Choice of mode for long distance travel

Current SP-based models from three European countries

Author(s):
Lapparent, Matthieu de; Frei, Andreas; Axhausen, Kay W.

Publication date:
2009

Permanent link:
https://doi.org/https://doi.org/10.3929/ethz-a-005888713

Originally published in:
Arbeitsberichte Verkehrs- und Raumplanung 579

Choice of mode for long distance travel: current SP-based models from three European countries

Matthieu de Lapparent* Andreas Frei†
Kay W. Axhausen‡

Abstract

We present the results of an econometric analysis of choice behaviours based on the RUM paradigm. We estimate and compare several parametric formulations of a panel mixed Logit model using SP data on long distance mode choice in three European countries (the Czech Republic, Switzerland, Portugal). The specification of the systematic part of the indirect utility function is flexible enough to account simultaneously for nonlinearities and interaction effects. Its random part is designed to allow existence of continuous unobserved taste heterogeneity and individual random effects.

We find clear evidence of the presence of nonlinearities in modelling long distance mode choice. It is also found that accounting for both observed and unobserved taste heterogeneity improves also our understanding of choice behaviours. Our results show that perception and attention paid to the travel attributes differ from one country to another but also accross the population of travellers of each of the considered countries.

*Corresponding author. Pôle Paris-Est. Institut National de Recherche sur les Transports et leur Sécurité. Département Economie et Sociologie des Transports. Bâtiment Descartes 2. 2 rue de la Butte Verte. F-93166 Noisy-le-Grand Cedex, France. E-mail: lapparent@inrets.fr. Tel: +33 1 459 255 72, Fax: +33 1 459 255 01

†ETH Zürich. Institute for Transport Planning and Systems. HIL F 51.3. Wolfgang-Pauli-Stra. 15. 8093 Zürich, Switzerland. E-mail: frei@ivt.baug.ethz.ch. Tel: +41 44 633 4102

‡ETH Zürich. Institute for Transport Planning and Systems. HIL F 31.3. Wolfgang-Pauli-Stra. 15. 8093 Zürich, Switzerland. E-mail: axhausen@ivt.baug.ethz.ch. Tel: +41 44 633 3943, Fax: +41 44 633 1057

1 Introduction

KITE stands for "a Knowledge base for Intermodal Transport in Europe". It was a FP6 research project that aimed at providing relevant information about passenger intermodal behaviour and that allowed to develop and to evaluate transport policies and measures to favour it. It integrates and it disseminates current existing and future information and data. We refer the reader to http://www.kite-project.eu/ for more detailed information. Part of the project developed a strategic approach for the collection of long distance and intermodal survey data. In the present article, we use data from an associated SP survey that was conducted at the end of year 2008 and the beginning of year 2009 in three countries: Switzerland, Czech Republic, and Portugal. The sample size is 900 individuals, of which were 40% in Czech Republic, 40% in Switzerland, and 20% in Portugal. Each decision maker is faced with a sequence of 4 choice experiments, giving a total number of 3600 observations. It focuses on the choice of a main mode of transport for long distance travel. Four modes of transport are considered: car, air, train, bus. Each is described with at most four attributes: in-vehicle time, access time, cost, and number of transfers. Even though it does not consider explicitly combination of modes of transport, it sheds light on potential demand for intermodal travel as air, train, and bus are modes for which travellers access and egress by almost systematically using other modes of transport.

We present the results of an econometric analysis of choice behaviours that we postulate to be consistent with random utility maximization. We estimate and compare several parametric formulations of a panel mixed Logit model. From a general point of view, the specification of the systematic part of the indirect utility function is flexible enough to account simultaneously for nonlinearities and interaction effects. Its random part is designed to allow existence of continuous unobserved taste heterogeneity and individual random effects. Fit statistics are presented for all the estimated models but only the detailed results of the most informative one according to some statistical criteria are reported and discussed.

Nonlinearities are modelled by means of Box-Cox transformations and multiplicative terms that incorporate the travel attributes. Interaction effects are captured by defining the parameters of the model as functions of the characteristics of the decision maker. Unobserved taste heterogeneity models random variation of the parameters that weigh the independent variables (i.e. each decision maker has specific tastes). We may also model unobserved taste heterogneity by means of random individual effects, also known as agent effects. In

our application, distributional assumptions will be set to keep consistent with economic demand theory (e.g. negativity of price and time coefficients).

As a result, the values of travel time savings may be defined as functions of travel attributes and socioeconomic characteristics. Not only are they key figures that will be used in the KITE project for development of an aggregate cost-benefit analysis tool to assess intermodal transport projects but they also give important information on disaggregate behavioural responses to modification of transport supply, particularly as it regards the incentives to relieve the use of a car for long distance travel. To that extent, a series of values on the basis of our work is formally derived, computed, and then discussed.

The outline of the article is as follows. After an introductory section, data are presented in a second section. The econometric specification is then developed in a third section. The results are discussed in a fourth section. In a last section are drawn conclusions and are discussed further extensions of the proposed approach.

2 Data

One of the work package of the European KITE project pertained to test a suitable survey methodology that intends to close remaining information gaps about long-distance travel behaviour by means of pilot surveys. These pilot surveys were carried out in three countries: Portugal, Switzerland, and the Czech Republic. They were carried out by means of computer assisted telephone interviews (CATI) in Portugal and in Switzerland and by means of face-to-face interviews in the Czech Republic.

One of the purposes of these pilot surveys was to test whether it would be possible to implement a common methodology in different countries in Europe and then to assess the quality of information that can be obtained through data collection. In particular, computation of figures to characterise demand for long distance travel and comparison with existing data sources were made to get a better idea of the promise of the used methodology.

Parallel to this approach, 2 stated preference (SP) surveys were designed to gather information about market potentials and user requirements. They focused on long distance main mode choice and long distance route choice given the main mode of transport. Preliminary investigation of the SP surveys were made to sketch profiles of users' needs. The SP surveys were built up on sampling individuals in the main survey (a revealed preference survey, i.e. RP survey) and using their answers to customise choice experiments to which they had to

answer. Actually, based on the answers in the first part of the survey, the SP surveys were sent to self-identified respondents. Those respondents which had undertaken a long distance journey during the last 8 weeks, which was not a regular journey[1], were asked if they were willing to participate in a written survey based on this telephone interview. Generation of hypothetical choice situations for these written self-completion stated preference surveys were based on one of the reported long distance journeys from the telephone interview.

The main target in these SP surveys is to discover and to analyse the preferences of the travellers who undertake long distance journeys. These preferences show the requirements of the users and their requirements towards a more sustainable use of transport means, e.g. under which circumstances they would change the transport mean and use public transport instead of car. The use of a transport mode is of course dependent on the available infrastructure in the different countries and regions etc.. It is not analyzed in this survey, but the results give key parameters which indicate under what kind of infrastructure change the population would accept to change their transport mode or their route choice. In the present article, we focus only on the choice of a main mode of transport.

To generate the experimental design for the SP questionnaires the software Ngene (e.g. Rose and Scarpa (2007)) was used. This software makes it possible to generate efficient experimental designs and therefore have small numbers of experiments by interviewee without losing goodness of fit in the models estimated with the data. Based on one of the reported journeys, the journeys' characteristics for the different modes were drawn and calculated using different data souces. Travel times and number of interchanges were drawn from the IVT Air Network, the IVT Road Network and the IVT TransEuropean Train Model. Travel cost were generated by implementing automatic internet requests that were manually corrected when necessary. With these observed/imputed values and the given characteristics from the experimental design, the different choice situations for the SP questionnaires were finally produced.

Table 1 reports simple descriptive statistics of the variables of interest. It is shown that we do not have available any socioeconomic

[1] A regular journey was defined as: at least once per week or journeys with the same destination during the last 8 weeks

and demographic variables for the Portuguese SP survey [2].

Table 1 about here

It is observed clear differences in the distributions of sociodemographic and economic characteristics accross the populations of decision makers in the Czech Republic and in Switzerland:

- Swiss people seem to pursue longer education;
- the [41; 60] years old interval is dominating in the Switzerland sample whereas the Czech sample has a larger proportion of younger travellers;
- even though the Switzerland and the Czech sample have the modal income class, it is observed that the mean income is larger in Switzerland and that its distribution has a larger right tail.

When regarding the distance of the reference trips (the ones draws from the RP survey to build the SP experiments), the results show that the sampled travellers from Portugal and from Switzerland make longer trips as compared to the sampled travellers from the Czech Republic.

Finally, concerning the levels of the attributes of the SP experiments, even though the access time variables take approximately the same value, we observe that the in-vehicle time variables takes larger average values for Swiss and for Portuguese travellers. We observe also that, excepted for the cost of the air mode, the cost variables of the other modes are really lower for the Czech travellers when compared to the travellers of the two other countries. The average number of interchanges is approximately the same in each country.

3 Model

3.1 Econometric specification

This section describes the specification of the discrete choice models that are used to explain the observed choices of the decision makers among modes of transport. It discusses also briefly estimation of such models by maximization of the associated simulated log-likelihood function.

[2] Actually, it was not possible to match socioeconomic and demographic variables from the RP survey with data from the SP survey because of the non correspondence between the ID keys that define the decision makers even though the latter participated to both surveys.

Consider a population of n decision makers, each labelled by $i = 1, \cdots, n$. Each of them is characterized by socioeconomic and demographic variables (e.g. age, gender, income class) that do not vary over choice situations and that we model by a column vector $\mathbf{z}_i$ (it includes also an intercept term). He is faced $t = 1, \cdots, T$ consecutive times with a choice between $M = 4$ modes of transport: car, train, air, and coach. The latter are described in terms of four attributes: in-vehicle time (labelled as "ivtime"), access time (labelled as "acctime"), cost (labelled as "cost"), and number of interchanges (labelled as "change"). These attributes may vary from one mode of transport to another, from one choice situation to another, and from one decision maker to another. The levels of these attributes for mode m in choice experiment t faced by responent i are therefore appropriately indexed by these subscripts.

The analysis presented in this paper makes use of standard discrete choice methodology that is based on Random Utility Maximisation (RUM). Within the framework of RUM, it is postulated that:

- preferences of any decision maker over a set of discrete alternatives are represented by a utility funtion U that is defined up to a monotone increasing transformation;
- any decision maker chooses the alternative that maximizes his/her level of utility;
- as modeller, one does not necessarily observe everything that may explain the level of utility for an alternative. One set up therefore a probabilistic framework within which it is postulated that the probability to choose an alternative in the given choice set is defined as the probability that it gives the maximum level of utility.

We refer the reader to Marschak (1959), McFadden and Richter (1970), McFadden (1974), Ben-Akiva and Lerman (1985), Train (2003), McFadden (2005) for detailed discussions and proofs of the necessary and sufficient theoretical conditions under which such a framework may hold. The utility that decision maker i would obtain from mode m in choice experiment t is mathematically defined here as:

$$\begin{aligned} U_{i,t,m} \quad = \quad & \mathbf{z}_i{}'\beta_m & + \\ & \alpha_1(\mathbf{z}_i)\, f_1(\text{cost}_{i,t,m}) & + \\ & \alpha_2(\mathbf{z}_i)\, f_2(\text{ivtime}_{i,t,m}) & + \\ & \alpha_3(\mathbf{z}_i)\, f_3(\text{acctime}_{i,t,m}) & + \\ & \alpha_4(\mathbf{z}_i)\, f_4\left(\text{change}_{i,t,m}\right) & + \\ & \sigma_m \eta_{i,m} & + \\ & \epsilon_{i,t,m}. \end{aligned} \tag{1}$$

In the present approach, we assume that the coefficients that weigh the attribute-specific variables are defined as functions of the characteristics of the decision maker. They vary accross the population of decision makers but not over choice expriments. Because it is likely that we do not observe all of these characteristics, we assume that there may exist some additional unobserved heterogeneity (which is known by the decision maker).

$$\alpha_j(\mathbf{z}_i) = -\exp\left(\mathbf{z}_i'\gamma_j + \psi_j\xi_{i,j}\right), \xi_{i,j} \overset{iid}{\to} \mathcal{N}(0,1). \tag{2}$$

The formula in equation 2 uses also, up to a preceding negative sign, an exponential transformation of the linear combination of these observed and unobserved characteristics of the decision maker. All in all, our assumptions account for deterministic and random heterogeneity in modelling the impacts of alternative-specific variables on the level of utility of the decision maker and they force consistency with theoretical constraints (negative impacts of these attributes). It allows virtually to model any interaction effects between socio-demographic variables and alternative-specific attributes, but not only. These interaction effects may be continuous or discrete, depending on the nature of the variables that one may enter in the specification of the coefficients associated to the alternative-specific variables.

Mackie et al. (2003) discussed in detail a very interesting a continuous interaction formulation. Let s be the observed value for a given socio-demographic or economic continuous variable (e.g. income, age and also distance travelled). Let $\bar{s}$ be also a reference value for this attribute. The choice of this reference is arbitrary but, to make things clear, assume that it is the mean value. Let x be the attribute with which one would like to model an interaction effect. The interaction between these 2 variables is defined as

$$g(s, x|\bar{s}, \delta, \tau) = \tau\left(\frac{s}{\bar{s}}\right)^{\delta} x. \tag{3}$$

It is obvious that the sensitivity to an alternative's attribute x varies with s. The choice of $\bar{s}$ implies that δ gives the sensitivity to x at the average value of s in the sample population. The latter formulation has already been used by Hess et al. (2008), Axhausen et al. (2008), and has proven to improve significantly the results of their discrete choice models. It appears as a great alternative to more simple interaction effects using segmentation of variables into different classes. While many modelling analyses allow for some interactions between socio-demographic characteristics and travel attributes, these generally come in the form of a simple segmented crossing of these variables. In the present approach, instead of making these approaches to

compete together, we combine them in the specification of the $\alpha_j(.)$ functions, thereby producing an elegant extension of the initial approach but at the cost of a formulation that applies only to strictly positive variables. Indeed, let $\mathbf{z}_i = \left(\tilde{\mathbf{z}}_i', \ln\left(\frac{s_i}{\bar{s}}\right)\right)'$ and $\gamma_j = \left(\tilde{\gamma}_j', \delta_j\right)'$. Equation 2 may therefore be rewritten as

$$\alpha_j(\mathbf{z}_i) = -\exp\left(\tilde{\mathbf{z}}_i'\tilde{\gamma}_j + \delta_j \ln\left(\frac{s_i}{\bar{s}}\right) + \psi_j \xi_{i,j}\right), \tag{4}$$

or

$$\alpha_j(\mathbf{z}_i) = -\exp\left(\tilde{\mathbf{z}}_i'\tilde{\gamma}_j + \psi_j \xi_{i,j}\right)\left(\frac{s_i}{\bar{s}}\right)^{\delta_j}, \tag{5}$$

which is similar to $\tau\left(\frac{s}{\bar{s}}\right)^{\delta}$ in equation 3.

A simple linear combination of the explanatory variables (whatever they are interacting or not each with the others), eventually in the presence of additional error terms, may not be the best way to model the utility function of the decision maker. The mathematical forms we use for the $f_j(.)$ functions are based on Box-Cox transformations. While preserving the property of additive separability of the the utility function, we use Box-Cox transformations on alternative-specific variables to capture possible curvatures of the utility function, thereby varying marginal rates of substitution between these variables with respect to their levels. The Box-Cox transformation (BCT), by statisticians Box and Cox, is one particular way of parameterizing a power transform that has advantageous properties. It is a restrictive case of the more general Box-Tukey transformation and it maps data from $\mathbb{R}_\star^+$ to $\mathbb{R}$ by using the following transformation:

$$\forall \mathrm{a} \in \mathbb{R}_\star^+, f_j(\mathrm{a}) = \begin{cases} \frac{\mathrm{a}^{\lambda_j}-1}{\lambda_j} & \text{if} \quad \lambda_j \neq 0 \\ \ln(\mathrm{a}) & \text{if} \quad \lambda_j = 0 \end{cases}. \tag{6}$$

In our approach, it is furthermore assumed that if $\mathrm{a} \in \mathbb{R}$, then $f_j(\mathrm{a}) = \mathrm{a}$. Since their seminal paper, the BCT has generated a great deal of interests, both from theoretical and empirical perspectives. In addition to theoretical investigation of its properties in various framework of analysis, it has been applied in several disciplines and on various topics of research. In particular, in transport analysis, Gaudry and Wills (1977), Gaudry and Dagenais (1979), Gaudry and Laferriere (1989) have shown that it improves significantly not only the goodness-of-fit of either standard regression models or MNL discrete choice models but also their interpretation. The Box-Cox transformation is of course not the only existing non linear transformation but it adds greater flexibility of the functional form at a rather low computational cost.

The random utility function that is defined in equation 1 may also exhibit individual random effects, also known as agent effects

(Walker et al. (2007)). The idea is that some unobserved characteristics do not enter the utility function only through the coefficients that weigh the alternative-specific variables but they may also enter directly the utility function. These unobserved characteristics are not necessarily the same accross alternatives, thereby different error terms. There is no prior theoretical aspect to bound the distribution of such effects. In the present approach, these agent effects are independently and identically distributed standard normal:

$$\eta_{i,m} \overset{iid}{\rightarrow} \mathcal{N}(0,1). \tag{7}$$

Finally, we build up our approach on the use of Multinomial Logit (MNL). Defining $\boldsymbol{\epsilon}_{i,t,\cdot} = (\epsilon_{i,t,1}, \cdots, \epsilon_{i,t,M})', \boldsymbol{\epsilon}_{i,\cdot,\cdot} = \left(\boldsymbol{\epsilon}'_{i,1,\cdot}, \cdots, \boldsymbol{\epsilon}'_{i,T,\cdot}\right)'$, it is assumed that

$$\boldsymbol{\epsilon}_{i,\cdot,\cdot} \overset{iid}{\rightarrow} F_{\boldsymbol{\epsilon}_{i,\cdot,\cdot};\kappa} = \prod_{t=1}^{T} \prod_{m=1}^{M} \exp\left(-\exp\left(-\kappa a_{i,t,m}\right)\right). \tag{8}$$

These error terms are independently and identically distributed type 1 extreme value with scale parameter κ.

As already stated earlier, RUM implies that the respondent chooses the mode of transport that provides the greater level of utility in each choice situation. Let $d_{i,t} \in \{1, \cdots, M\}$ denote the respondent's chosen alternative in experiment t, and let $\mathbf{d}_{i,\cdot} = (d_{i,1}, \cdots, d_{i,T})$ denote the respondent's sequence of choices. Since the error terms in $\boldsymbol{\epsilon}_{i,\cdot,\cdot}$ are identically and independently distributed type 1 extreme value, the choice probability conditional on $\boldsymbol{\xi}_i, \boldsymbol{\eta}_i$ and other observed variables that the decision maker i chooses the m-th mode in situation t is a MNL choice probability (McFadden, 1974). Furthermore, since the error terms in $\boldsymbol{\epsilon}_{i,\cdot,\cdot}$ are independent over choice experiments, the joint probability of the decision maker's sequence of choices, still conditional on $\boldsymbol{\xi}_i, \boldsymbol{\eta}_i$ and other observed variables, is the product of these MNL marginal choice probabilities:

$$\begin{aligned} &\Pr\left(\mathbf{d}_{i,\cdot} | \mathbf{z}_i, \mathbf{x}_{i,\cdot,\cdot}, \boldsymbol{\eta}_{i,\cdot}, \boldsymbol{\xi}_i; \boldsymbol{\theta}, \kappa\right) = \\ &\prod_{t=1}^{T} \prod_{m=1}^{M} \left[\frac{\exp\left(\kappa V_{i,t,m}\left(\mathbf{z}_i, \mathbf{x}_{i,t,m}, \eta_{i,m}, \boldsymbol{\xi}_i; \boldsymbol{\theta}\right)\right)}{\sum_{k=1}^{M} \exp\left(\kappa V_{i,t,k}\left(\mathbf{z}_i, \mathbf{x}_{i,t,k}, \eta_{i,k}, \boldsymbol{\xi}_i; \boldsymbol{\theta}\right)\right)} \right]^{y_{i,t,m}} \end{aligned} \tag{9}$$

where $V_{i,t,m}\left(\mathbf{z}_i, \mathbf{x}_{i,t,m}, \eta_{i,m}, \boldsymbol{\xi}_i; \boldsymbol{\theta}\right) = U_{i,t,m} - \epsilon_{i,t,m}$, and where $y_{i,t,m} = 1$ if m is chosen or 0 otherwise.

The researcher does not observe $\boldsymbol{\xi}_i$ and $\boldsymbol{\eta}_i$. As statistical inference is based on only observed data, the target quantity is therefore the expectation of the choice probabilities presented in equation 9 with respect to the joint distribution of the unobserved terms. These conditional choice probabilities are integrated over all possible values of

$\boldsymbol{\xi}_i$ and $\boldsymbol{\eta}_i$ using the latter's probability density function, here modelled by h and defined as the product of univariate standard normal distributions:

$$\begin{aligned}&\Pr\left(\mathbf{d}_{i,\bullet}|\mathbf{z}_i, \mathbf{x}_{i,\bullet,\bullet}; \boldsymbol{\theta}, \kappa\right) = \\ &\quad \int_{\mathbb{R}^{4+M}} \Pr\left(\mathbf{d}_{i,\bullet}|\mathbf{z}_i, \mathbf{x}_{i,\bullet,\bullet}, \boldsymbol{\eta}_{i,\bullet}, \boldsymbol{\xi}_i; \boldsymbol{\theta}, \kappa\right) h\left(\boldsymbol{\eta}_{i,\bullet}, \boldsymbol{\xi}_i\right) \mathrm{d}\boldsymbol{\eta}_{i,\bullet}\mathrm{d}\boldsymbol{\xi}_i\end{aligned} \tag{10}$$

Such choice probabilities are called mixed Logit probabilities because they are defined as mixtures of Logit formulas. The log-likelihood function that one would like to maximise may then be written as:

$$\ell\left(\boldsymbol{\theta}, \kappa|\mathbf{d}, \mathbf{z}, \mathbf{x}\right) = \sum\nolimits_{i=1}^{n} \ln\left(\Pr\left(\mathbf{d}_{i,\bullet}|\mathbf{z}_i, \mathbf{x}_{i,\bullet,\bullet}; \boldsymbol{\theta}, \kappa\right)\right). \tag{11}$$

The integral in the mixed logit probability that is presented in equation 10 does not have a closed form. It defines actually the expectation of the conditional probability in equation 9 with respect to the joint distribution of the unobserved components $\boldsymbol{\eta}_{i,\bullet}$ and $\boldsymbol{\xi}_i$. We make therefore use of Monte-Carlo integration techniques: the multivariate integral is approximated through simulation. In particular, for each decision maker i, R draws of $\boldsymbol{\eta}_{i,\bullet}, \boldsymbol{\xi}_i$ are taken from the probability density function $h\left(\boldsymbol{\eta}_{i,\bullet}, \boldsymbol{\xi}_i\right)$. For each draw, the joint probability in equation 9 is then calculated and the results are averaged over draws. The objective is then to maximize the simulated log-likelihood function over $\boldsymbol{\theta}$. Such a function may be written as

$$\begin{aligned}&\ell^R\left(\boldsymbol{\theta}, \kappa|\mathbf{d}, \mathbf{z}, \mathbf{x}, \hat{\boldsymbol{\eta}}, \hat{\boldsymbol{\xi}}\right) = \\ &\quad \sum\nolimits_{i=1}^{n} \ln\left(\tfrac{1}{R} \sum\nolimits_{r=1}^{R} \Pr\left(\mathbf{d}_{i,\bullet}|\mathbf{z}_i, \mathbf{x}_{i,\bullet,\bullet}, \hat{\boldsymbol{\eta}}_{i,\bullet,r}, \hat{\boldsymbol{\xi}}_{i,r}; \boldsymbol{\theta}, \kappa\right)\right)\end{aligned} \tag{12}$$

where

$$\begin{aligned}&\Pr\left(\mathbf{d}_{i,\bullet}|\mathbf{z}_i, \mathbf{x}_{i,\bullet,\bullet}, \hat{\boldsymbol{\eta}}_{i,\bullet,r}, \hat{\boldsymbol{\xi}}_{i,r}; \boldsymbol{\theta}, \kappa\right) = \\ &\quad \prod\nolimits_{t=1}^{T} \prod\nolimits_{m=1}^{M} \left[\frac{\exp\left(\kappa V_{i,t,m}\left(\mathbf{z}_i, \mathbf{x}_{i,t,m}, \hat{\eta}_{i,m,r}, \hat{\boldsymbol{\xi}}_{i,r}; \boldsymbol{\theta}\right)\right)}{\sum_{k=1}^{M} \exp\left(\kappa V_{i,t,k}\left(\mathbf{z}_i, \mathbf{x}_{i,t,k}, \hat{\eta}_{i,k,r}, \hat{\boldsymbol{\xi}}_{i,r}; \boldsymbol{\theta}\right)\right)}\right]^{y_{i,t,m}}.\end{aligned} \tag{13}$$

As already stated by Gouriéroux and Monfort (1996), Train (2003), if each draw is independent each from the others and from the probability in equation 9, then the simulated probability converges almost surely to the "true" probability, with variance inversely proportional to R. In maximum simulated likelihood (MSL) estimation, if R rises faster than the square root of the number of observation, then the effects of simulation disappear asymptotically, and MSL is equivalent to maximum likelihood with exact probabilities (see, e.g., Hajivassiliou and Ruud (1994); Hajivassiliou (1997), Lee (1995)). Under these regularity conditions (and some more), the MSL estimator is asymptotically unbiased, consistent, normal and efficient.

However, given a number of replications R, simulation bias and variance stays inherent to estimation. Furthermore, Pakes and Pollard (1989) suggest to use the same draws at each evaluation of the simulated log-likelihood function while estimating the parameters of interest (the population parameters in $\boldsymbol{\theta}$). In the application, which results are presented in the next section, we use Halton draws for the simulation (Train (2000)). This quasi-random number generation technique has been found to provide greater accuracy than standard pseudo-random number draws in simulation-based estimation of discrete choice models. Of course, as also stated in Bhat (2003), Hess et al. (2006), it is not the only way to generate appropriate draws. On may cite for instance the Modified Latin Hypercube Sampling technique, the Sobol sequences, or the Faure sequences. But it appears to be rather simple to set up for empirical purpose and, given a number of replications R to approximate choice probabilities, it diminishes the simulation bias and variance as compared to standard pseudo-random uniform numbers that may be used to generate the necessary draws.

Another point that pertains to estimation is identification of the parameters of interest. Because the utility function models preference orderings up to a monotone increasing transformation and because what determines choice are the differences between utility levels (see for instance the books of Ben-Akiva and Lerman (1985), Train (2003)), one must define additional exclusion constraints to ensure a one-to-one mapping between the log-likelihood function and the set of parameters of interest. One must select an alternative of reference from which are excluded all the variables that enter directly the utility function and that model the socioeconomic and demographic characteristics of the decision maker. However, due to the panel dimension of out model, Walker et al. (2007) showed that we do not need to exlude the agent effect from this alternative of reference.

One may also be attentive to computation of the standard deviation of the estimates of the model in the presence of Box-Cox transformations. Indeed, Spitzer (1984) has shown that the standard deviations of the parameters that weigh Box-Cox transformed variables are not invariant with respect to the unit of measurement used for these variables. A strategy is therefore to compute the standard deviation of these parameters conditionally on the Box-Cox transformation at the point of convergence of the log-likelihood function, i.e. assuming they are fixed parameters.

3.2 Modelling values of travel time savings

As already highlighted by Mackie et al. (2001), Lapparent et al. (2002), Hensher (2006), Hess et al. (2008) but also many other authors, reliable measures of the valuation of travel time savings (VTTS) are a crucial input into transport policy planning. Indeed, travel time savings are one of the dimension of the benefits arising from network expansion and capacity additions. The VTTS is a key willingness-to-pay indicator, used for example for cost-benefit analysis in the context of planning new transport systems. It enters the appraisal scheme as a value for use within project evaluation. The VTTS is one aspect of the total value of travel time, see for instance Becker (1965), DeSerpa (1971), Jara-Diaz (2000) for a theoretical microeconomic framework from which one may derive the VTTS.

In this subsection, on the basis of the model developed in the former subsection 3.1, we develop the theoretical analysis for the in-vehicle travel time variable. The analysis for the access travel time variable is obtained by just replacing the subscript ivtime by the subscript acctime.

The VTTS function is defined as the marginal rate of substitution between travel time and travel cost. It models the price the decision maker is willing to pay to save one unit of travel time sych as to maintain his/her level of utility. From a technical perspective, it is computed as the ratio between the marginal (dis)utility of travel time and the marginal (dis)utility of travel cost. According to our model specification, one obtains therefore

$$\begin{aligned}
&\psi_{\text{ivtime}} \equiv \\
&\quad \psi\left(\tilde{\mathbf{z}}_i, \text{distance}_i, \text{income}_i, t_{\text{ivtime},i,t,m}, c_{\text{cost},i,t,m}, \xi_{i,\text{ivtime}}, \xi_{i,\text{cost}}\right) = \\
&\qquad \exp\left(\tilde{\mathbf{z}}_i'\left(\tilde{\gamma}_{\text{ivtime}} - \tilde{\gamma}_{\text{cost}}\right) + \psi_{\text{ivtime}}\xi_{i,\text{ivtime}} - \psi_{\text{cost}}\xi_{i,\text{cost}}\right) \times \\
&\qquad \text{distance}_i^{\delta_{\text{distance,ivtime}} - \delta_{\text{distance,cost}}} \times \\
&\qquad \text{income}_i^{\delta_{\text{income,ivtime}} - \delta_{\text{income,cost}}} \times \\
&\qquad t_{\text{ivtime},i,t,m}^{\lambda_{\text{ivtime}}-1} \times \\
&\qquad c_{\text{cost},i,t,m}^{1-\lambda_{\text{cost}}}.
\end{aligned} \tag{14}$$

As stated earlier in the paper, statistical inference is based on only observed data. What is of interest is therefore the expected value of the ψ_{ivtime} function with respect to the joint distribution of the unobserved term $\xi_{i,\text{ivtime}}$ and $\xi_{i,\text{cost}}$:

$$\begin{aligned}
&VTTS_{\text{ivtime}}\left(\tilde{\mathbf{z}}_i, \text{distance}_i, \text{income}_i, t_{\text{ivtime},i,t,m}, c_{\text{cost},i,t,m}\right) = \\
&\quad \int_{\mathbb{R}^2} \psi_{\text{ivtime}} \varphi\left(\xi_{i,\text{ivtime}}, \xi_{i,\text{cost}}\right) \mathrm{d}\xi_{i,\text{ivtime}} \mathrm{d}\xi_{i,\text{cost}}
\end{aligned} \tag{15}$$

The effects of the characteristics of the decision maker on his/her willingness to pay for saving travel time depends on the difference

between the corresponding parameters that model observed taste heterogeneity that may affect in-vehicle travel time and travel cost. If the difference is positive, then the effect is positive. If the difference is negative, then the effect is negative.

The shape of the function that models the value of travel time savings with respect to the income of the decision maker depends on the difference $\Delta_{\text{income}} = \delta_{\text{income,ivtime}} - \delta_{\text{income,cost}}$. If $\Delta_{\text{income}} \in [0, 1]$, the VTTS function is increasing and concave with respect to income. If $\Delta_{\text{income}} > 1$, it is an increasing and convex function of income. Finally, if $\Delta_{\text{income}} < 0$, it is a decreasing and convex function of income.

The results that pertain to the shape of the VTTS function with respect to the reference distance are obtained by an equivalent approach. Let $\Delta_{\text{distance}} = \delta_{\text{distance,ivtime}} - \delta_{\text{distance,cost}}$. If $\Delta_{\text{distance}} \in [0, 1]$, the VTTS function is increasing and concave with respect to distance. If $\Delta_{\text{distance}} > 1$, it is an increasing and convex function of distance. Finally, if $\Delta_{\text{distance}} < 0$, it is a decreasing and convex function of distance.

The shape of the VTTS function with respect to travel time is driven by the sign and the magnitude of λ_{ivtime}. If $\lambda_{\text{ivtime}} > 2$, the VTTS function is increasing and strictly convex with respect to in-vehicle travel travel time. If $\lambda_{\text{ivtime}} \in]1, 2]$, it is an increasing and concave function of in-vehicle travel time. If $\lambda_{\text{ivtime}} < 1$, it is a decreasing and convex function of in-vehicle travel time. Finally, note that $\lambda_{\text{ivtime}} = 1$ means that the VTTS function does not depend on in-vehicle travel time.

The shape of the VTTS function with respect to travel cost is driven by the sign and the magnitude of λ_{cost}. If $\lambda_{\text{cost}} > 1$, the VTTS function is decreasing and concave with respect to travel cost. If $\lambda_{\text{cost}} \in [0, 1]$, it is an increasing and concave function of travel cost. If $\lambda_{\text{cost}} < 0$, it is an increasing and convex function of travel cost. Finally, note that $\lambda_{\text{cost}} = 1$ means that the VTTS function does not depend on travel cost.

Figure 1 depicts the theoretical shapes of the VTTS function with respect to travel time and travel cost. The theoretical shapes of the VTTS function with respect to distance and income are similar.

Figure 1 about here

4 Results

4.1 Estimation strategy and model selection

In our approach, we have not estimated directly the whole specification as described in subsection 3.1. We have carried out the work step by step. Our methodology is summarized in table 2. Actually, we have estimated several nested models for each country to build our comments on the "most reliable" one. Reliability is here defined by statistical goodness-of-fit criteria.

Table 2 about here

Firstly, we have distinguished MNL models from Box-Cox MNL models. Secondly, we have also distinguished between models with deterministic taste heterogeneity and models without deterministic taste heterogeneity. Thirdly, we have distinguished between models with unobserved taste heterogeneity and models without unobserved taste heterogeneity. Finally, we have distinguished models with random agent effects and models without random agent effects. All in all, we have estimated 16 models for each country.

The sequence of estimation of the various models is important as it defines the way we have set the initial guesses to implement the optimization algorithms to obtain the estimates on which we build up our results and comments. The baseline model is the MNL model without observed and unobserved heterogeneity, without agent effect, and without Box-Cox transformation. As the associated log-likelihood function is known to be globally concave, any vector of initial values for the parameters can be used to get the estimates. A vector of 0's is used here.

The maximum likelihood estimates have then be used as initial values to estimate the MNL model with unobserved taste heterogeneity and the MNL model with random agent effects.

The former maximum simulated likelihood estimates have been used to implement the MNL specification with deterministic taste heterogeneity and unobserved taste heterogeneity and the latter maximum simulated likelihood estimates have been used to estimate the MNL model specification with deterministic taste heterogeneity and random agent effects. The initial values of the coefficients associated to unobserved heterogeneity and individual random effects have all been set to 0.

The MNL specification with deterministic taste heterogeneity, unobserved taste heterogeneity, and random agent effects, has been estimated by using the maximum simulated likelihood estimates of the

MNL specification with deterministic taste heterogeneity and unobserved taste heterogeneity. Here again, the initial values of the coefficients associated to individual random effects have all been set to 0.

The specifications with Box-Cox transformations have been implemented by using the final estimates of the corresponding specifications without Box-Cox transformations as initial values. Every times, the initial values of the Box-Cox transformations have all been set to 1.

All the considered models have been estimated with BIOGEME software (Bierlaire (2006)) using either the baseline Trust Region algorithm or the Donlp2 algorithm (Spellucci (1993)). When necessary, 500 Halton draws have been used to approximate the multivariate integrals. Tables 3, 4, 5 report fit statistics about the estimated models which specifications are presented in table 2.

Tables 3, 4, 5 about here

For each sample, we select here the model with the lowest AIC criterion. The AIC is a penalised likelihood criterion that is used to find the model that best explains the data with a minimum of free parameters. It is not the only one that may be considered (e.g. the Schwartz criterion, also known as the Bayesian information criterion, or the Hannan-Quinn criterion). Fortunately, the selected models correspond to the most complete specification. Actually, we should reestimate each model while considering only the significant variables and then make the final selection. In the present approach, model selection is based on model specifications with the full set of explanatory variables.

4.2 MSL Estimates

For identification purpose, the mode of reference is the car mode and the baseline traveller has the following characteristics: she is a woman, aged less than 30 years old and whose level of education is a training degree. Tables 6, 7, 8 report the estimates of the selected models.

Tables 6, 7, 8 about here

4.2.1 The Czech Republic

Looking at the results of the selected model for the Czech Republic, we observe that none of the socioeconomic and demographic variables that may affect the baseline levels of the utility function are significant

except the logartihm of the income level for the air mode and the coach mode and the [31; 40] age class for the air mode.

The Box-Cox transformations are all significantly different from 0 but also from 1. They all lie in the $[0, 1]$ interval, thereby denoting that the utility function is a decreasing and strictly convex function of travel times and travel cost.

Education variables are found to play no significant roles on sensitivity to time and cost variables whereas they have significant impacts on sensitivity to the number of interchanges. As compared to education by training, all the travellers with other levels/types of education are less sensitive to the number of interchanges. One may also remark that highly educated travellers (technical level or university level) are the less sensitive to the number interchanges.

Gender plays a significant role only on sensivity to the number of interchanges. Men are less sensitive than women as it regards the number of interchanges. There is no other significant difference between men and women as it regards sensitivity to the travel attributes.

The effects of the age variables on sensitivities to the travel attributes differ from one of the latter to another. As compared to travellers less than 30 years old, travellers in between 31 and 40 years old are less sensitive to the cost variable and more sensitive to the number of interchanges. Travellers in between 41 and 60 years old are less sensitive to the "in-vehicle" time variable. There is no other significant difference between the different age class as it regards sensitivity to the travel attributes.

Income and distance explain significantly how tastes for travel attributes distribute accross the population of travellers.

Unobserved taste heterogeneity is significant when considering the "access time" variable.

4.2.2 Switzerland

Looking at the results of the selected model for Switzerland, we observe that none of the Box-Cox transformations is statistically different from 1. *Ceteris paribus*, the utility function is decreasing in a linear fashion with respect to any of the travel attributes.

The results show also that the sociodemographic and economic characteristics may significantly explain taste heterogeneity and the level of the baseline utility function.

The attention that is paid to the travel and cost attributes decreases significantly and the attention that is paid to the number of interchanges increases significantly when travellers have a university degree. Furthermore, this decrease in attention is lower when considering the time atributes than when considering the cost attribute.

Also, as compared to the others, travellers who pursued only obligatory school or a technical formation pay less more attention to access time.

Travellers in between 41 and 60 years old pay less attention to time attributes as compared to younger or older travellers. They pay also more attention to the number of interchanges. People aged older than 60 years pay also more attention the cost of the trip as compared to people aged less than 31 years old.

As compared to women, men pay less attention to the number of interchanges.

Income plays a significant role on tastes for travel time (in-vehicle and access). The richer the traveller, the higher his/her sensitivity to the time attributes.

The travelled distance plays also a significant role on the tastes for all the trip attributes. The longer the distance to travel, the lesser sensitivity of the traveller to time and cost attributes but the higher his/her sensitivity to the number of interchanges.

We observe also that unobserved taste heterogeneity is significant for the in-vehicle time and the number of interchanges variables.

As it regards the alternative specific effects, we observe that older travellers have a stronger preference to use air than car as compared to travellers who are younger than 31 years old. The results show also that travellers in between 41 and 60 years old consider less the bus alternative than the car alternative as compared to travellers who are younger than 31 years old. Finally, the travellers who are in between 31 and 40 years old have (*all other things being equal*) a lower probability to choose the train alternative as compared to the other travellers.

The travellers who pursued only obligatory school have different baseline levels of utility. *Ceteris paribus*, they have larger probability to use the train alternative or the coach alternative as compared to the other travellers.

We remark that agent effects are all statistically significant, therevy suggesting that there exists unconsidered traveller's characteristics that may affect his/her baseline levels of utility for the different modes of transport.

4.2.3 Portugal

Looking at the results of the selected model for Portugal, we observe also both significant observed and unobserved heterogeneity as well as as significant non linearity on travel attributes. Indeed, the Box-Cox transformations and their associated Student statistics show

that the utility function is decreasing and convex with respect to the "in-vehicle" travel time attribute: despite a high numerical value of the Box-Cox parameter, it is not significantly different from 0, thereby suggesting a logarithmic form of the "in-vehicle" time variable. The utility function is also decreasing but linear (i.e. the Box-Cox transformation is not statistically different from 1) with respect to the "access" travel time variable. The results show also that the Box-Cox transformation of the cost variable, although numerically greater than 1, is not statistically different from 1.

The results show also that the attention that is paid to the travel attributes decreases significantly with respect to the travel distance. Furthermore, this decrease in attention is greater when considering the time atributes than when considering the cost attribute.

Interestingly, we notice that the importance of the access time variable is significantly distributed accross the population of decision-makers whereas it is not the case for the other travel attributes. We notice also that the agent effects are significant for the train and the plane modes of transport but not for the car and the coach modes of transport, thereby suggesting that there may exist additional determinants that explain the observed choices. This might be due to the lack of information as it regards the sociodemographic and economic characteristics of the decision makers but also to some unmeasurable aspects of their behaviours.

4.3 Results on the values of travel time savings

Looking at the estimates of our selected models and computing for each individual the associated VTTS functions that have been defined in subsection 3.2[3], we clearly see that the results are heterogenous between individuals, between sociodemographic groups, and between countries.

As it regards Portugal, since there is no sociodemographic and economic variables to cluster the population, the VTTS functions depend only on the levels of travel times, the levels of travel cost, the travelled distances, and the unobserved tase heterogeneity. As the distributions of the levels of travel attributes differ from one mode of transport to another, we obtain therefore specific values of travel time savings for each of them. As it regards Switzerland and the Czech Republic, it is also the case but we can also distinguish likely differences between sociodemographic groups and the levels of income of the travellers as

[3]We use 10000 pseudo random draws to approximate the integrals

we consider that the attention that is paid to the travel attributes is function of these sociodemographic and economic characteristics.

There is however a major difficulty in presenting estimates of the VTTS as they appear to be individual-specific (through the socioeconomic characteristics), mode-specific (through the levels of travel attributes), and defined as continuous functions of the travel distance, the level of income, travel times, and travel costs. We prefer to discuss some of the estimates of the parameters of the functional form in equation 15 and presents some estimated shapes.

4.3.1 The Czech Republic

Both the "in-vehicle" VTTS function and the "access" VTTS functions are found to be decreasing functions of the travelled distance ($\delta_{\text{distance,ivtime}} - \delta_{\text{distance,cost}} = -0.87$ and $\delta_{\text{distance,acctime}} - \delta_{\text{distance,cost}} = -0.48$).

The results show also that the "in-vehicle" VTTS function is not elastic with respect to the level of income ($\delta_{\text{income,ivtime}} - \delta_{\text{income,cost}} = -0.06$). The "access" VTTS function is also not elastic with respect to the level of income (it is decreasing but not significantly with respect to it, $\delta_{\text{income,acctime}} - \delta_{\text{distance,cost}} = -0.15$).

As the utility function is decreasing and convex with respect to the time and cost attributes, the derived VTTS functions are decreasing and convex with respect to the time attributes and increasing and concave with respect to the cost attribute.

Ceteris paribus and as compared to travellers who are less than 30 years old, travellers who are in between 31 and 40 years old have a higher "in-vehicle" VTTS, travellers who are in between 41 and 60 years old have a lower "in-vehicle" VTTS, and travellers who are aged more than 61 years old have the same "in-vehicle" VTTS. Also, travellers who are more than 31 years old have a higher "access" VTTS.

Still assuming all other things being equal, there is no difference between men and women in evaluation of the VTTS functions. We remark also that, as compared to travellers with a "training" level in education, travellers with other education levels have higher "in-vehicle" and "access" VTTS.

Figures 2 and and 3 depict how evolve the VTTS functions with respect to the level of income and the travel distance for a woman aged less than 30 years old with a training level of education. It is computed at the average times and costs of each mode of transport to give a better representation of how the VTTS functions evolve also with respect to the travel attributes. Note that whatever are the characteristics of the traveller, these functions have the same shapes. They differ only up to a vertical translation on the z-axis.

Figures 2 and 3

4.3.2 Switzerland

The results show that the "in-vehicle" VTTS function is not elastic to the travelled distance ($\delta_{\text{distance,ivtime}} - \delta_{\text{distance,cost}} = 0.07$) but that the "access" VTTS function is decreasing significantly with respect to the travelled distance ($\delta_{\text{distance,acctime}} - \delta_{\text{distance,cost}} = -2.18$).

The results show also that the "in-vehicle" and the "access" VTTS functions are increasing with respect to the level of income ($\delta_{\text{income,ivtime}} - \delta_{\text{income,cost}} = 0.42$ and $\delta_{\text{income,acctime}} - \delta_{\text{income,cost}} = 1.48$)

As the Box-Cox transformations on the time and the cost attributes are not significantly different from 1, the derived VTTS functions does not depend on the latter.

Ceteris paribus and as compared to travellers who are less than 30 years old, travellers who are in between 31 and 40 years old have a higher "in-vehicle" VTTS, travellers who are in between 41 and 60 years old have a lower "in-vehicle" VTTS, and travellers who are aged more than 61 years old have a higher "in-vehicle" VTTS. Also, travellers who are in between 31 and 60 years old have a lower "access" VTTS and travellers who are aged more than 61 years old have a higher "in-vehicle" VTTS

Still assuming all other things being equal, there is no difference between men and women in evaluation of the VTTS functions. We remark also that, as compared to travellers with a "training" level in education, travellers with an "obligatory school" level of education and travellers with an "technical" level of education levels have higher "access" VTTS whereas travellers with an "university" level of education and travellers with an "highschool" level of education levels have lower "access" VTTS. In the same vein, only travellers with a "highschool" level of education have lower "in-vehicle" VTTS whereas the travellers with any other level of education have higher "in-vehicle" VTTS.

Figures 4 depicts how evolves the VTTS functions with respect to the level of income and the travel distance for a woman aged less than 30 years old with a training level of education.

Figure 4

4.3.3 Portugal

Both the "in-vehicle" VTTS function and the "access" VTTS functions are found to be decreasing functions of the travelled distance

($\delta_{\text{distance,ivtime}} - \delta_{\text{distance,cost}} = -5.40$ and $\delta_{\text{distance,acctime}} - \delta_{\text{distance,cost}} = -0.27$).

As the Box-Cox transformation on the "in-vehicle" time is not significantly different from 0, the derived VTTS functions is decreasing and convex with respect to "in-vehicle" time. The two VTTS functions do not depend on the cost variable and the "access" VTTS function does not depend on the "access" time variable.

Figure 5 depicts how evolves the "in-vehicle" VTTS functions with respect to travel distance. Figure 6 depicts how the "access" VTTS function evolve with travel distance.

Figures 5 and 6

5 Concluding remarks

There is clear evidence of the presence of nonlinearities in modelling long-distance mode choice. It is found that accounting both for observed and for unobserved taste heterogeneity improves also our understanding of choice behaviours. As it regards long distance mode choice in the three studied European countries, the results show that perception and attention paid to the travel attributes differ from one country to another but also accross the population of travellers of each of the considered countries.

There are several ways to extend our approach. For instance, further research may focus on comparing mixing distributions that model the different dimensions of unobserved heterogeneity. Actually, even though the log-normal distribution is bounded on one side, it has a long tail on the other side. Furthermore, nothing states that it is the best way to model unobserved taste heterogeneity.

Another point is that we have not analyzed the choice of a more appropriate GEV distribution to model substitution patterns between the modes of transport. It may be possible that the considered modes of transport, and also because of the aggregate nature of their definition, share some common unobserved attributes. To that extent, further research may focus on comparing either different nesting or cross-nesting structures or error component structures and on testing whether they add significant understanding of observed choice behaviors.

References

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

Becker, G. S. (1965). A theory of the allocation of time. *The Economic Journal*, 75(299):493–517.

Ben-Akiva, M. and Lerman, S. R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. The MIT Press, Cambridge, Massachusetts.

Bhat, C. R. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled halton sequences. *Transportation Research Part B: Methodological*, 37(9):837–855.

Bierlaire, M. (2006). BIOGEME: a free package for the estimation of discrete choice models. In *Swiss Transport Research Conference, Ascona, Switzerland*.

DeSerpa, A. C. (1971). A theory of the economics of time. *Economic Journal*, 81(324):828–46.

Gaudry, M. and Dagenais, M. (1979). Heteroscedasticity and the use of box-cox transformations. *Economics Letters*, 2(3):225–229.

Gaudry, M. and Laferriere, R. (1989). The box-cox transformation : Power invariance and a new interpretation. *Economics Letters*, 30(1):27–29.

Gaudry, M. and Wills, M. (1977). Estimating the functional form of travel demand models. Cahiers de recherche 7702, Universite de Montreal, Departement de sciences economiques.

Gouriéroux, C. and Monfort, A. (1996). *Simulation-based econometric methods*. Oxford University Press.

Hajivassiliou, V. A. (1997). Some practical issues in maximum simulated likelihood. STICERD - Econometrics Paper Series /1997/340, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Hajivassiliou, V. A. and Ruud, P. A. (1994). Classical estimation methods for ldv models using simulation. In Engle, R. F. and McFadden, D., editors, *Handbook of Econometrics*, volume 4, pages 2384–2441.

Hensher, D. (2006). The signs of the times: Imposing a globally signed condition on willingness to pay distributions. *Transportation*, 33(3):205–222.

Hess, S., Erath, A., and Axhausen, K. (2008). Estimates of the valuation of travel time savings in switzerland obtained from pooled data. *Transportation Research Record*, (2082):43–55.

Hess, S., Train, K. E., and Polak, J. W. (2006). On the use of a modified latin hypercube sampling (mlhs) method in the estimation of a mixed logit model for vehicle choice. *Transportation Research Part B: Methodological*, 40(2):147–163.

Jara-Diaz, S. (2000). Allocation and valuation of travel time savings. In Hensher, D. and K. Button, K., editors, *Handbook of Transportation*, pages 303–319. Pergamon Press, Oxford.

Lapparent, M. d., de Palma, A., and Fontan, C. (2002). Non-linearities in the valuations of time estimates. In *Proceedings of the PTRC Annual Meeting*.

Lee, L.-F. (1995). Asymptotic bias in simulated maximum likelihood estimation of discrete choice models. *Econometric Theory*, 11(3):437–483.

Mackie, P., Jara-diaz, S., and Fowkes, A. (2001). The value of travel time savings in evaluation. *Transportation research part E*, 37:91–106.

Mackie, P., Wardman, M., Fowkes, A., Whelan, G., Nellthorp, J., and Bates, J. (2003). Values of travel time savings in the UK. Technical report, Institute of Transport Studies, University of Leeds, Leeds, UK.

Marschak, J. (1959). Binary choice constraints on random utility indicators. Cowles Foundation Discussion Papers 74, Cowles Foundation, Yale University.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarembka, P., editor, *Frontiers in Econometrics*, pages 105–142. Academic Press, New York.

McFadden, D. (2005). Revealed stochastic preference: a synthesis. *Economic Theory*, 26(2):245–264.

McFadden, D. and Richter, M. (1970). On the extension of a probability to the boolean algebra generated by a family of events. Technical report, University of California, Berkeley.

Pakes, A. and Pollard, D. (1989). Simulation and the asymptotics of optimization estimators. *Econometrica*, 57(5):1027–57.

Rose, J. and Scarpa, R. (2007). Designs efficiency for non-market valuation with choice modelling: How to measure it, what to report and why. Working Papers in Economics 07/21, University of Waikato, Department of Economics.

Spellucci, P. (1993). A SQP method for general nonlinear programs using only equality constrained subproblems. *Mathematical Programming*, 82:413–448.

Spitzer, J. (1984). Variance estimates in models with the box-cox transformations: Implications for estimation and hypothesis testing. *Review of Economics and Statistics*, 66:645–652.

Train, K. (2000). Halton sequences for mixed logit. Economics Working Papers E00-278, University of California at Berkeley.

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

Walker, J. L., Ben-Akiva, M., and Bolduc, D. (2007). Identification of parameters in normal error component logit-mixture (neclm) models. *Journal of Applied Econometrics*, 22(6):1095–1125.

Figure 1: Shapes of the VTTS as function of the Box-Cox transformations, *ceteris paribus*

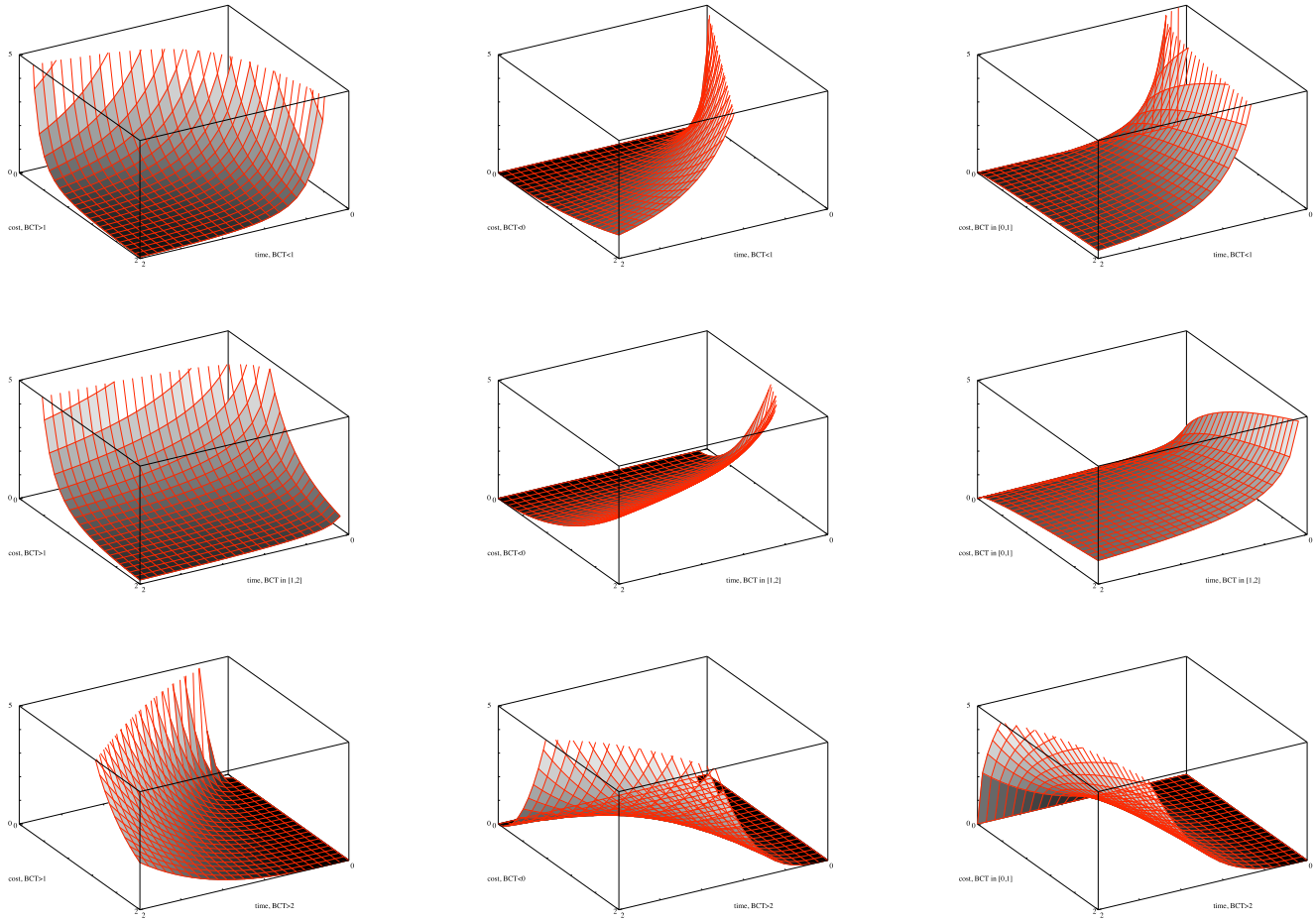


Table 1: Data description, SP sample

Label	Czech Republic #[a] of obs.[b] = 2044 # of DM.[c] = 511 mean	std.dev.[d]	freq.	Switzerland # of obs. = 916 # of DM. = 229 mean	std.dev.	freq.	Portugal # of obs. = 148 # of DM. = 37 mean	std.dev.	freq.
SP variables[e]									
Choice: car mode			1488			528			112
IV.[f] time in mn.[g], car	341.59	201.17		458.49	157.75		467.34	176.50	
Cost in €[h], car	43.69	25.56		151.90	51.08		159.20	59.18	
Choice: train mode			216			252			12
IV time, train	374.15	224.19		497.83	171.65		499.21	198.93	
Acc.[i] time in mn, train	9.16	4.05		8.88	3.98		8.65	3.98	
Cost, train	27.12	16.30		138.87	48.58		141.98	56.87	
# of interchanges, train	0.89	0.85		0.87	0.84		0.82	0.88	
Choice: air mode			44			104			12
IV time, air	86.06	51.30		114.25	39.77		116.53	45.40	
Acc. time, air	119.97	25.83		119.38	25.99		120.41	26.88	
Cost, air	318.49	55.42		310.31	52.91		317.03	56.40	
# of interchanges, air	1.00	0.86		0.97	0.85		1.01	0.90	
Choice: coach mode			268			28			12
IV time, coach	325.83	196.20		423.17	146.73		433.33	172.65	
Acc. time, coach	58.30	24.78		57.35	24.65		57.57	25.11	
Cost, coach	20.84	12.46		159.55	24.91		164.49	26.32	
# of interchanges, coach	1.00	0.86		1.03	0.88		0.99	0.88	
Socioeconomic variables[j]									
Dist.[k] of ref.[l] trip in km.[m]	258.05	145.47		344.03	104.41		347.99	120.48	
Income	2826.06	1341.41		4521.03	1761.29			–	
Income< 2000€			128			25		–	
Income∈ [2000; 5999]€			211			121		–	
Income≥ 6000€			18			35		–	
Age≤ 30 years old			120			33		–	
Age∈ [31; 40]			100			44		–	
Age∈ [41; 60]			197			106		–	
Age≥ 61			93			46		–	
Education: obli.[n] school			101			13		–	
Education: training			166			111		–	
Education: highschool			182			12		–	
Education: tech.[o]			8			27		–	
Education: university			53			56		–	
Gender: man			248			115		–	
Gender: woman			263			114		–	

[a] #: number
[b] obs.: observations
[c] DM.: decision makers, i.e. individuals
[d] std.dev.: standard deviation
[e] Descriptive statistics based on the number of observations
[f] IV.: in-vehicle
[g] mn.: minutes
[h] €: Euro
[i] Acc.: access
[j] Descriptive statistics based on the number of individuals
[k] Dist.: distance of baseline trip used to generate SP experiments
[l] ref.: reference
[m] km.: kilometres
[n] obli.: obligatory
[o] tech.: further technical formation

Table 2: Description of model specifications

Model name	Box-Cox transf.[a]	Determ.[b] het.[c]	Random het.	Random agent effects
M_1	–	–	–	–
M_2	–	x	–	–
M_3	–	–	x	–
M_4	–	–	–	x
M_5	–	x	x	–
M_6	–	x	–	x
M_7	–	–	x	x
M_8	–	x	x	x
M_9	x	–	–	–
M_10	x	x	–	–
M_11	x	–	x	–
M_12	x	–	–	x
M_13	x	–	x	x
M_14	x	x	x	–
M_15	x	x	–	x
M_16	x	x	x	x

[a]transf.: transformation
[b]Determ: deterministic
[c]het.: heterogeneity

Table 3: Goodness-of-Fit of models, The Czech Republic

Model name	ℓ_{final} [a]	$\ell_{\text{int\_only}}$ [b]	ℓ_0 [c]	# of par. P [d]	LR. stat. [e]	Adj. ρ^2 [f]	AIC [g]
M_1	-1632.88	-2217.00	-2833.59	34	2401.42	0.41	3333.76
M_2	-1453.25	-2217.00	-2833.59	74	2760.67	0.46	3054.50
M_3	-1437.60	-2217.00	-2833.59	38	2791.98	0.48	2951.20
M_4	-1422.94	-2217.00	-2833.59	38	2821.29	0.48	2921.88
M_5	-1365.31	-2217.00	-2833.59	78	2936.55	0.49	2886.62
M_6	-1309.78	-2217.00	-2833.59	78	3047.62	0.51	2775.56
M_7	-1398.65	-2217.00	-2833.59	42	2869.87	0.49	2881.30
M_8	-1300.34	-2217.00	-2833.59	82	3066.49	0.51	2764.68
M_9	-1514.57	-2217.00	-2833.59	37	2638.03	0.45	3103.14
M_10	-1399.44	-2217.00	-2833.59	77	2868.29	0.48	2952.88
M_11	-1364.94	-2217.00	-2833.59	41	2937.29	0.50	2811.88
M_12	-1340.55	-2217.00	-2833.59	41	2986.07	0.51	2763.10
M_13	-1318.37	-2217.00	-2833.59	45	3030.44	0.52	2726.74
M_14	-1352.77	-2217.00	-2833.59	81	2961.64	0.51	2866.54
M_15	-1298.35	-2217.00	-2833.59	81	3070.48	0.51	2757.70
M_16	-1274.04	-2217.00	-2833.59	85	3119.10	0.52	2718.08

[a] Value of the loglikelihood function at point of convergence
[b] Value of the loglikelihood function with only intercept terms as variables
[c] Value of the loglikelihood function with parameters all equal to 0
[d] Number of parameters
[e] Likelihood Ratio statistic, $2 * (\ell_{\text{final}} - \ell_0)$
[f] Adjusted pseudo ρ^2, $(\ell_0 - \ell_{\text{final}} + P) / \ell_0$
[g] Akaike Information Criterion, defined as $2P - 2\ell_{\text{final}}$

Table 4: Goodness-of-Fit of models, Switzerland

Model name	ℓ_{final}	$\ell_{\text{int\_only}}$	ℓ_0	# of par. P	LR. stat.	Adj. ρ^2	AIC
M_1	-751.22	-1035.33	-1269.85	34	1037.25	0.38	1570.44
M_2	-679.33	-1035.33	-1269.85	74	1181.02	0.41	1506.66
M_3	-603.92	-1035.33	-1269.85	38	1331.85	0.49	1283.84
M_4	-578.68	-1035.33	-1269.85	38	1382.34	0.51	1233.36
M_5	-561.81	-1035.33	-1269.85	78	1416.07	0.50	1279.62
M_6	-515.18	-1035.33	-1269.85	78	1509.32	0.53	1186.36
M_7	-556.19	-1035.33	-1269.85	42	1427.31	0.53	1196.38
M_8	-509.93	-1035.33	-1269.85	82	1519.84	0.53	1183.86
M_9	-739.62	-1035.33	-1269.85	37	1060.45	0.39	1553.24
M_10	-672.97	-1035.33	-1269.85	77	1193.75	0.41	1499.94
M_11	-595.50	-1035.33	-1269.85	41	1348.70	0.50	1273.00
M_12	-566.66	-1035.33	-1269.85	41	1406.37	0.52	1215.32
M_13	-556.48	-1035.33	-1269.85	45	1426.72	0.53	1202.96
M_14	-550.56	-1035.33	-1269.85	81	1438.58	0.48	1263.12
M_15	-510.80	-1035.33	-1269.85	81	1518.10	0.51	1183.60
M_16	-504.06	-1035.33	-1269.85	85	1531.58	0.53	1178.12

Table 5: Goodness-of-Fit of models, Portugal

Model name[a]	ℓ_{final}	$\ell_{\text{int\_only}}$	ℓ_0	# of par. P	LR. stat.	Adj. ρ^2	AIC
M_1	-116.45	-160.93	-205.17	7	177.44	0.40	246.90
M_2	-100.61	-160.93	-205.17	11	209.124	0.46	223.22
M_3	-96.31	160.93	-205.17	11	217.72	0.48	214.62
M_4	-93.40	-160.93	-205.17	11	223.54	0.49	208.80
M_5	-91.35	-160.93	-205.17	15	227.64	0.48	212.70
M_6	-84.82	-160.93	-205.17	15	240.71	0.51	199.64
M_7	-86.79	-160.93	-205.17	15	236.76	0.50	203.58
M_8	-81.75	-160.93	-205.17	19	246.85	0.51	201.5
M_9	-113.35	-160.93	-205.17	10	183.64	0.40	246.70
M_10	-99.78	-160.93	-205.17	14	210.78	0.45	227.56
M_11	-93.68	-160.93	-205.17	14	222.98	0.48	215.36
M_12	-91.25	-160.93	-205.17	14	227.84	0.49	210.5
M_13	-84.42	-160.93	-205.17	18	241.51	0.50	204.84
M_14	-90.05	-160.93	-205.17	18	230.25	0.47	216.10
M_15	-82.51	-160.93	-205.17	18	245.33	0.51	201.02
M_16	-78.05	-160.93	-205.17	22	254.24	0.51	200.10

[a]Reminder: as it was not possible to draw the socioeconomic variables that characterize the decision makers who participated to the SP survey from the RP survey, the models with deterministic taste heterogeneity include only the logarithm of the distance.

Table 6: MSL estimates (500 Halton draws), The Czech Republic

Taste heterogeneity								
	IV time		Acc time		Cost		# interchanges	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
Intercept	-2.02	-4.38	-3.47	-5.51	-6.98	-15.85	-77.20	-8.66
ln Dist. trip in km.	-0.66	-4.87	-0.27	-1.95	0.21	3.95	9.85	33.42
ln Income in €	0.34	3.21	0.25	4.95	0.40	31.85	0.59	1.74
Age$\in [31;40]$	-0.41	-1.13	-0.29	-0.63	-0.64	-1.82	17.70	2.16
Age$\in [41;60]$	-0.33	-2.36	0.06	0.19	-0.17	-1.03	-4.80	-0.61
Age≥ 61	-0.18	-0.48	0.18	0.21	-0.16	-0.23	12.90	1.55
Education: obli.	-0.13	-0.35	0.19	0.39	-0.30	-1.42	-5.34	-6.11
Education: highschool	-0.21	-1.29	0.14	0.47	-0.28	-0.67	-5.90	-5.79
Education: tech.	-0.04	-0.07	0.53	0.72	0.25	0.22	-18.10	-3.44
Education: university	-0.03	-0.05	0.01	0.05	-0.25	-0.82	-17.80	-12.98
Gender: man	0.03	0.19	-0.05	-0.15	0.003	0.01	-5.22	-5.91
Unobserved terms	0.19	1.26	0.43	2.37	0.06	0.69	0.13	0.43
Box-Cox transformations								
	IV time		Acc time		Cost		# interchanges	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
	0.74	50.22	0.58	4.32	0.29	4.28	–	–
Alternative specific effects								
	Car		Train		Air		Coach	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
Intercept	–	–	0.70	1.13	-14.50	-6.87	1.87	1.14
ln Income in €	–	–	-0.20	-0.66	2.97	25.18	-0.89	-7.12
Age$\in [31;40]$	–	–	-0.14	-0.15	-5.10	-2.12	0.61	0.22
Age$\in [41;60]$	–	–	-0.24	-0.12	1.03	0.53	0.63	0.61
Age≥ 61	–	–	-0.54	-0.64	0.42	0.15	0.98	0.27
Education: obli.	–	–	-0.28	-0.38	-2.41	-0.70	0.61	0.14
Education: highschool	–	–	-0.20	-0.11	-1.16	-0.43	0.90	0.65
Education: tech.	–	–	1.70	1.12	5.40	0.55	0.80	0.17
Education: university	–	–	0.54	0.39	-3.51	-1.17	-0.07	-0.02
Gender: man	–	–	-0.18	-0.10	-0.76	-0.73	-0.14	-0.08
Agent effects	1.50	1.36	1.44	1.53	0.44	1.35	0.19	0.08

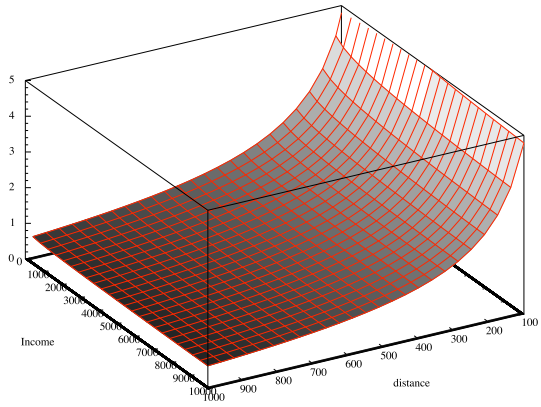
Table 7: MSL estimates (500 Halton draws), Switzerland

Taste heterogeneity								
	IV time		Acc time		Cost		# interchanges	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
Intercept	-0.92	-0.43	3.17	1.21	0.75	0.39	-18.20	-1.78
ln Dist. trip in km.	-1.01	-5.60	-3.26	-4.69	-1.08	-5.04	2.98	1.88
ln Income in €	0.46	2.01	1.52	3.52	0.04	0.19	-0.38	-0.56
Age$\in [31;40]$	-0.31	-1.15	-21.50	-1.04	0.52	1.98	0.89	2.35
Age$\in [41;60]$	-0.60	-2.38	-1.09	-2.81	0.09	0.40	1.62	2.03
Age≥ 61	0.20	0.67	0.89	1.52	0.58	2.18	-0.29	-0.63
Education: obli.	-0.37	-0.97	1.95	3.30	-5.45	-0.06	-0.41	-0.28
Education: highschool	-0.02	-0.05	0.002	0.01	0.13	0.35	1.91	1.27
Education: tech.	-0.07	-0.24	1.27	3.03	-0.60	-2.82	-12.60	-1.16
Education: university	-0.57	-3.03	-0.99	-2.52	-0.74	-3.01	3.54	2.63
Gender: man	-0.03	-0.20	0.02	0.11	0.07	0.40	-2.97	-3.29
Unobserved terms	0.36	7.09	0.06	1.20	0.02	0.84	2.49	2.56
Box-Cox transformations								
	IV time		Acc time		Cost		# interchanges	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
	0.99	5.75	0.77	3.02	1.38	5.34	–	–
Alternative specific effects								
	Car		Train		Air		Coach	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
Intercept	–	–	-16.80	-1.50	1.22	0.10	-9.48	-0.81
ln Income in €	–	–	2.13	3.52	-0.26	0.19	0.85	-0.56
Age$\in [31;40]$	–	–	-2.68	-2.05	4.49	3.00	-5.18	-2.53
Age$\in [41;60]$	–	–	-1.03	-0.65	3.44	1.93	-2.25	-1.76
Age≥ 61	–	–	-0.52	-0.30	5.82	1.98	0.52	0.23
Education: obli.	–	–	4.14	2.59	-2.97	-0.87	5.25	2.15
Education: highschool	–	–	-2.78	-2.02	0.73	0.30	-2.92	-1.84
Education: tech.	–	–	0.33	0.20	3.28	0.94	1.76	0.96
Education: university	–	–	-0.08	-0.06	-1.69	-0.84	-0.01	-0.01
Gender: man	–	–	0.80	0.81	-0.23	-0.22	0.68	0.72
Agent effects	2.34	2.94	3.12	3.44	4.84	3.56	1.86	3.82

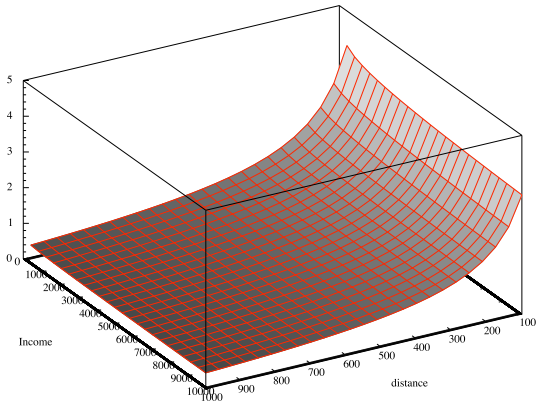
Table 8: MSL estimates (500 Halton draws), Portugal

Taste heterogeneity								
	IV time		Acc time		Cost		# interchanges	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
Intercept	13.40	1.70	4.79	2.64	-1.06	-0.66	-2.04	-0.48
ln Dist. trip in km.	-6.16	-4.13	-1.03	-3.05	-0.76	-2.66	0.33	0.43
Unobserved terms	0.30	1.30	0.41	3.32	0.07	1.11	0.97	0.84
Box-Cox transformations								
	IV time		Acc time		Cost		# interchanges	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
	3.87	0.78	0.66	2.44	1.46	2.57	–	–
Alternative specific effects								
	Car		Train		Air		Coach	
	est.	t-stat.	est.	t-stat.	est.	t-stat.	est.	t-stat.
Intercept	–	–	2.57	1.96	21.50	2.43	0.64	0.40
Agent effects	1.57	1.29	3.93	2.17	2.68	2.72	0.44	0.80

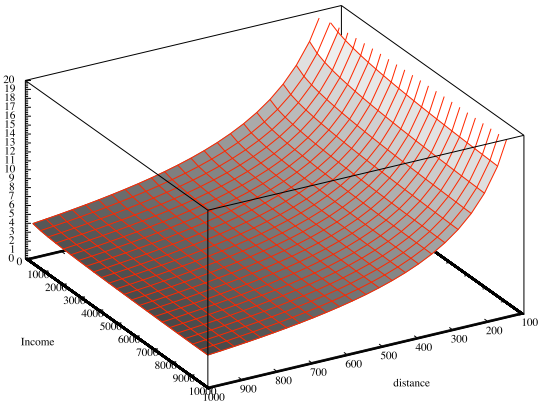
Figure 2: The Czech Republic: VTTS in € per mn at the mean IV time and mean cost for each mode



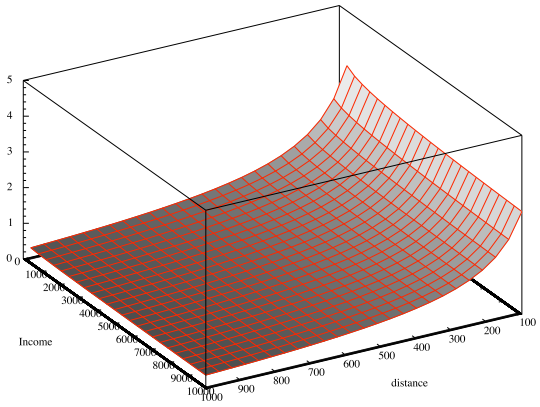
Car mode



Train mode

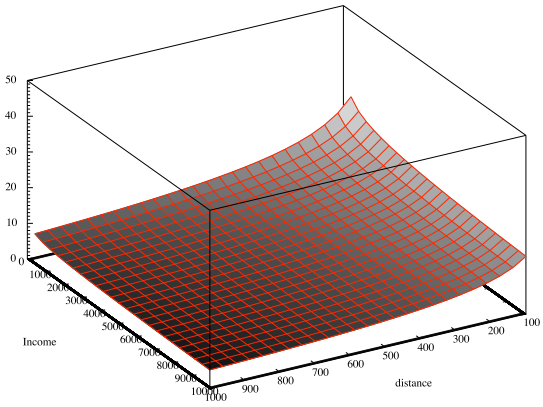


Air mode

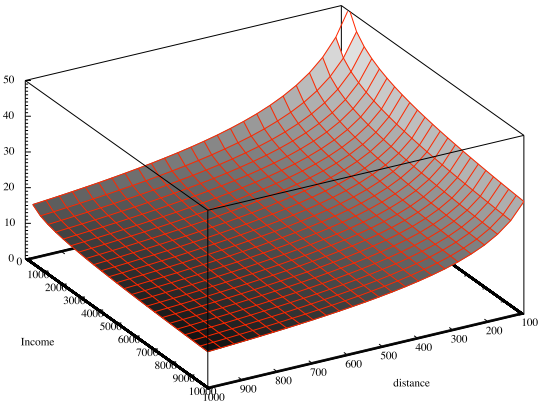


Coach mode

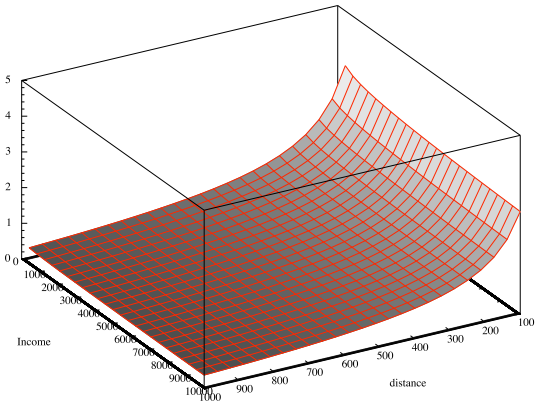
Figure 3: The Czech Republic: VTTS in € per mn at the mean access time and mean cost for each mode



Train mode

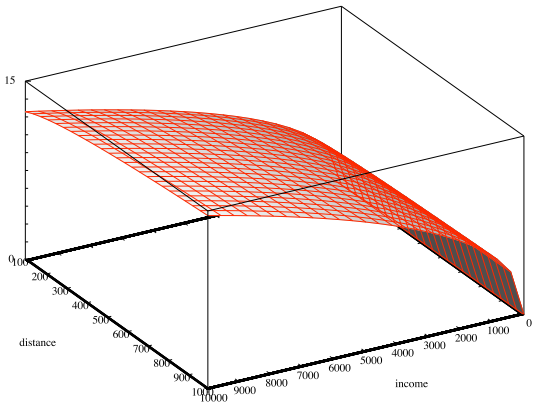


Air mode

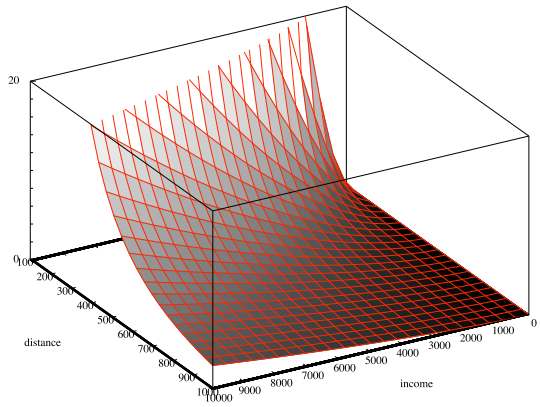


Coach mode

Figure 4: Switerland: VTTS in € per mn

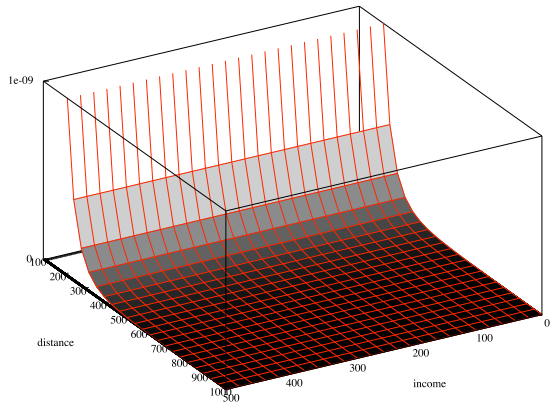


IV time

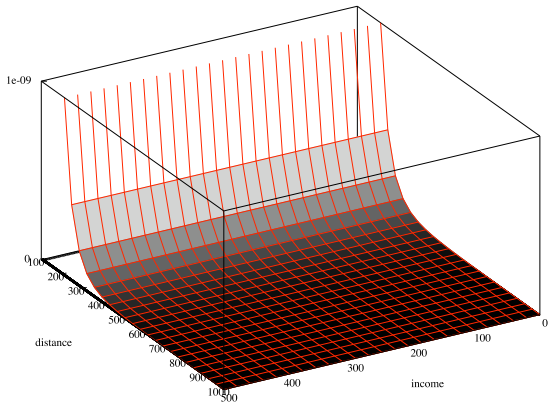


Access time

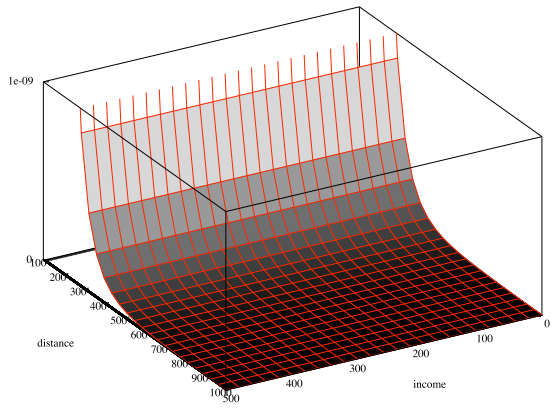
Figure 5: Portugal:IV VTTS in € per mn



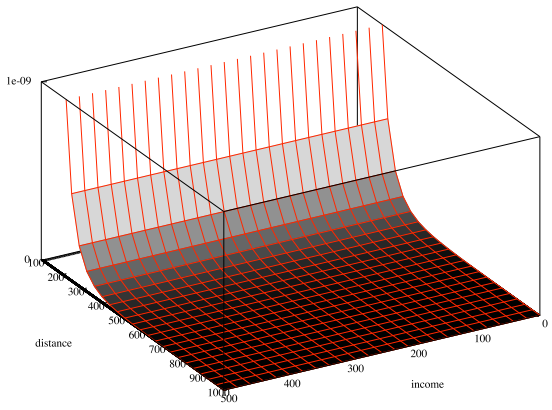
Train mode



Air mode



Air mode



Coach mode

Figure 6: Portugal:Access VTTS in € per mn