Correcting for endogeneity using the EMIS method: a case study with revealed preference data

Anna Fernández Antolín

C. Angelo Guevara-Cue

Michel Bierlaire

Transportation and Mobility Laboratory, ENAC, EPFL

April 2015

STRC

15th Swiss Transport Research Conference

Monte Verità / Ascona, April 15 – 17, 2015

Transportation and Mobility Laboratory, ENAC, EPFL

Correcting for endogeneity using the EMIS method: a case study with revealed preference data

Anna Fernández Antolín,
Michel Bierlaire
Transport and Mobility Laboratory
Ecole Polytechnique Fédérale de Lausanne
Station 18
CH-1015 Lausanne
phone: +41-21-693 93 29
fax: +41-21-693 80 60
{anna.fernandezantolin,michel.bierlaire}@epfl.ch

C. Angelo Guevara-Cue

Universidad de los Andes, Chile

phone: +56226181622
fax:
aguevara@uandes.cl

April 2015

Abstract

Endogeneity is an important issue that often arises in discrete choice models leading to biased estimates of the parameters. We propose the extended multiple indicator solution (EMIS) methodology to correct for it and exemplify it with a case study using revealed preference data about mode choice in Switzerland. The same data is used as one of the first applications of the multiple indicator solution (MIS) method. These two methodologies - EMIS and MIS - are then compared to an integrated choice and latent variable model (ICLV), and a model without any correction. In order to compare the different methodologies between them, the value of time and the time elasticity in public transportation estimates are computed and reported for each of the methods.

Keywords

Discrete choice models, endogeneity, integrated choice and latent variable, multiple indicators solution, extended multiple indicators solution

1 Introduction

In interurban mode choice modeling, different alternatives such as bus, private modes and soft modes compete among them. Thus, a premium alternative will be more highly priced but, in compensation, it may offer shorter travel or waiting time, fewer transfers, larger space between seats and so on. The choice maker will take into consideration these amenities, but the researcher will usually be able to account only for some of them. A choice model omitting some of the various dimensions of the level of service will suffer of endogeneity.

Various methods have been proposed to address endogeneity. The purpose of this research is to integrate two of these methods: (i) the integrated choice and latent variable (ICLV) developed by Walker (2001) where a latent factor captures an unobserved qualitative attribute (e.g. the comfort of a transport mode), and (ii) the multiple indicator solution (MIS) introduced by Guevara and Polanco (2013). It relies on having a couple of suitable indicators, which are measured variables that depend on the latent variable that causes endogeneity, but not on other measured attributes.

The integration of both methodologies will be referred to as extended multiple indicator solution (EMIS) and is the main contribution of this research. This work is also one of the fist applications of the MIS method with real data. We present a comparative analysis between the use of ICLV, MIS and EMIS methods to account for endogeneity biases. We use a transportation mode choice case study in Switzerland (Glerum *et al.* (2014)) to show how the omission of comfort in public transportation biases the estimation of the parameter associated to time. The same case study is then used to illustrate the theoretical framework that integrates psychological variables and endogeneity correction in choice models, and how the methods under study perform.

The idea behind the combined method is to consider the MIS method to account for the heterogeneity of the omitted factors between modes and individuals, and to use the latent variables approach to account for the heterogeneity in the perceptions among individuals.

2 Literature review

Kraus (1991) initiated the study of discomfort in public transportation in the academic area. He took into account the passengers' relative dislike for standing. Researchers also use crowding in public transportation to model discomfort (de Lapparent and Koning (2015), Wardman and Whelan (2011)). However, in practice we do not always have this data (crowding, standing vs seating) in demand modeling. Therefore, the model will suffer from misspecification due to the omitted variable which will cause endogeneity.

Louviere *et al.* (2005) present the 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. The focus of this paper is in endogenous explanatory variables. The methods presented below differ in many aspects, including their tractability and their applicability to different contexts.

A very used methodology is the BLP (Berry *et al.* (1995), Berry *et al.* (2004)) which receives its name from the names of the authors. This approach consists on removing the endogeneity from the non-linear choice model and dealing with it in linear regressions. This is done by adding an alternative specific constant (ASC) for each product and each market. By doing this, the instrumental variables method can be used in the linear regression. A description of the instrumental variable methodology can be found in most of the basic econometric textbooks such as Baum (2006) or Lancaster (2004). Guevara-Cue (2010) describes in his thesis why it is more complex to deal with endogeneity in discrete choice models compared to linear models. The problem encountered when trying to correct for endogeneity in non-linear models is that 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 will impact the peer group will also influence the decision maker and this will cause correlation between the field effect variable and the error. Train and 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 and Petrin (2004) where they examine the direct broadcast satellites as a competitor to cable TV.

A second very used 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 and Robb Jr. (1985). Petrin and 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 and Winer (1999) in brand choice models and Park and Gupta (2012). In particular, Park and 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 Control and Latent Variable (ICLV) approach (Walker and Ben-Akiva (2002), Walker (2001), Glerum *et al.* (2014)), 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 (2002) in the context of linear models and generalized in Guevara and Polanco (2013) for discrete choice. As opposed to the control-function method, the MIS method does not need of instrumental variables. Instead, it uses indicators to introduce a factor of correction in the choice model in order to obtain unbiased estimators.

There are also other methods, but are less used because they are 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.

3 EMIS method

This section consists on the theoretical presentation of the proposed methodology. As stated in Section 1, this research aims at integrating two methods mitigating parameter biases resulting from the omission of unobserved factors in discrete choice models. The ICLV approach is briefly described in section 3.1, the MIS method is derived in subsection 3.2 and finally our contribution is the integration of both: the extended multiple indicator solution (EMIS), developed in subsection 3.3. Subsection 3.4 shows that this methodology can also be generalized to take into account interactions between an attribute and the unobserved factor.

3.1 Integrated choice and latent variable (ICLV) model framework

Let us consider an integrated choice and latent variable model where the choice of an alternative i depends on an economic factor t_{in} which is correlated with an unobserved attribute ξ_{in}, and on a set of other explanatory factors x_{in}. The utility of this alternative is specified as follows

$$U_{in} = \mathrm{ASC}_i + \beta_x x_{in} + \beta_t t_{in} + \beta_\xi \xi_{in} + e_{in}, \tag{1}$$

where ASC_i, β_x β_t and β_ξ are parameters to estimate and e_{in} is a random error term. We assume that t_{in} is correlated with ξ_{in}, so that the above model is endogenous. For instance, i could be a transport mode alternative, where the travel time t_{in} could be correlated with a variable representing the perception of comfort in a transport mode ξ_{in}. The structural equation of the latent variable model is given as follows

$$\xi_{in} = \eta_0 + \eta s_n + e_\xi, \tag{2}$$

where η_0, η are (vectors of) parameters to estimate, s_n is a vector of socio-economic characteristics of the respondent n, and e_ξ is an error term.

The measurement model specifies the following k measurement equations

$$I_{kin} = \alpha_{k0} + \alpha_{k\xi}\xi_{in} + e_{I_{kin}}, \tag{3}$$

where α_{k0} and $\alpha_{k\xi}$ are parameters to estimate, and $e_{I_{kin}}$ is a random error term.

The ICLV method has the drawback that it does not fully capture endogeneity bias. For example, if we now assume that variable ξ_{in} represents comfort in a mode i instead of the mode's perception of comfort, equation (2) could also depend on attributes of the alternative. Therefore it is unclear that e_ξ is uncorrelated with the other explanatory variables in the utility (as it should be to obtain consistent parameters). The method proposed below addresses this issue.

3.2 MIS method

Instead of using the ICLV method to account for the omission of ξ_{in}, let us now consider a model with the same formulation of utility as in equation (1).

We assume that we have two indicators I_{1in} and I_{2in} which are related to the omitted variable ξ_{in} by the following relations

$$I_{1in} = \alpha_0 + \alpha_\xi \xi_{in} + e_{I_{1in}}, \tag{4}$$

$$I_{2in} = \delta_0 + \delta_\xi \xi_{in} + e_{I_{2in}}. \tag{5}$$

Given equation (4), we can replace ξ_{in} by $\frac{I_{1in}}{\alpha_\xi} - \frac{\alpha_0}{\alpha_\xi} - \frac{e_{I_{1in}}}{\alpha_\xi}$ in equation (1), which becomes

$$U_{in} = \mathrm{ASC}_i - \theta_\xi \alpha_0 + \beta_x x_{in} + \beta_t t_{in} + \theta_\xi I_{1in} - \theta_\xi e_{I_{1in}} + e_{in}, \tag{6}$$

where we have defined the following relation: $\theta_\xi = \frac{\beta_\xi}{\alpha_\xi}$. The above model is still endogeneous since I_{1in} is correlated with $e_{I_{1in}}$. We will hence apply the control function method (similarly as in Guevara-Cue (2010)) and use I_{2in} as an instrument for I_{1in}. Since I_{2in} is correlated with I_{1in} by equations (4) and (5) but uncorrelated with $e_{I_{1in}}$, we can define the following relations

$$I_{1in} = \gamma_0 + \gamma_1 I_{2in} + \gamma_t t_{in} + \gamma_x x_{in} + \delta_{in}, \tag{7}$$

$$e_{I_{1in}} = \beta_\delta \delta_{in} + \nu_{in}, \tag{8}$$

where δ_{in} captures the part of $e_{I_{1in}}$ which is correlated with I_{1in} and ν_{in} is an error term. Given these equations, the utility function in equation (6) can be rewritten as follows

$$U_{in} = \tilde{\mathrm{ASC}}_i + \beta_x x_{in} + \beta_t t_{in} + \theta_\xi I_{1in} + \theta_\delta \delta_{in} + \tilde{e}_{in}, \tag{9}$$

where we have defined new variables, that is, $\tilde{\mathrm{ASC}}_i := \mathrm{ASC}_i - \theta_\xi \alpha_0$ and $\theta_\delta := -\theta_\xi \beta_\delta$. In the above equation, we have the following error term $\tilde{e}_{in} := -\theta_\xi \nu_{in} + e_{in}$.

The model with the MIS correction is estimated in two stages. First δ_{in} is obtained by taking the residual values of equation (7). Second, all parameters of equation (9) are estimated by maximum likelihood. Note that using the full information maximum likelihood would render a one-stage estimation possible.

3.3 EMIS method

The two methods just described are designed to address different goals. On the one hand the purpose of the ICLV model is to assess the impact of a perceptional variable on choice. On the other hand the aim of the model with the MIS correction is to correct endogeneity biases. We propose here a more complete framework which achieves both goals.

The utility of equation (9) is assumed to be the same, except that we now consider a random coefficient for θ_ξ. More precisely, this coefficient's mean is assumed to be a function of socioeconomic characteristics of the respondent. We hence assume the following relation

$$\theta_\xi \sim \text{Distr}(\eta_0 + \eta s_n, \sigma), \tag{10}$$

where Distr could typically be a normal distribution, and parameters η_0, η and σ should be estimated to assess the heterogeneity of the impact of the latent psychological variable on choice. The above model hence combines a measurement of the effect of an unobserved variable and a correction for endogeneity.

Note that the mean of θ_ξ can not depend on the attributes that enter the utility function due to the definition of $\tilde{e}_{in}$. If it did, the model presented would still be endogenous. Also, the assumptions that have to be done on the distributions of ν_{in} and e_{in} in order to consider $\tilde{e}_{in} \sim$ EV(0,1) are unclear and are left for future work.

3.4 Interaction between travel time and comfort

In terms of specification of the utility of public transportation, there is evidence in the literature that suggests to use the interaction between travel time an comfort. This is discussed in de Lapparent and Koning (2015), Wardman and Whelan (2011). The results in de Lapparent and Koning (2015) show that there is evidence using statistical tools to show that not considering the non-linearity is better. However, they argue that such interaction is to be kept given that this is, a priory, more intuitive.

The methodology presented in Section 3.3 is also valid in this case, by adding one extra assumption of the indicators, as is shown below. In what follows we use the same notation that in the sections above.

The utility function can be expressed as

$$U_{in} = ASC_i + \beta_t t_{in} + \beta_x x_{in} + \beta_\xi t_{in}\xi_{in} + e_{in}. \tag{11}$$

The extra assumption that has to be done is that $t_{in}I_{in}$ is an indicator of $t_{in}\xi_{in}$. Thus, we have the following equations for the indicators

$$t_{in}I_{1in} = \alpha_0 + \alpha_\xi t_{in}\xi_{in} + e_{I_{1in}}, \tag{12}$$

$$t_{in}I_{2in} = \delta_0 + \delta_\xi t_{in}\xi_{in} + e_{I_{2in}}. \tag{13}$$

From equation (12) we obtain $\xi_{in} = (t_{in}I_{1in} - \alpha_0 - e_{I_{1in}})/\alpha_\xi t_{in}$. By substituting this expression in equation (11) and denoting $\theta_\xi = \frac{\beta_\xi}{\alpha_\xi}$ we obtain

$$U_{in} = ASC_i + \beta_t t_{in} + \beta_x x_{in} + \theta_\xi t_{in}I_{1in} - \theta_\xi\alpha_0 - \theta_\xi e_{I_{1in}} + e_{in}. \tag{14}$$

We now proceed to the control function stage of the MIS method. We regress one indicator (which is now $t_{in}I_{1in}$) on the other ($t_{in}I_{2in}$), obtaining the following expression

$$t_{in}I_{1in} = \gamma_0 + \gamma_1 t_{in}I_{2in} + \gamma_t t_{in} + \gamma_x X_{in} + \delta_{in}. \tag{15}$$

We can now write

$$e_{I_{1in}} = \beta_\delta\delta_{in} + \nu_{in}, \tag{16}$$

where β_δ captures all of $e_{I_{1in}}$ that is correlated with I_{1in} and ν_{in} is exogenous. Substituting equation (16) to (14) we obtain

$$U_{in} = (ASC_i - \theta_\xi\alpha_0) + \beta_t t_{in} + \beta_x x_{in} + \theta_\xi t_{in}I_{1in} - \theta_\xi\beta_\delta\delta_{in} - \theta_\xi\nu_{in} + e_{in}. \tag{17}$$

By 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

$$U_{in} = A\tilde{S}C_i + \beta_t t_{in} + \beta_x x_{in} + \theta_\xi t_{in}I_{1in} + \theta_\delta\delta_{in} + \tilde{e}_{in}, \tag{18}$$

where there is no endogeneity anymore.

4 Case study

This section presents the case study used to apply the different methodologies described in section 3. In section 4.1 there is a description of the data used. It is followed by the model specification in section 4.2. Finally, the results are presented in section 4.3

4.1 Data collection

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. The work by Glerum *et al.* (2014) consists in translating these qualitative indicators - the adjectives - to quantitative indicators. These are the indicators that will be used in this research. They can take continuous values between -2 and 2.

Figure 1 shows these indicators, that will be used for the ICLV, MIS and EMIS methods. The upper plots show the indicators itself (I_{1in} and I_{1in}, also denoted Adj1 and Adj2). Even if the values that they can take are between -2 and 2, most of them are in the range between -1 and 1. The lower plots in Figure 1 show $t_{in}I_{1in}, t_{in}I_{2in}$ which will be used as indicators for the interaction between comfort and travel time.

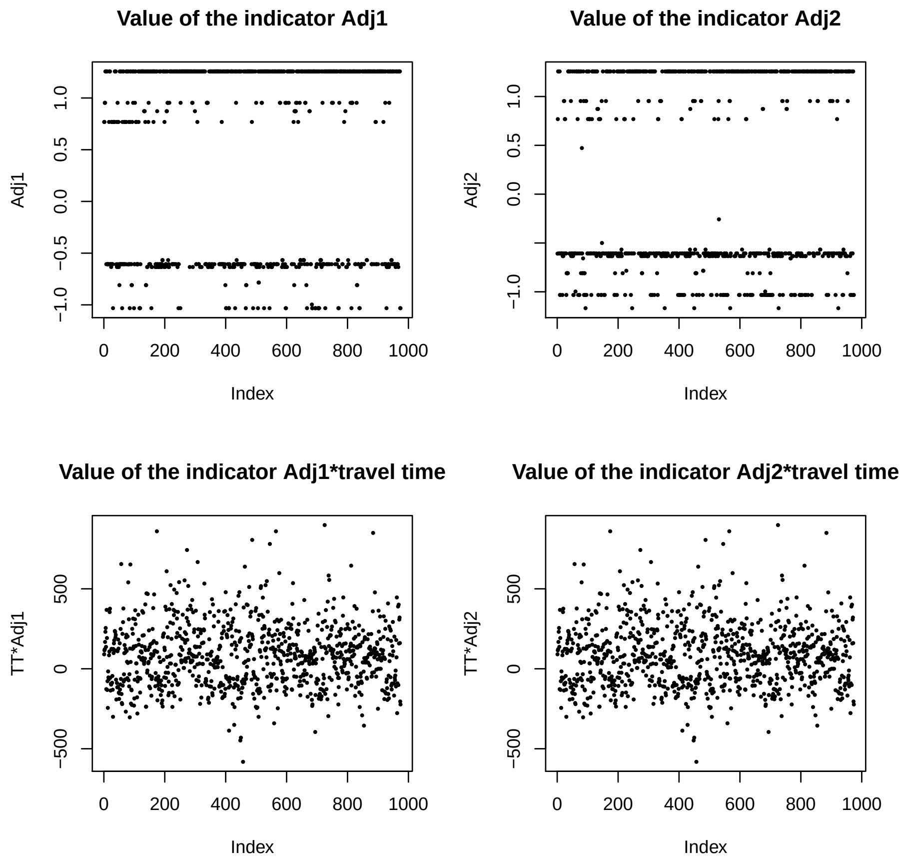

Figure 1: Upper left to lower right: Plots of the values of the indicators (Adj1 and Adj2) for each respondent; Plot of the indicators used in the model with interaction: the original indicators multiplied by the travel time of each respondent (tt· Adj1 and tt· Adj2).

The mode alternatives are public transportation (PT), private motorized modes (PMM) (car, motorbike, etc.) and slow modes (SM) (bike, walk). PMM will also be referred to as *Car*. Table 1 shows the market shares in the dataset for each of the three considered modes. These are the results after excluding the respondents who did not give an answer for the selected mode and who stated that went by car and not to have access to a car.

	PT	PMM	SM	Total
Observed market shares (%)	66	28	6	100
Number of observations	536	1249	114	1899

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).

4.2 Model specification

Table 2 shows the model specification used as the base model for the case study. It is a model with 19 parameters. In the slow modes utility function, only the distance of the trip is considered as an explanatory variable.

In the public transportation utility, there is the alternative specific constant (ASC), some socioeconomic variables related to the age of the respondent and the travel cards that s/he owns, as well as attributes of the mode such as cost and time. The parameter for cost is an alternative specific one, while the parameters related to travel time - which is interacted with the distance of the trip - are generic for both alternatives, but interacted with the occupation of the respondent.

In the car utility function there is also an ASC, two socioeconomic variables which are if the respondent is from a French speaking part of Switzerland or not and the number of cars in the respondent's household, and time and cost of the trip. The parameters related to the marginal cost of the trip are alternative specific and interacted with socioeconomic variables, while the one related to cost is analogous to how it appears in public transportation.

The specifications used for the other four models (MIS, MIS with interaction, EMIS and ICLV) are the same except for the parameters associated to each methodology.

Since not all the respondents answered the semi-open questions with adjectives related to comfort, some observations do not have associated indicators. In the dataset there are 831 observations with indicators and 1068 without. In order to use all the respondents for the estimation of the model, an extra parameter γ is introduced for the MIS and EMIS methods. In these methods, we define two utility functions for public transportation, one with the MIS/EMIS correction for the respondents that have associated indicators, and one without it for the rest. For those respondents without the correction, the parameter associated to travel time is shifted by this new parameter γ. Note that by adding only one γ to the parameters associated to travel time in public transportation we are assuming that the rest of parameters are not correlated with comfort in public transportation. If this was not the case, then we should exclude the respondents for which we do not have indicators. In this particular case study this latter approach results in several non-significant parameters. We therefore decided to apply the approach where two utility functions for public transportation were defined.

Parameter	Public transportation	Car	Slow modes
ASC_{PT}	1	0	0
β_1	Age < 45	0	0
β_2	Age $\in (45, 65)$	0	0
β_3	Marginal cost [CHF]	0	0
β_4	TimePT$\cdot$ log(1 + dist.(km))/1000, if full time worker	Time car$\cdot$ log(1 + dist.(km))/1000, if full time worker	0
β_5	TimePT$\cdot$ log(1 + dist.(km))/1000, if part time worker	Time car$\cdot$ log(1 + dist.(km))/1000, if part time worker	0
β_6	TimePT$\cdot$ log(1 + dist.(km))/1000, if other occupation	Time car$\cdot$ log(1 + dist.(km))/1000, if other occupation	0
β_7	Season tickets	0	0
β_8	Half fare travel card	0	0
β_9	Line related travel card	0	0
β_{10}	Area related travel card	0	0
β_{11}	Other travel cards	0	0
ASC_{car}	0	1	0
β_{12}	0	Number of cars	0
β_{13}	0	Gasoline cost [CHF] if HWH	0
β_{14}	0	Gasoline cost [CHF] if other	0
β_{15}	0	Gasoline cost [CHF] if male	0
β_{16}	0	French speaking	0
β_{17}	0	0	dist. (km)

Table 2: Base model specification.

4.3 Results

In what follows we present an analysis of the five models estimated and a comparison of the estimates of value of time (VOT) and time elasticity in public transportation. All the models are estimated using the software Biogeme (Bierlaire (2003)).

The five estimated models are:

Base model A logit model, including travel times, travel costs, distance, socio-economic characteristics of the respondent.

ICLV An ICLV model, which has the same specification as the logit model, but it additionally includes a latent variable capturing the perception of comfort in public transportation.

MIS A logit model including the MIS correction to remedy the omission of the variable capturing the perception of comfort in public transportation.

MIS with interaction A logit model including the MIS correction to remedy the omission of the variable capturing the perception of comfort in public transportation where an interaction between the omitted comfort and the travel time is considered.

EMIS A logit model including the MIS correction with the mixed parameter to both remedy the omission of the variable capturing the perception of comfort in public transportation and capturing the perception of comfort.

4.3.1 Base model

Table 3 shows the estimation results for the model specification defined in Table 2. The signs are in line with our expectations: the parameters associated to travel time, travel cost and distance are negative. Note that the parameter associated to travel time is interacted with distance and is generic for both utilities. Moreover, the disutility towards longer travel times is stronger for the segment of full time workers, then part time workers and finally of people with other occupations. However, the parameter associated to the marginal cost of public transportation has a p-value of 0.19 suggesting that it is only significantly different from zero at a 20 % significance level. This might be due to the fact that the only cost that is available is the marginal cost. For all the respondents that have travel cards, the marginal cost will be zero, making the modeling more difficult, as it can not be translated to a cost per trip. To capture this, parameters 8 to 12 have been introduced. They are all positive as expected meaning that having a public transportation travel card will increase the utility of using public transportation.

For the utility of car, the number of cars per household is positive, as expected. All else being equal, a larger number of cars in a respondent's household will translate in a higher utility of

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC_{PT}	-1.09	0.406	-2.69	0.01
2	Age 0-45 (PT)	-0.0151	0.00691	-2.18	0.03
3	Age 45-65 (PT)	0.0202	0.0110	1.84	0.07
4	Marginal cost [CHF] (PT)	-0.0136	0.0105	-1.30	0.19
5	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if full time job	-2.61	0.429	-6.08	0.00
6	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if other occupation	-0.880	0.363	-2.42	0.02
7	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if part time job	-2.06	0.573	-3.59	0.00
8	Season ticket dummy (PT)	2.92	0.322	9.06	0.00
9	Half fare travel card dummy (PT)	0.589	0.165	3.57	0.00
10	Line related travel card dummy (PT)	1.89	0.260	7.29	0.00
11	Area related travel card (PT)	2.76	0.247	11.18	0.00
12	Other travel cards dummy (PT)	1.57	0.260	6.06	0.00
13	ASC_{CAR}	-0.619	0.352	-1.76	0.08
14	Number of cars in household (Car)	0.733	0.107	6.85	0.00
15	Gasoline cost [CHF], if trip purpose HWH (Car)	-0.164	0.0347	-4.71	0.00
16	Gasoline cost [CHF], if trip purpose other (Car)	-0.0795	0.0275	-2.89	0.00
17	Gasoline cost [CHF], if male (Car)	-0.0370	0.0198	-1.87	0.06
18	French speaking (Car)	0.794	0.173	4.60	0.00
19	Distance [km] (Slow modes)	-0.203	0.0504	-4.02	0.00

Summary statistics

Number of observations = 1899

Number of estimated parameters = 19

$$\mathcal{L}(\beta_0) = -1032.079$$
$$\mathcal{L}(\hat{\beta}) = -922.117$$
$$-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})] = 219.924$$
$$\rho^2 = 0.107$$
$$\bar{\rho}^2 = 0.088$$

Table 3: Estimation results for the base model.

this respondent towards car. In a similar way as in the utility of public transportation, only the marginal cost of the trip is considered, which is the gasoline cost. The parameter associated to the marginal cost is more negative for home to work trips than for other types of trips. It can also be seen from the parameter estimates that male have higher disutility towards high prices of gasoline compared to women. All else being equal, respondents from the French speaking part of Switzerland have higher utility towards car compared to public transportation or soft modes. This can be due to the fact that the public transportation infrastructure is more developed in the German speaking part of Switzerland. Finally, all else being equal, respondents have higher preference towards soft modes and car than towards public transportation, with respondents below the age of 45 having a slightly lower utility for public transportation and respondents with ages between 45 and 65 having a slightly higher utility for public transportation compared to respondents over the age of 65.

For slow modes, the only explanatory variable is the distance of the trip. The parameter associated to it is negative, meaning that respondents will have a higher disutility towards slow modes as the distance increases.

4.3.2 ICLV method

Table 8 in the Appendix shows the estimation results for the integrated choice and latent variable methodology. The choice model specification is the same one used in the base model, plus one parameter capturing comfort. Surprisingly, this parameter is negative and significantly different from zero. The latent variable model used is the same as in Glerum *et al.* (2014) and a detailed interpretation of the model estimates can be found there.

4.3.3 MIS method

Table 4 shows the results for the estimation of the model with the MIS correction. Most of the coefficients can be interpreted as in the base model case. However, in this case there are two alternative specific constants for public transportation because two utility functions for public transportation are used, as explained in section 4.2. The three parameters relative to the correction are β_δ, γ and θ_ξ.

The parameter associated to the residuals of the regression (β_δ) is a result of the mathematical derivation, but can not be interpreted behaviorally. The parameter associated to the image of comfort in public transportation is positive, as expected, meaning that higher levels of comfort will translate in higher utilities towards public transportation. The value of γ is negative, also as expected. The behavioral interpretation of γ is that the parameters associated to travel time by public transportation are shifted by this value when endogeneity is not addressed, this is, for the respondents for which no indicators are available. In other words, when there is no correction for the omission of comfort in the model, the parameters related to travel time for public transportation are more negative than when we address the issue. The intuition behind this is is that when omitting comfort from the model, the travel time coefficient will absorb the disutility of both higher travel times and of discomfort. Therefore, when comfort is explicitly modeled, we expect the travel time parameter to be less negative.

However, the p-values associated to this parameters are higher than what would be desirable.

MIS method with interaction Table 9 in the Appendix shows the estimation results of the model with the MIS correction and considering an interaction between the image of comfort and the travel time by public transportation. This non-linearity is, a priory, more intuitive. However, in line with what de Lapparent and Koning (2015) describe, statistical tools show differently. In our case, the p-values of the parameters related to the MIS correction are higher, meaning that these parameters are less significant. Therefore, for the application of the EMIS method we

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC_{PT}	-1.07	0.410	-2.61	0.01
2	$AS\tilde{C}_{PT}$	-1.09	0.424	-2.58	0.01
3	Age 0-45 (PT)	-0.0166	0.00695	-2.38	0.02
4	Age 45-65 (PT)	0.0203	0.0110	1.85	0.06
5	Marginal cost [CHF] (PT)	-0.0142	0.0107	-1.33	0.18
6	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if full time job	-2.32	0.434	-5.35	0.00
7	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if other occupation	-0.575	0.381	-1.51	0.13
8	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if part time job	-1.93	0.581	-3.32	0.00
9	Season ticket dummy (PT)	2.84	0.322	8.83	0.00
10	Half fare travel card dummy (PT)	0.559	0.168	3.33	0.00
11	Line related travel card dummy (PT)	1.91	0.259	7.37	0.00
12	Area related travel card (PT)	2.79	0.248	11.22	0.00
13	Other travel cards dummy (PT)	1.52	0.266	5.71	0.00
14	ASC_{CAR}	-0.622	0.353	-1.76	0.08
15	Number of cars in household (Car)	0.745	0.106	7.02	0.00
16	Gasoline cost [CHF], if trip purpose HWH (Car)	-0.172	0.0380	-4.53	0.00
17	Gasoline cost [CHF], if trip purpose other (Car)	-0.0890	0.0314	-2.83	0.00
18	Gasoline cost [CHF], if male (Car)	-0.0387	0.0208	-1.87	0.06
19	French speaking (Car)	0.744	0.173	4.29	0.00
20	Distance [km] (Slow modes)	-0.203	0.0504	-4.02	0.00
21	Residuals of the regression (β_δ)	0.413	0.473	0.87	0.38
22	Correction coefficient for endogeneity (γ)	-0.429	0.272	-1.58	0.11
23	Image Comfort PT (θ_ξ)	0.202	0.137	1.47	0.14

Summary statistics
Number of observations = 1899
Number of estimated parameters = 23

$$\begin{aligned}\mathcal{L}(\beta_0) &= -1086.151\\ \mathcal{L}(\hat{\beta}) &= -917.799\\ -2[\mathcal{L}(\beta_0)-\mathcal{L}(\hat{\beta})] &= 336.703\\ \rho^2 &= 0.155\\ \bar{\rho}^2 &= 0.134\end{aligned}$$

Table 4: Estimation results for the MIS method.

decide to keep the model without the interaction. To study the effects of the interaction between travel time and comfort is considered as future work.

4.3.4 EMIS method

The last method applied to this case study was the EMIS method. It is the first application in the literature of this methodology. The model specification is the same as in the MIS method, except that θ_ξ is assumed to be distributed. The notation is the same used for equation 9 where $\tilde{\mathrm{ASC}}_i := \mathrm{ASC}_i - \theta_\xi\alpha_0$, $\theta_\delta := -\theta_\xi\beta_\delta$ and $\theta_\xi \sim \mathcal{N}(\eta_0, 1)$. The variance of θ_ξ is fixed to one because it could not be estimated.

We can see that the parameters not related to the methodology are in line with what is obtained for the base model. About the parameters related to the EMIS, we can see that β_δ and α_0 are

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC_{PT}	-1.07	0.410	-2.61	0.01
2	Age 0-45 (PT)	-0.0166	0.00695	-2.38	0.02
3	Age 45-65 (PT)	0.0203	0.0110	1.85	0.06
4	Marginal cost [CHF] (PT)	-0.0142	0.0107	-1.33	0.18
5	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if full time job	-2.32	0.434	-5.35	0.00
6	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if other occupation	-0.575	0.381	-1.51	0.13
7	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if part time job	-1.93	0.581	-3.32	0.00
8	Season ticket dummy (PT)	2.84	0.322	8.83	0.00
9	Half fare travel card dummy (PT)	0.559	0.168	3.33	0.00
10	Line related travel card dummy (PT)	1.91	0.259	7.37	0.00
11	Area related travel card (PT)	2.79	0.248	11.22	0.00
12	Other travel cards dummy (PT)	1.52	0.266	5.71	0.00
13	ASC_{CAR}	-0.622	0.353	-1.76	0.08
14	Number of cars in household (Car)	0.745	0.106	7.02	0.00
15	Gasoline cost [CHF], if trip purpose HWH (Car)	-0.172	0.0380	-4.53	0.00
16	Gasoline cost [CHF], if trip purpose other (Car)	-0.0890	0.0314	-2.83	0.00
17	Gasoline cost [CHF], if male (Car)	-0.0387	0.0208	-1.87	0.06
18	French speaking (Car)	0.744	0.173	4.29	0.00
19	Distance [km] (Slow modes)	-0.203	0.0504	-4.02	0.00
20	Residuals of the regression (β_δ)	0.413	0.473	0.87	0.38
21	Correction coefficient for endogeneity (γ)	-0.429	0.272	-1.58	0.11
22	Term shifting the ASC for PT (α_0)	0.101	1.03	0.10	0.92
23	Mean of θ_ξ (η_0)	0.202	0.137	1.47	0.14

Summary statistics

Number of observations = 1899

Number of estimated parameters = 23

$$\mathcal{L}(\beta_0) = -1086.151$$
$$\mathcal{L}(\hat\beta) = -917.799$$
$$-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat\beta)] = 336.703$$
$$\rho^2 = 0.155$$
$$\bar\rho^2 = 0.134$$

Table 5: Estimation results of the EMIS method.

not significantly different from zero even at a 20% level. However, as long as η_0 and γ are significantly different from zero, it means that endogeneity is being corrected for. As for the MIS, ideally these parameters would be more significant. In the case study they are only significantly different from zero at a 15% level.

γ has the expected negative sign, and η_0 the positive expected sign. However the expectations would be for it to be larger as the variance of the distribution is 1 meaning that θ_ξ will have a high probability of being negative.

4.3.5 Comparison of the different methodologies

In this section the elasticities of the travel time parameter for public transportation as well as the value of time (VOT) estimates will be compared across the five methods presented above.

The software Biogeme (Bierlaire (2003)) was also used for the simulation of these estimates. It gives as an output the value of the point estimate for each respondent as well as the 5% the 95% percentiles of the point estimates.

Value of Time Table 6 contains the mean of the point estimates, the 5% and the 95% percentiles of value of time across all observations. We can see that for the base model and for the MIS correction when considering the interaction with travel time, we obtain the highest mean of the point estimate of VOT (25.5 and 25.4 respectively). For the EMIS, MIS and ICLV it is lower. However, the 5% and 95% percentiles are very far away from the point estimate. This means that the values obtained are not significantly different from each other. This can be due to the fact that the parameter associated to travel cost by public transportation is not significant at a 10% level in any of the models, and since it is in the denominator of the VOT, it will make the standard errors for VOT increase. For this reason we decide to study the estimate for time elasticity in public transportation.

	Base model	ICLV	MIS	MIS interaction	EMIS
5% percentile	-72.3	-82.2	-65.7	-82.7	-65.7
Point estimate	25.5	23.4	24.1	25.4	24.2
95% percentile	143	100	108	111	108

Table 6: Means of the point estimate, the 5% and the 95% percentiles of the VOT [CHF/h] for each methodology.

Figure 2 shows a boxplot for each methodology containing the point estimates of VOT for each respondent. It is always positive and sometimes larger than 60 CHF/h. This is in line with the findings by Axhausen *et al.* (2008) where they find that the mean VOT in Switzerland in 2008 for public transportation varies between 18 and 50 CHF/h depending on the trip purpose.

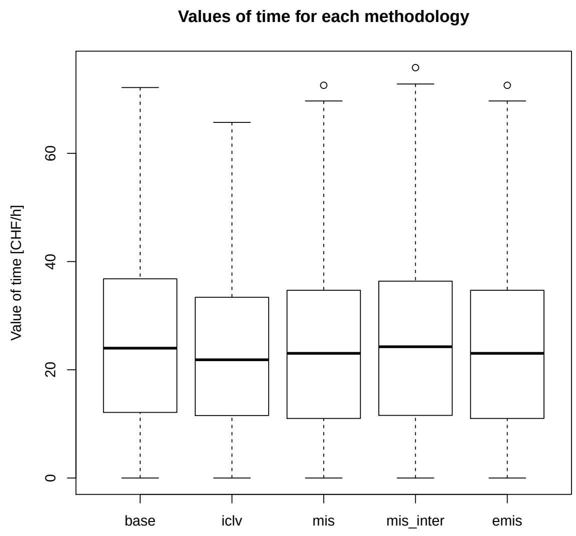

Figure 2: Boxplots of the VOT point estimates [CHF/h] for each methodology.

Travel time elasticity in public transportation Table 7 shows the average of the point estimates, the 5% and the 95% percentiles of time elasticity for public transportation across all observations. The results shown are multiplied by 100 to make it more legible.

	Base model	ICLV	MIS	MIS interaction	EMIS
5% percentile	-120	-115	-122	-113	-123
Point estimate	-84.9	-81.7	-85.3	-75.9	-85.6
95% percentile	-52.9	-47.7	-50.1	-41.5	-50.1

Table 7: Means of the point estimate, the 5% and the 95% percentiles of the time elasticity by public transportation for each methodology ($\cdot 10^2$)

The elasticity of travel time represents the variation in the probability of choosing the public transportation alternative following an increase in the travel time of this mode. Its value is negative as expected. The magnitudes, however, are not as expected. The hypothesis is that in a logit model, the travel time parameter contains the disutility towards travel time but also towards discomfort. By correcting for the endogeneity caused by the omitted variable of comfort, the intuition is that the elasticity for travel time in public transportation will be less negative. This is not the case for the MIS method or for the EMIS method. It is the case instead for both the ICLV and the MIS considering the interaction. It is considered future work to be able to explain these differences.

Figure 3 contains the boxplots of the disaggregate point estimates of time elasticity in public transportation. The upper plot contains the outliers. In the lower plot these have been excluded to make it more readable. Both the mean and the spread of all the boxplots look very similar, except for the MIS method with the interaction with travel time. For the latter one the variability of the point estimates is smaller. However, this result is difficult to interpret as we have not found reference values in the literature to compare it with.

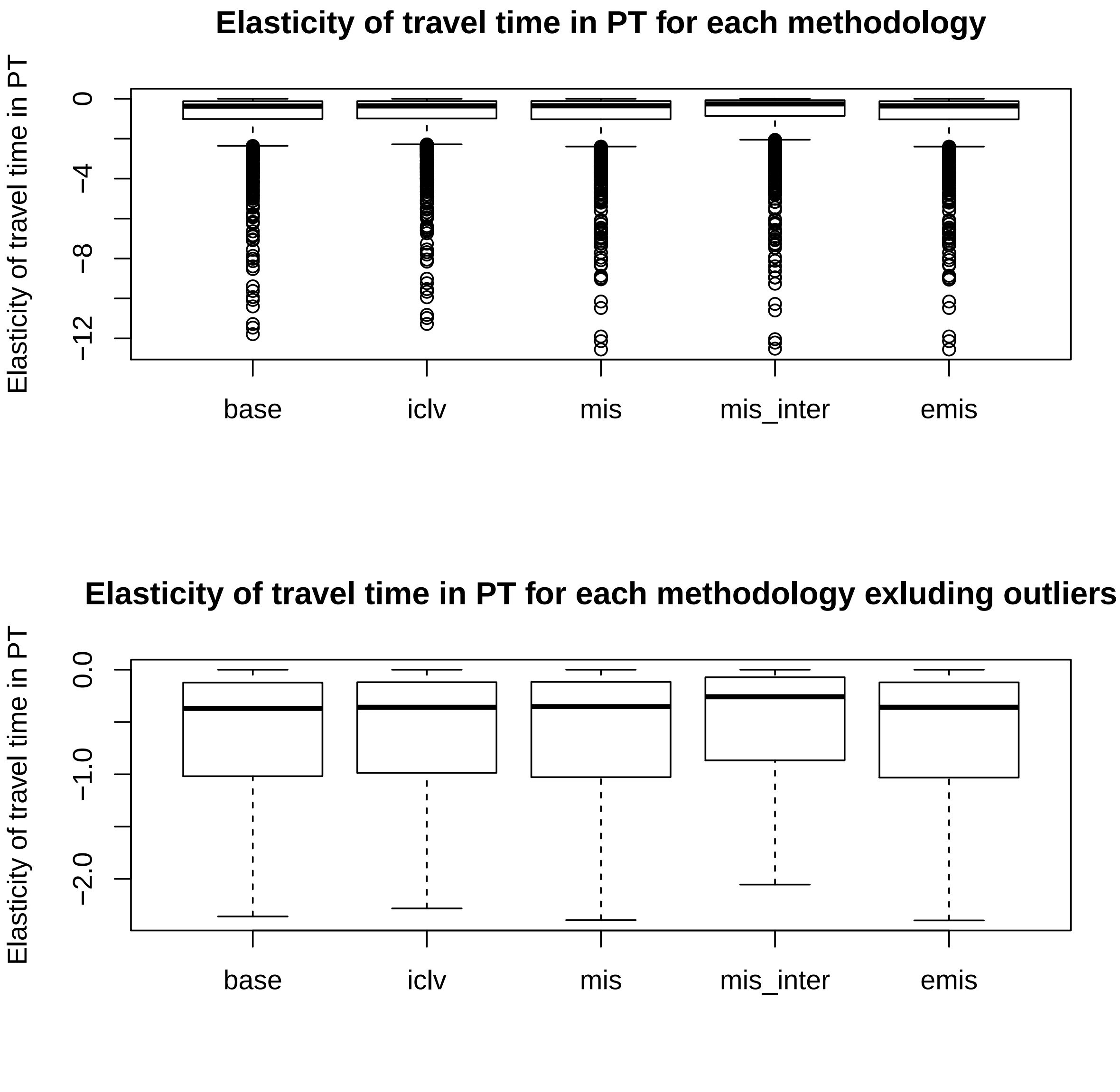

Figure 3: Boxplots of the elasticity of travel time by public transportation point estimates for each methodology. Upper figure: with outliers. Lower figure: without outliers.

5 Conclusions and future work

We introduce the EMIS methodology to correct for endogeneity in discrete choice models, which integrates two existing ones: the ICLV and the MIS. The Optima dataset, about mode choice in Switzerland, was used to illustrate the first application of the EMIS method, and one of the first ones of the MIS method. The results were compared to the ones obtained using a logit model and an ICLV model. The goal was to show how the above methods allow to obtain more realistic values of time than when the perception of comfort in public transportation is omitted.

However, we found that the different methodologies, including a logit model, do not perform very differently among them. This might be due to a number of reasons. It is possible that the indicators used are not good indicators. A good indicator would be one highly correlated with the unobserved variable. Moreover, there are only indicators related to comfort for 831 out of the 1899 observations. It might also be due to the fact that we only have access to the marginal cost of the trips, since a large proportion of the respondents have public transportation travel cards with which the marginal cost of each trip is zero. This explains why the parameter related to cost in public transportation is not as significant as we would expect it to be, and this implies directly that the standard errors of the value of time estimates will be large. Finally, it could also be that the unobserved comfort does not play a role in the respondents' choice. Therefore, in this particular case study there would be no endogeneity to correct for.

The next steps are to analyze in more detail the results, to be able to tell which of the previous reasons are causing them. It is also important to study how strong the assumptions needed for the EMIS method to work are, in particular about the distributions of the different error terms. To do so, Montecarlo experimentation will be performed. It will also be studied in more detail whether the interaction between comfort and travel time should be kept or dismissed, and what are the consequences of disregarding it if it is part of the true model. This will also be done by means of Montecarlo experiments.

6 References

Amemiya, T. (1978) The estimation of a simultaneous equation generalized probit model, *Econometrica*, **46** (5) 1193–1205, September 1978.

Axhausen, K. W., S. Hess, A. König, G. Abay, J. J. Bates and M. Bierlaire (2008) Income and distance elasticities of values of travel time savings: New Swiss results, *Transport Policy*, **15** (3) 173–185, May 2008, ISSN 0967070X.

Baum, C. F. (2006) *An Introduction to Modern Econometrics Using Stata*, Stata Press, August 2006, ISBN 9781597180139.

Berry, S., J. Levinsohn and A. Pakes (1995) Automobile prices in market equilibrium, *Econometrica*, **63** (4) 841–890, July 1995.

Berry, S., O. B. Linton and A. Pakes (2004) Limit theorems for estimating the parameters of differentiated product demand systems, *The Review of Economic Studies*, **71** (3) 613–654, July 2004, ISSN 0034-6527, 1467-937X.

Bierlaire, M. (2003) Biogeme: a free package for the estimation of discrete choice models, paper presented at the *Proceedings of the Swiss Transport Research Conference*, Ascona, Switzerland.

Bierlaire, M., A. Curchod, A. Danalet, E. Doyen, P. Faure, A. Glerum, V. Kaufmann, K. Tabaka and M. Schuler (2011) Projet de recherche sur la mobilité combinée, Rapport définitif de l'enquête de préférences révélées, *Technical Report*.

Crawford, G. S. (2000) The impact of the 1992 cable act on household demand and welfare, *SSRN Scholarly Paper*, **ID 235290**, Social Science Research Network, Rochester, NY, July 2000.

de Lapparent, M. and M. Koning (2015) Travel discomfort-time tradeoffs in Paris subway: an empirical analysis using interval regression models, *Technical Report*.

Glerum, A., B. Atasoy and M. Bierlaire (2014) Using semi-open questions to integrate perceptions in choice models, *Journal of Choice Modelling*, **10**, 11–33, March 2014, ISSN 17555345.

Goolsbee, A. and A. Petrin (2004) The consumer gains from direct broadcast satellites and the competition with cable TV, *Econometrica*, **72** (2) 351–381, March 2004, ISSN 1468-0262.

Guevara, C. A. and D. Polanco (2013) Correcting for endogeneity without instruments in discrete choice models: the multiple indicator solution, paper presented at the *Proceedings of the Third International Choice Modelling Conference*, Sydney, Australia.

Guevara-Cue, C. A. (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, J. A. (1978) Specification tests in econometrics, *Econometrica: Journal of the Econometric Society*, 1251–1271.

Heckman, J. J. (1978) Dummy endogenous variables in a simultaneous equation system, *Working Paper*, **177**, National Bureau of Economic Research, May 1978.

Heckman, J. J. and R. Robb Jr. (1985) Alternative methods for evaluating the impact of interventions: An overview, *Journal of Econometrics*, **30** (1-2) 239–267, October 1985, ISSN 0304-4076.

Jiang, R., P. Manchanda and P. E. Rossi (2009) Bayesian analysis of random coefficient logit models using aggregate data, *Journal of Econometrics*, **149** (2) 136–148, April 2009, ISSN 0304-4076.

Kraus, M. (1991) Discomfort externalities and marginal cost transit fares, *Journal of Urban Economics*, **29** (2) 249–259.

Lancaster, T. (2004) *Introduction to Modern Bayesian Econometrics*, Wiley, June 2004, ISBN 9781405117203.

Louviere, J., K. Train, M. Ben-Akiva, C. Bhat, D. Brownstone, T. A. Cameron, R. T. Carson, J. R. Deshazo, D. Fiebig and W. Greene (2005) Recent progress on endogeneity in choice modeling, *Marketing Letters*, **16** (3-4) 255–265.

Nevo, A. (2001) Measuring market power in the ready-to-eat cereal industry, *Econometrica*, **69** (2) 307–342, March 2001, ISSN 1468-0262.

Newey, W. K. (1985) Semiparametric estimation of limited dependent variable models with endogenous explanatory variables, *Annales de l'inséé*, (59/60) 219–237, July 1985, ISSN 00190209.

Newey, W. K. (1987) Efficient estimation of limited dependent variable models with endogenous explanatory variables, *Journal of Econometrics*, **36** (3) 231–250, November 1987, ISSN 0304-4076.

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

Petrin, A. and K. Train (2009) A control function approach to endogeneity in consumer choice models, *Journal of Marketing Research*, **47** (1) 3–13, February 2009, ISSN 0022-2437.

Train, K. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press, ISBN 9780521017152.

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

Villas-Boas, J. M. and R. S. Winer (1999) Endogeneity in brand choice models, *Management Science*, **45** (10) 1324–1338, October 1999, ISSN 0025-1909.

Walker, J. and M. Ben-Akiva (2002) Generalized random utility model, *Mathematical Social Sciences*, **43** (3) 303–343.

Walker, J. L. (2001) Extended discrete choice models: integrated framework, flexible error structures, and latent variables, Ph.D. Thesis, Massachusetts Institute of Technology.

Walker, J. L., E. Ehlers, I. Banerjee and E. R. Dugundji (2011) Correcting for endogeneity in behavioral choice models with social influence variables, *Transportation Research Part A: Policy and Practice*, **45** (4) 362–374, May 2011, ISSN 09658564.

Wardman, M. and G. Whelan (2011) Twenty Years of Rail Crowding Valuation Studies: Evidence and Lessons from British Experience, *Transport Reviews*, **31** (3) 379–398, May 2011, ISSN 0144-1647, 1464-5327.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press, January 2002, ISBN 9780262232197.

Yang, S., Y. Chen and G. M. Allenby (2003) Bayesian analysis of simultaneous demand and supply, *Quantitative Marketing and Economics*, **1** (3) 251–275, September 2003, ISSN 1570-7156, 1573-711X.

A Estimation results

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC_{PT}	-1.08	0.412	-2.62	0.01
2	$AS\tilde{C}_{PT}$	-0.978	0.428	-2.29	0.02
3	Age 0-45 (PT)	-0.0167	0.00704	-2.36	0.02
4	Age 45-65 (PT)	0.0215	0.0109	1.96	0.05
5	Marginal cost [CHF] (PT)	-0.0138	0.0107	-1.29	0.20
6	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if full time job	-2.40	0.447	-5.37	0.00
7	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if other occupation	-0.621	0.430	-1.44	0.15
8	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if part time job	-1.98	0.612	-3.24	0.00
9	Season ticket dummy (PT)	2.86	0.324	8.84	0.00
10	Half fare travel card dummy (PT)	0.575	0.167	3.44	0.00
11	Line related travel card dummy (PT)	1.91	0.258	7.40	0.00
12	Area related travel card (PT)	2.76	0.246	11.23	0.00
13	Other travel cards dummy (PT)	1.53	0.265	5.77	0.00
14	ASC_{CAR}	-0.625	0.353	-1.77	0.08
15	Number of cars in household (Car)	0.748	0.107	6.98	0.00
16	Gasoline cost [CHF], if trip purpose HWH (Car)	-0.169	0.0379	-4.47	0.00
17	Gasoline cost [CHF], if trip purpose other (Car)	-0.0871	0.0311	-2.80	0.01
18	Gasoline cost [CHF], if male (Car)	-0.0397	0.0210	-1.89	0.06
19	French speaking (Car)	0.749	0.173	4.32	0.00
20	Distance [km] (Slow modes)	-0.203	0.0504	-4.02	0.00
21	Residuals of the regression (β_δ)	-0.00118	0.0203	-0.06	0.95
22	Correction coefficient for endogeneity (γ)	-0.386	0.282	-1.37	0.17
23	Image Comfort PT (θ_ξ)	0.0374	0.121	0.31	0.76

Summary statistics

Number of observations = 1899

Number of estimated parameters = 23

$$\mathcal{L}(\beta_0) = -1086.151$$
$$\mathcal{L}(\hat{\beta}) = -918.769$$
$$-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})] = 334.763$$
$$\rho^2 = 0.154$$
$$\bar{\rho}^2 = 0.133$$

Table 9: Estimation results for the MIS method with interaction.

Parameter number	Description	Coeff. estimate	Robust Asympt. std. error	t-stat	p-value
1	ASC_{PT}	-1.85	0.482	-3.83	0.00
2	Age 0-45 (PT)	-0.0123	0.00695	-1.77	0.08
3	Age 45-65 (PT)	0.0105	0.0117	0.90	0.37
4	Marginal cost [CHF] (PT)	-0.0143	0.0103	-1.39	0.16
5	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if full time job	-2.50	0.432	-5.80	0.00
6	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if other occupation	-0.917	0.362	-2.53	0.01
7	Travel time [min] $\times$ log(1+ distance[km]) / 1000, if part time job	-1.93	0.558	-3.47	0.00
8	Season ticket dummy (PT)	2.94	0.326	9.00	0.00
9	Half fare travel card dummy (PT)	0.582	0.168	3.46	0.00
10	Line related travel card dummy (PT)	1.94	0.265	7.35	0.00
11	Area related travel card (PT)	2.83	0.253	11.18	0.00
12	Other travel cards dummy (PT)	1.58	0.264	5.98	0.00
13	ASC_{CAR}	-0.592	0.352	-1.68	0.09
14	Number of cars in household (Car)	0.718	0.106	6.77	0.00
15	Gasoline cost [CHF], if trip purpose HWH (Car)	-0.169	0.0351	-4.81	0.00
16	Gasoline cost [CHF], if trip purpose other (Car)	-0.0814	0.0272	-2.99	0.00
17	Gasoline cost [CHF], if male (Car)	-0.0367	0.0200	-1.84	0.07
18	French speaking (Car)	0.745	0.175	4.25	0.00
19	Distance [km] (Slow modes)	-0.203	0.0503	-4.03	0.00
20	Image Comfort PT (θ_{ξ})	-0.117	0.0405	-2.88	0.00
21	λ_{active}	1.01	0.245	4.13	0.00
22	$\lambda_{age_{50}}$	1.13	0.258	4.39	0.00
23	α_2	-0.00118	0.152	-0.01	0.99
24	α_3	-0.308	0.269	-1.14	0.25
25	α_4	-1.32	0.268	-4.93	0.00
26	α_5	-1.22	0.260	-4.68	0.00
27	α_6	-1.59	0.368	-4.31	0.00
28	α_7	-0.487	0.177	-2.76	0.01
29	α_8	-0.705	0.265	-2.66	0.01
30	α_9	-1.31	0.394	-3.31	0.00
31	λ_1	-0.146	0.0203	-7.17	0.00
32	λ_2	-0.0902	0.0232	-3.89	0.00
33	λ_3	-0.104	0.0352	-2.96	0.00
34	λ_4	-0.268	0.0198	-13.56	0.00
35	λ_5	-0.184	0.0313	-5.88	0.00
36	λ_6	-0.257	0.0400	-6.41	0.00
37	λ_7	-0.207	0.0276	-7.49	0.00
38	λ_8	-0.199	0.0222	-8.97	0.00
39	λ_9	-0.248	0.0368	-6.74	0.00
40	λ_{mean}	-7.08	0.907	-7.80	0.00
41	λ_{French}	0.755	0.280	2.70	0.01
42	σ_1	-0.463	0.0529	-8.74	0.00
43	σ_2	-0.159	0.0262	-6.07	0.00
44	σ_3	-0.158	0.0409	-3.85	0.00
45	σ_4	-0.580	0.124	-4.68	0.00
46	σ_5	-0.395	0.0882	-4.47	0.00
47	σ_6	-0.787	0.294	-2.68	0.01
48	σ_7	-0.520	0.112	-4.64	0.00
49	σ_8	-0.242	0.0603	-4.01	0.00
50	σ_9	-0.450	0.182	-2.48	0.01
51	λ_{cars}	0.458	0.246	1.86	0.06

Summary statistics

Number of observations = 1899

Number of estimated parameters = 51

$$\mathcal{L}(\beta_0) = -5792.712$$
$$\mathcal{L}(\hat{\beta}) = -4188.043$$
$$-2[\mathcal{L}(\beta_0) - \mathcal{L}(\hat{\beta})] = 3209.337$$
$$\rho^2 = 0.277$$
$$\bar{\rho}^2 = 0.268$$

Table 8: Estimation results for the ICLV method.