* is referred to as *convenience indicator*. The correlation between them is 0.88. In the reminder of the paper, the expression *likert indicator* is used when referring to the 1 to 5 indicators, and the expression *composite indicator* is used to refer to the product of this indicators and travel time.

4.2 Model specification

Table 2 shows the model specification used as the base model for the case study. It is a model with 13 parameters. In the slow modes utility function, only the distance of the trip and the number of bicycles in the household are considered as explanatory variables.

In the public transportation utility, there is the alternative specific constant (ASC), some socioeconomic variables related to the type of neighborhood (rural vs urban) and to the occupation (student or not), as well as attributes of the mode such as cost and time, where cost is interacted with the income of the respondent. The parameter for time is an alternative specific one, while the parameter related to travel cost is generic for both alternatives.

In the car utility function there is also an ASC and three socioeconomic variables which are if the respondent is from a French speaking part of

Parameter	Public transportation	Car	Slow modes
β_1 (ASC_{PT})	1	0	0
β_2	0	Time car [min]	0
β_3	Travel time by PT [min]	0	0
β_4 (ASC_{car})	0	1	0
β_5	0	Number of children	0
β_6	0	Number of cars	0
β_7	$\frac{\text{Marginal cost of PT}}{\text{Income}}$	$\frac{\text{Marginal cost of car}}{\text{Income}}$	0
β_8	0	Work-related trip	0
β_9	0	French speaking	0
β_{10}	Student	0	0
β_{11}	Urban area	0	0
β_{12}	0	0	dist. [km]
β_{13}	0	0	Number of bicycles

Table 2: Base model specification.

Switzerland or not, the number of cars in the respondent's household and the number of children in the household. There are also the time and cost of the trip, where the cost is the gasoline cost, and it is again interacted with the income of the respondent. There is also a dummy variable for the trip purpose (if it is work-related or not).

The specifications used for the other two models (MIS and ICLV) are the same except for the parameters associated with each methodology. The base model specification is suspected to suffer from endogeneity issues for the reasons discussed earlier.

4.3 Results

The presentation of the results is divided in several sections. Sections 4.3.1, 4.3.2, 4.3.3 present the estimation results of the logit, logit with MIS correction and ICLV methodology respectively. They are followed by Sections 4.3.4 and 4.3.5 where a comparison of the results obtained is performed. All models are estimated using Biogeme, an open source software designed for the estimation of discrete choice models (Bierlaire, 2003).

4.3.1 Base model: Logit

Table 3 shows the estimation results for the model specification defined in Table 2. The signs are in line with our expectations and the literature. The parameters associated with travel time, travel cost and distance are negative. Moreover, travel time in private modes causes more disutility than travel time in public transportation. This is justified by the fact that the time in public transportation can be used to do other things, while when a person is driving s/he can not do any other activity.

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC (PT)	1.08	0.399	2.71	0.01
2	Travel time [min] (Car)	-0.0272	0.00507	-5.37	0.00
3	Travel time [min] (PT)	-0.00878	0.00169	-5.19	0.00
4	ASC (Car)	0.257	0.440	0.58	0.56
5	No. of children in household (Car)	0.181	0.0699	2.59	0.01
6	Number of cars in household (Car)	1.04	0.125	8.32	0.00
7	$\frac{\text{Marginal cost}}{\text{Income}}$	-0.334	0.0817	-4.08	0.00
8	Work related trip (Car)	-0.659	0.130	-5.06	0.00
9	French speaking (Car)	1.01	0.175	5.79	0.00
10	Student (PT)	2.94	0.481	6.10	0.00
11	Household in urban area (PT)	-0.202	0.134	-1.50	0.13
12	Distance [km] (SM)	-0.204	0.0505	-4.04	0.00
13	No. of bikes in household (SM)	0.390	0.0607	6.43	0.00

Summary statistics

Number of observations = 1686

Number of excluded observations = 579

Number of estimated parameters = 13

$$\begin{aligned} \mathcal{L}(\beta_0) &= -1337.224 \\ \mathcal{L}(\hat{\beta}) &= -880.350 \\ -2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})] &= 913.749 \\ \rho^2 &= 0.342 \\ \bar{\rho}^2 &= 0.332 \end{aligned}$$

Table 3: Estimation results for the logit base model.

4.3.2 Multiple Indicator Solution method

Table 4 shows the estimation results of using the MIS methodology when there is an interaction between travel time and the *car loving* attitude. The approach introduced in Section 3.3 is used. All the parameters that appear also in the logit can be interpreted in a similar way, except for travel time by car. The likert flexibility indicator can take values from 1 to 5, so the travel time parameter is in the range $(-0.0777 + 1 \cdot 0.0121, -0.0777 + 5 \cdot 0.0121) = (-0.0656, -0.0172)$, which includes the travel time parameter that is obtained in the logit model. The β_δ parameter does not have a direct behavioral interpretation, but is derived by the mathematical formulation.

It is introduced in equation (14).

A likelihood ratio test between the logit and the MIS shows that they are not statistically equivalent:

$$-2(\mathcal{L}_r(\hat{\beta}) - \mathcal{L}_u(\hat{\beta})) = -2(-880.350 - (-864.915)) = 30.87 \geq 5.99 = \chi^2_{2,0.05}.$$

Therefore, the model with the MIS correction is preferred.

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC (PT)	1.06	0.398	2.66	0.01
2	Travel time [min] (Car)	-0.0777	0.0167	-4.66	0.00
3	Travel time [min] (PT)	-0.00905	0.00182	-4.97	0.00
4	ASC (Car)	0.488	0.446	1.09	0.27
5	No. of children in household (Car)	0.169	0.0710	2.38	0.02
6	Number of cars in household (Car)	0.866	0.147	5.88	0.00
7	$\frac{\text{Marginal cost}}{\text{Income}}$	-0.330	0.0914	-3.61	0.00
8	Work related trip (Car)	-0.708	0.132	-5.36	0.00
9	French speaking (Car)	0.998	0.173	5.75	0.00
10	Student (PT)	2.78	0.457	6.07	0.00
11	Household in urban area (PT)	-0.240	0.136	-1.76	0.08
12	Distance [km] (Slow modes)	-0.206	0.0509	-4.05	0.00
13	No. of bikes in household (SM)	0.383	0.0605	6.34	0.00
14	β_δ (Car)	0.348	0.177	1.97	0.05
15	Likert flex. ind. × travel time [min] (Car)	0.0121	0.00348	3.48	0.00

Summary statistics

Number of observations = 1686

Number of excluded observations = 579

Number of estimated parameters = 15

$$\begin{aligned} \mathcal{L}(\beta_0) &= -1337.224 \\ \mathcal{L}(\hat{\beta}) &= -864.915 \\ -2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})] &= 944.619 \\ \rho^2 &= 0.353 \\ \bar{\rho}^2 &= 0.342 \end{aligned}$$

Table 4: Estimation results for the MIS method.

4.3.3 Integrated Choice and Latent Variable method

Finally, an ICLV model is estimated. Results are shown in Table 5. Parameters 1-13 can be interpreted as in the case of the logit (except for the travel time by car, that is discussed below). In order to understand the rest of the parameters, the structural and measurement equations are introduced. The measurement equation for the car-loving attitude is defined as follows:

$$\text{Car loving} = \eta_{Carloving} + \omega, \tag{18}$$

where $\omega \sim \mathcal{N}(0, \sigma^2)$ and $\eta_{Carloving}$ is a parameter to estimate. In a classical ICLV approach this structural equation could be more complex. In the case study we consider it as shown in Equation 18 so that the results can be compared to those of the MIS method.

The measurement equations are as follows:

$$t \cdot I_1 = \alpha_1 + \lambda_1 \cdot t \cdot \text{Car loving} + \omega_1, \tag{19}$$

$$t \cdot I_2 = \alpha_2 + \lambda_2 \cdot t \cdot \text{Car loving} + \omega_2, \tag{20}$$

where $\omega_1 \sim \mathcal{N}(0, \sigma_1^2)$ and $\omega_2 \sim \mathcal{N}(0, \sigma_2^2)$. For identification reasons, α_1 is normalized to 0, and λ_1 and σ_1 to 1.

Parameter 14, corresponding to the interaction between *Car loving* and travel time, is positive, as expected.

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC (PT)	1.05	0.391	2.69	0.01
2	Travel time [min] (Car)	-4.08	0.671	-6.08	0.00
3	Travel time [min] (PT)	-0.548	0.104	-5.29	0.00
4	ASC (Car)	0.0870	0.421	0.21	0.84
5	No. of children in household (Car)	0.199	0.0692	2.87	0.00
6	No. of cars in household (Car)	1.09	0.121	9.00	0.00
7	$\frac{\text{Marginal cost}}{\text{Income}}$	-0.346	0.0890	-3.89	0.00
8	Work related trip (Car)	-0.703	0.129	-5.45	0.00
9	French speaking (Car)	0.963	0.171	5.65	0.00
10	Student (PT)	3.38	0.433	7.79	0.00
11	Household in urban area (PT)	-0.216	0.134	-1.62	0.11
12	Distance [km] (SM)	-0.206	0.0500	-4.11	0.00
13	No. of bikes in household (SM)	0.374	0.0598	6.25	0.00
14	Car loving × travel time [min] (Car)	0.870	0.182	4.78	0.00
15	$\eta_{Carloving}$	2.68	0.0735	36.42	0.00
16	σ	0.589	0.0176	33.50	0.00
17	α_2	0.000575	0.00766	0.08	0.94
18	λ_2	1.53	0.0453	33.88	0.00
19	σ_2	0.142	0.0189	7.49	0.00

Summary statistics

Number of observations = 1686

Number of excluded observations = 579

Number of estimated parameters = 19

$$\begin{aligned}
\mathcal{L}(\beta_0) &= -23121.351 \\
\mathcal{L}(\hat{\beta}) &= -4545.965 \\
-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})] &= 37150.773 \\
\rho^2 &= 0.803 \\
\bar{\rho}^2 &= 0.803
\end{aligned}$$

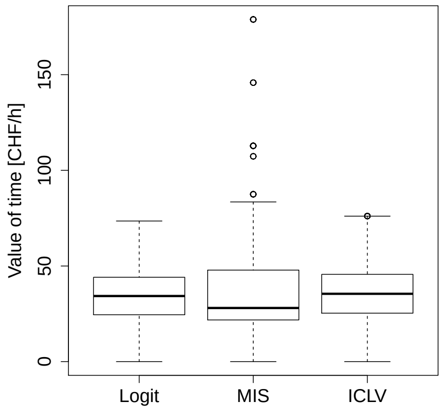
Table 5: Estimation results for the ICLV method.

4.3.4 Comparison of the methodologies: value of time

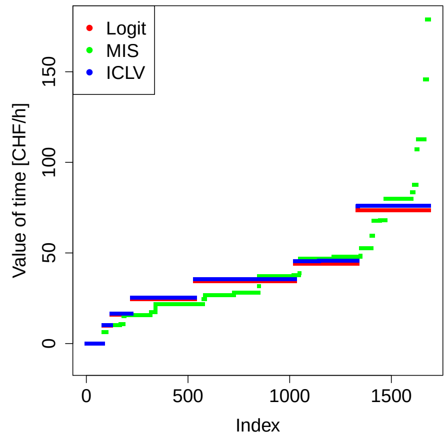
In this section the value of time (VOT) estimates are compared across the three methods presented above. The software Biogeme (Bierlaire, 2003) is also used for the simulation of these estimates. It gives as an output the value of the point estimate for each respondent.

Figure 2(a) shows a boxplot containing the disaggregate values of VOT of the respondents. We can see that the results obtained with the logit model have a lower spread compared to those of MIS and ICLV, which is expected since the *car loving* attitude is not taken into account. These values have a wider spread that those found by Axhausen *et al.* (2008), that range between 25 and 50 CHF/h for car users, depending on their trip purpose.

Figure 2(b) is an alternative representation of the same values, where the VOT have been reordered from the lowest to the highest value. We can see that for the logit model we obtain six different values of the VOT, one for each level of income. The results obtained for the ICLV are very similar, since the structural equation is given only by the mean plus an error term (see Equation (18)). Finally, for the MIS we obtain 30 different values, one per level of income and per answer to the likert indicator. The higher rate an individual gave to the statement *With my car I can go whenever and wherever*, the lower is his/her VOT. This is in line with what is expected, since a *car lover* is willing to pay less to save a minute of travel time by car compared to a someone with lower affection towards car. This is better seen in Figure 3, which is a graphical representation of the VOT for each of the *car loving* and income levels. The value of time for the category of low income and low *car loving* attitude is equal zero since none of these respondents has access to car. It is interesting to notice that the diagonals of this rectangle have almost the same value of time. For example, an individual with a monthly income of 5,000 CHF that gave the lowest value to the flexibility likert indicator has the same VOT than a person with a monthly income of 7,000 CHF that rated the indicator with the second value, and the same as an individual with a monthly income of 9,000 CHF that answered with a 3 out of 5 to the flexibility likert indicator. As expected the highest value of time corresponds to the respondents with the highest income and that gave a lower value to the flexibility likert indicator. The VOT decreases as income level decreases and as *car lovingness* –represented by the indicator– increases. In this sense, it is interesting to see how a respondent with an income level of at least 15,000 CHF per month has the same VOT as a respondent with a monthly income of 3,250 CHF if the first one rated the indicator with a 5 out of 5, and the second with a 1 out of 5.



(a) Boxplot of the VOT for the different methodologies.



(b) Plot of the ordered VOT for the different methodologies.

Figure 2: Representation of the VOT [CHF/h] for car.

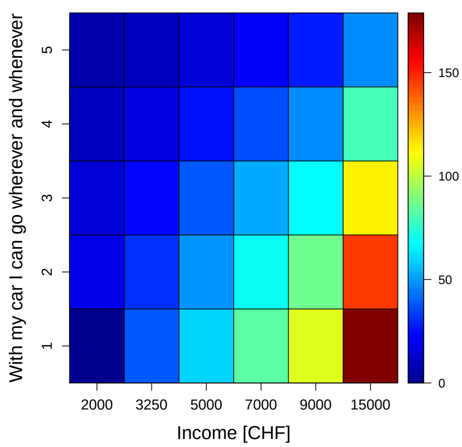


Figure 3: Representation of the VOT [CHF/h] for car per income and attitude level using the MIS method.

4.3.5 Comparison of the methodologies: travel time elasticity

The elasticity of travel time represents the percentage of variation in the probability of choosing an alternative following an increase of one percent in the travel time of this alternative.

Table 6 shows the weighted average of travel time elasticity (TE) for both the car and the public transportation alternatives for each of the three methodologies: a logit model, a model with the MIS correction and an ICLV model. Note that to compute the aggregate indicators of demand, the observations have to be weighted to coincide with the real population. Weights calculated by Atasoy *et al.* (2013) by age, gender and education level using the iterative proportional fitting algorithm are used.

In all the cases it is negative, as expected, meaning that an increase of travel time in a transportation mode decreases the probability of choosing it. It is also observed that the time elasticity for public transportation is larger in absolute value than that of car. This is not what is expected from the parameter estimates, since in Table 3 it can be seen that the parameter related to travel time for public transportation is smaller in absolute value than the parameter related to travel time by car. It becomes clearer by looking at the formula of the elasticity of travel time for an alternative i:

$$E_{t_{in}}^{P_n(i)} = \frac{\partial P_n(i)}{\partial t_{in}} \frac{t_{in}}{P_n(i)}, \tag{21}$$

where $P_n(i)$ is the probability of respondent n to choose alternative i with $i \in \{\mathrm{Car}, \mathrm{PT}\}$ and t_{in} is the travel time for respondent n and alternative$i$. As shown in Figure 1, travel time by public transportation is usually longer than by car, so this results in the mean time elasticity for public transportation being larger in absolute value than the mean time elasticity for car.

	Logit	MIS	ICLV
Car	-0.37	-0.48	-0.43
PT	-0.96	-0.99	-0.98

Table 6: Weighted average of the travel time elasticity for car and public transportation for each of the methodologies used.

We can see that the logit model underestimates the time elasticity towards car compared to both the MIS and the ICLV methodologies. Indeed, a 1% change in travel time by car will have an impact of -0.37% on the probability of choosing car, according to the logit model while after correcting for endogeneity with either the MIS or the ICLV methodologies we see that the decrease would be between 0.43% and 0.48%.

It is also interesting to look at the distribution of elasticities across the population, rather than the mean value. Figure 4 shows the boxplots across the three different methodologies. Since the spread is very wide – the minimum values are -13.5, -31.6 and -14,4 for the logit, MIS and ICLV values respectively– the boxplot is zoomed in the range $(-1, 0)$. The red cross represents the weighted mean value of TE. We can see that the spread of the boxplot without taking into account the outliers is larger for the ICLV methodology, due to the error terms in the structural and measurement equations. The shape is similar for the MIS and the logit models, but as discussed above, the average is not, and the tail of the distribution, related to the minimum values, is a lot more negative for the MIS methodology than for the ICLV and the logit, capturing better the extreme values.

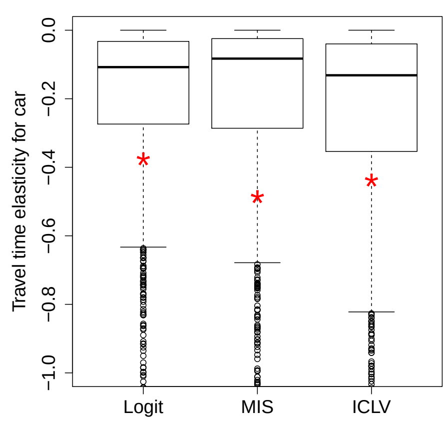


Figure 4: Boxplot of the TE for the different methodologies with a red cross representing the mean value.

5 Conclusions and future work

We have shown that the Multiple Indicator Solution can also be applied in discrete choice models in the presence of interactions between observed and unobserved attributes in the utility function. Moreover we have tested this methodology with a case study using real data collected in Switzerland. This is the first application of the MIS methodology with revealed preference data. The estimation results obtained are comparable to what is obtained by applying the same correction using the ICLV methodology, and the values of time obtained have larger spread than the results found in the literature since we are taking into account both income and the *car loving* attitude. The distribution of demand indicators such as value of time and time elasticity are also studied. Results reveal that the logit model underestimates the mean travel time elasticity for car compared to both the ICLV and the MIS method. Thanks to the MIS method we can also derive the VOT for different levels of *car lovingness* and income which also reveals interesting results. Moreover, a likelihood ratio test shows that the model with the MIS correction is significantly better than the logit model. In conclusion, the MIS performs as the ICLV or better, and is easier and faster to estimate. The purpose of this case study is to show that the MIS method is operational and that it can be adapted to model interactions between observed and unobserved attributes.

However, the MIS methodology is not free of limitations. An important limitation is that an indicator, as well as the residuals of a regression, appear directly in the utility function. How to do forecasting using this methodology is therefore not trivial. As mentioned, a possibility is to estimate a measurement equation for the unobserved indicators as a function of socioeconomic characteristics of the respondent, and then use these in the utility function.

The difficulty of using the MIS for forecasting might not be a problem if the interest of the application is to compute trade-offs such as VOT estimates, or elasticities. From a modeling point of view, the MIS method is a logit model with a correction factor. Therefore it has a closed form, and it is computationally a lot faster than the ICLV approach (the estimation time is of less than a second for the MIS method and of around 5 minutes for the ICLV). A potential solution when the model is to be used for forecasting would be to use the MIS approach to identify endogeneity and to find a good model specification, and then apply the ICLV method with the same specification and indicators once it is confirmed.

References

Abou-Zeid, Maya, Ben-Akiva, Moshe, Bierlaire, Michel, Choudhury, Charisma, & Hess, Stephane. 2010. Attitudes and Value of Time Heterogeneity. *Pages 523–545 of:* de Voorde, Eddy Van, & Vanelslander, Thierry (eds), *Applied Transport Economics A Management and Policy Perspective*. de boeck.

Ajzen, Icek. 2001. Nature and Operation of Attitudes. *Annual Review of Psychology*, **52**(1), 27–58.

Amemiya, Takeshi. 1978. The Estimation of a Simultaneous Equation Generalized Probit Model. *Econometrica*, **46**(5), 1193–1205.

Atasoy, Bilge, Glerum, Aurélie, & Bierlaire, Michel. 2013. Attitudes towards mode choice in Switzerland. *disP - The Planning Review*, **49**(2), 101–117.

Axhausen, Kay W., Hess, Stephane, König, Arnd, Abay, Georg, Bates, John J., & Bierlaire, Michel. 2008. Income and distance elasticities of values of travel time savings: New Swiss results. *Transport Policy*, **15**(3), 173–185.

Baum, Christopher F. 2006. *An Introduction to Modern Econometrics Using Stata*. Stata Press.

Ben-Akiva, M., & Boccara, B. 1987. Integrated framework for travel behavior analysis. *In: International Association of Travel Behavior Research (IATBR) Conference, Aix-en-Provence, France.*

Ben-Akiva, Moshe, Bolduc, Denis, & Bradley, Mark. 1993. Estimation of travel choice models with randomly distributed values of time. *Transportation Research Record*.

Berry, Steve, Linton, Oliver B., & Pakes, Ariel. 2004. Limit Theorems for Estimating the Parameters of Differentiated Product Demand Systems. *The Review of Economic Studies*, **71**(3), 613–654.

Berry, Steven, Levinsohn, James, & Pakes, Ariel. 1995. Automobile Prices in Market Equilibrium. *Econometrica*, **63**(4), 841–890.

Bierlaire, Michel. 2003. BIOGEME: a free package for the estimation of discrete choice models. *In: Proceedings of the Swiss Transport Research Conference*.

Bierlaire, Michel, Curchod, Anne, Danalet, Antonin, Doyen, Etienne, Faure, Prisca, Glerum, Aurélie, Kaufmann, Vincent, Tabaka, Kamila, & Schuler, Martin. 2011. *Projet de recherche sur la mobilité combinée, Rapport définitif de l'enquête de préférences révélées*. Tech. rept.

Bollen, Kenneth A. 1989. A New Incremental Fit Index for General Structural Equation Models. *Sociological Methods & Research*, **17**(3), 303–316.

Crawford, Gregory S. 2000 (July). *The Impact of the 1992 Cable Act on Household Demand and Welfare*. SSRN Scholarly Paper ID 235290. Social Science Research Network, Rochester, NY.

Dupuit, Jules. 1844. On the measurement of the utility of public works. *International Economic Papers*, **2**(1952), 83–110.

Dupuit, Jules. 1849. On tolls and transport charges. *Pages 7–31 of: Annales des ponts et chaussées*, vol. 11.

Everitt, B. S. 1984. *An Introduction to Latent Variable Models*. Dordrecht: Springer Netherlands.

Fosgerau, Mogens. 2006. Investigating the distribution of the value of travel time savings. *Transportation Research Part B: Methodological*, **40**(8), 688–707.

Glerum, Aurélie, Atasoy, Bilge, & Bierlaire, Michel. 2014. Using semi-open questions to integrate perceptions in choice models. *Journal of Choice Modelling*, **10**(Mar.), 11–33.

Golob, Thomas F. 2001. Joint models of attitudes and behavior in evaluation of the San Diego I-15 congestion pricing project. *Transportation Research Part A: Policy and Practice*, **35**(6), 495–514.

Goolsbee, Austan, & Petrin, Amil. 2004. The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV. *Econometrica*, **72**(2), 351–381.

Grange, Louis de, González, Felipe, Vargas, Ignacio, & Troncoso, Rodrigo. 2015. A Logit Model With Endogenous Explanatory Variables and Network Externalities. *Networks and Spatial Economics*, **15**(1), 89–116.

Guevara, C. A., & Polanco, D. 2013. Correcting for endogeneity without instruments in discrete choice models: the multiple indicator solution. *In: Proceedings of the Third International Choice Modelling Conference*.

Guevara, C. Angelo. 2015. Critical assessment of five methods to correct for endogeneity in discrete-choice models. *Transportation Research Part A: Policy and Practice*, **82**(Dec.), 240–254.

Guevara, C. Angelo, & Polanco, Daniel. 2016. Correcting for endogeneity due to omitted attributes in discrete-choice models: the multiple indicator solution. *Transportmetrica A: Transport Science*, **12**(5), 458–478.

Guevara-Cue, Cristian Angelo. 2010. *Endogeneity and Sampling of Alternatives in Spatial Choice Models*. Thesis, Massachusetts Institute of Technology. Thesis (Ph. D.)–Massachusetts Institute of Technology, Dept. of Civil and Environmental Engineering, 2010.

Hausman, Jerry A. 1978. Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, 1251–1271.

Heckman, James J. 1978 (May). *Dummy Endogenous Variables in a Simultaneous Equation System*. Working Paper 177. National Bureau of Economic Research.

Heckman, James J., & Robb Jr., Richard. 1985. Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics*, **30**(1-2), 239–267.

Hess, Stephane, & Axhausen, Kay W. 2004. *Checking our assumptions in value-of-travel-time modelling: Recovering taste distributions*. Tech. rept. CTS Working Paper, Centre for Transport Studies, Imperial College London.

Jiang, Renna, Manchanda, Puneet, & Rossi, Peter E. 2009. Bayesian analysis of random coefficient logit models using aggregate data. *Journal of Econometrics*, **149**(2), 136–148.

Koppelman, Frank S., & Hauser, John R. 1978. Destination choice behavior for non-grocery-shopping trips. *Transportation Research Record*.

Lancaster, Tony. 2004. *Introduction to Modern Bayesian Econometrics*. Wiley.

Louviere, Jordan, Train, Kenneth, Ben-Akiva, Moshe, Bhat, Chandra, Brownstone, David, Cameron, Trudy Ann, Carson, Richard T., Deshazo, J. R., Fiebig, Denzil, & Greene, William. 2005. Recent progress on endogeneity in choice modeling. *Marketing Letters*, **16**(3-4), 255–265.

McFadden, Daniel. 1986. The Choice Theory Approach to Market Research. *Marketing Science*, **5**(4), 275–297.

Nevo, Aviv. 2001. Measuring Market Power in the Ready-to-Eat Cereal Industry. *Econometrica*, **69**(2), 307–342.

Newey, Whitney K. 1985. Semiparametric Estimation of Limited Dependent Variable Models with Endogenous Explanatory Variables. *Annales de l'inséé*, July, 219–237.

Newey, Whitney K. 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. *Journal of Econometrics*, **36**(3), 231–250.

Office fédéral de la statistique, Office fédéral du développement territorial. 2012. *La mobilité en Suisse: Résultats du microrecensement mobilité et transports 2010*. Neuchâtel: Office fédéral de la statistique.

Olson, James M., & Zanna, Mark P. 1993. Attitudes and Attitude Change. *Annual Review of Psychology*, **44**(1), 117–154.

Outwater, Maren, Castleberry, Steve, Shiftan, Yoram, Ben-Akiva, Moshe, Shuang Zhou, Yu, & Kuppam, Arun. 2003. Attitudinal Market Segmentation Approach to Mode Choice and Ridership Forecasting: Structural Equation Modeling. *Transportation Research Record: Journal of the Transportation Research Board*, **1854**(Jan.), 32–42.

Park, Sungho, & Gupta, Sachin. 2012. Handling endogenous regressors by joint estimation using copulas. *Marketing Science*, **31**(4), 567–586.

Petrin, Amil, & Train, Kenneth. 2009. A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research*, **47**(1), 3–13.

Proussaloglou, Kimon E., & Koppelman, Frank S. 1989. Use of travelers' attitudes in rail service design. *Transportation Research Record*.

Train, Kenneth. 2003. *Discrete Choice Methods with Simulation*. Cambridge University Press.

Train, Kenneth E., & Winston, Clifford. 2007. Vehicle choice behavior and the declining market share of US automakers. *International Economic Review*, **48**(4), 1469–1496.

Villas-Boas, J. Miguel, & Winer, Russell S. 1999. Endogeneity in Brand Choice Models. *Management Science*, **45**(10), 1324–1338.

Vredin Johansson, Maria, Heldt, Tobias, & Johansson, Per. 2006. The effects of attitudes and personality traits on mode choice. *Transportation Research Part A: Policy and Practice*, **40**(6), 507–525.

Walker, Joan L., Ehlers, Emily, Banerjee, Ipsita, & Dugundji, Elenna R. 2011. Correcting for endogeneity in behavioral choice models with social influence variables. *Transportation Research Part A: Policy and Practice*, **45**(4), 362–374.

Walker, Joan Leslie. 2001. *Extended discrete choice models: integrated framework, flexible error structures, and latent variables*. Ph.D. thesis, Massachusetts Institute of Technology.

Wood, Wendy. 2000. Attitude Change: Persuasion and Social Influence. *Annual Review of Psychology*, **51**(1), 539–570.

Wooldridge, Jeffrey M. 2010. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Yang, Sha, Chen, Yuxin, & Allenby, Greg M. 2003. Bayesian Analysis of Simultaneous Demand and Supply. *Quantitative Marketing and Economics*, **1**(3), 251–275.