

Using semi-open questions to integrate perceptions in choice models

Aurélie Glerum * Bilge Atasoy * Michel Bierlaire *

13th December 2013

Report TRANSP-OR 120325
Transport and Mobility Laboratory
Ecole Polytechnique Fédérale de Lausanne
`transp-or.epfl.ch`

Abstract

This research investigates the measurement of perceptions by means of adjectives freely reported by respondents in semi-open questions. It involved the use of semi-open responses of 1'763 Swiss individuals to develop indicators for a latent variable representing the perception of comfort of public transportation. The indicators are then incorporated into a discrete choice model of revealed mode choices.

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

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

Key words

Mode choice models, latent variables, perceptions, qualitative data, hybrid choice models.

*ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE (EPFL), School of Architecture, Civil and Environmental Engineering (ENAC), Transport and Mobility Laboratory (TRANSP-OR), {aurelie.glerum, bilge.kucuk, michel.bierlaire}@epfl.ch

1 Introduction

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

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

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

Open questions can be structured in various ways. In this research we focus on a subset of them, which we denote by *semi-open questions*, and where we ask the respondents to report several adjectives describing best a variable of interest (e.g. a transportation mode). This data

collection technique has been used in psychology to obtain a representation of a person (Potkay and Allen, 1973) and more recently in social sciences to obtain representations of transportation modes (Kaufmann et al., 2001, 2010). Though this data collection method and the well-known *word association* technique may show some common features, we distinguish them. In word association surveys, individuals are required to report the ‘first single word which comes to mind’ (Cramer, 1968) when they are exposed to a stimulus word. In semi-open questions, respondents are required to reply to a more precise question and give a more structured answer.

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

In this paper, we consider the HCM framework and illustrate our methodology by modeling the impact of the LV ‘perception of comfort of public transportation’ on choice. Since we are using responses to semi-open questions as measurements of the LV, a method to quantify the responses to the semi-open questions is required. We propose a method which consists of asking a set of individuals (the ‘evaluators’) to rate the adjectives on a scale relative to the LV, where a positive number indicates a positive representation of the LV and a negative number implies a negative representation of it. The use of individuals to evaluate qualitative concepts is common practice in computer science studies (*Amazon’s Mechanical Turk*, 2005-2012; Franklin et al., 2011) and research on natural language processing (Snow et al., 2008). Humans happen to be more effective than computers for some specific tasks, such as evaluating a language (Snow et al., 2008; Sorci

et al., 2010; Robin et al., 2011; Venetis et al., 2012). Moreover, in medical and social sciences scores from a number of raters are used for decision-making. In such domains it is critical to have an agreement between the raters in order to be fair to the subjects. Therefore there is a well-studied literature on how to minimize the impact of the rater factor on the data quality. We refer to the review by Banerjee et al. (1999) for the analysis of agreement between raters which is called as *inter-rater reliability*. However in our case, we are interested in the heterogeneity of raters and we analyze the impact of different ratings on the results. Therefore we do not seek for inter-rater reliability but rather analyze the differences between ratings.

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

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

This research is illustrated by a case study which aims at analyzing individuals' transportation mode choice in low-density areas of Switzerland. A revealed preferences (RP) survey was conducted in the context of a joint project between PostBus, one of the major companies operating in such regions, and EPFL's Transportation Center (TraCe). In this survey, inhabitants were asked to describe all trips performed on a particular day as well as the chosen transport modes. Moreover, they were asked to report adjectives which describe best a list of transportation modes. The field of possible answers for such questions is wide and therefore allows for the analysis of a variety

of themes related to perception. For instance, one could model perception of flexibility, environmental impact or reliability of a particular transport mode. In this paper the analysis of the impact on choice of individuals' perception of comfort of public transportation is presented. For previous work on this research, we refer to Glerum et al. (2011).

The paper is structured as follows. Section 2 presents the data collection phase. Section 3 describes the specification of the HCM and explains how qualitative adjective data are integrated into the framework. Section 4 presents the estimation results. Section 5 presents a validation of the HCM. Section 6 presents a forecasting analysis by means of demand indicators derived for all evaluators. Section 7 discusses the implications of the present data collection and modeling approach on future questionnaire designs. Section 8 concludes by presenting some possible further developments.

2 Data collection

In order to analyze the transportation mode choices of inhabitants of regions which are loosely connected, an RP survey was conducted between 2009 and 2010. Two copies of a questionnaire were sent to each household of 57 towns or villages connected by post busses. The towns and villages were selected in order to be representative of the whole network of PostBus and respondents of 16 years and over were asked to answer the questionnaire. Individuals had to answer the survey in German or French, depending on the language region they were living in. In total, 20'138 households were surveyed and 1'763 valid questionnaires from 1'326 different households were eventually collected. Further details about the data collection procedure can be found in Bierlaire et al. (2011).

The following sections present the structure of the survey and the quantification process of the indicators of perception.

2.1 Revealed preferences survey

In the RP survey, information was collected on all the trips respondents performed in one day, including the mode(s) they used, the trip duration and the cost of fuel or public transport ticket, on their opinions on a list of statements, on their mobility habits, household structure and socio-economic information. In total, 2'265 trips were identified.

In addition, respondents were questioned about their perceptions of transport modes. Precisely, they were asked to report three adjectives describing best each transport mode of a given list. The answers to these questions provide spontaneous indicators of individuals' perceptions, since they

consist of words which they freely reported. We refer to these type of questions as *semi-open questions*.

Let us note that due to the inaccuracy of the durations and costs reported by the respondents for each of their trips, we imputed times and costs based on the websites of the Swiss railways (SBB) <http://www.cff.ch> and of ViaMichelin <http://fr.viamichelin.ch>. The same websites were used to impute the times and costs for the non-chosen alternatives. For the travel cost, the marginal cost of the trip is considered. More precisely, for car/motorbike trips, the only considered costs are the fuel costs for the associated trip. Similarly for public transport alternatives, the cost of the trip is calculated by accounting for the possible ownership of passes, i.e. for individuals that do not own any pass, this cost is set to the standard trip fare indicated by the SBB website, while for individuals who do own one, it is set to 0. However costs of subscriptions to the PT system (i.e. yearly passes, monthly passes, etc.) are ignored.

In addition to the imputation of the travel time and costs, we aggregated the modes reported by the individuals in three groups: public transportation modes, private motorized modes and soft modes (see Section 3 for the description). This aggregation was done since the interest of the paper is to take into account the perceptions of individuals in order to better understand their preferences towards public, private and soft modes, rather than the specific choices of train, bus, motorbike, etc.

Some bias occurred in the responses to the survey. In particular, individuals with a high education level, with ages between 40 and 79 years, or men had a higher response rate than other categories. To evaluate the demand for the different transportation modes, we correct the response bias by introducing a sample weight w_n for each individual n in the computation of the indicators. These weights are computed using the *iterative proportional fitting (IPF)* algorithm. This procedure enables us to make our sample representative of the population with respect to its marginal distributions of education, age and gender. We have made the assumption that self-selection only occurs according to exogenous socio-economic characteristics. However it could also occur in the response behavior, implying that individuals using public transportation may show a higher response rate than car users (Brownstone, 1998). For that purpose, we would also need to weight the sample according to the transport mode shares in the Swiss population. We leave this for future research.

More detailed information on the survey can be found in Atasoy et al. (2013).

2.2 Semi-open questions

In the section of the questionnaire on the perception of the transport modes, respondents had to report three adjectives describing best the following set of transport modes: car, train, bus/metro/tram, post bus, bike or walk. As an illustration, Table 1 presents the semi-open questions as they appeared in the questionnaire.

For each of the following transport modes, give three adjectives that describe them best according to you.

		Adjective 1	Adjective 2	Adjective 3
1	The car is:			
2	The train is:			
3	The bus, the metro and the tram are:			
4	The post bus is:			
5	The bicycle is:			
6	The walk is:			

Table 1: Semi-open questions on the perception of several transport modes.

With the help of social scientists, we grouped these adjectives into several themes including comfort, perception of time, perception of cost, difficulty of access, flexibility, efficiency, reliability, environmental impact, appreciation, feeling or look. The adjectives classified within each theme provide information which is assumed to reflect closely each respondent's perception of each topic, as they are freely reported.

In the model presented in this paper (see Section 3 for the specification), we are interested in evaluating the effect of the perception of one of the characteristics of transport modes listed above, that is *comfort of public transportation*, on the mode choice. Hence, we use the adjective data classified in the theme of comfort and relative to modes train, bus/metro/tram and post bus as indicators of this particular attribute. As respondents reported three adjectives for each of the three public transportation alternatives we have at most nine indicators of the perception of comfort of public transportation per individual. The variable of comfort of public transportation was selected due to the fact that a large number of related adjectives were identified in the procedure of classification into themes.

About 16% of the answers to the nine indicators were classified as related to comfort (see Table 2). Despite this fairly high proportion of relevant adjectives, about 10% of the respondents

who provided answers to the semi-open questions did not provide any adjective related to comfort (see Table 3). These 10% cases are hence lost and will not provide information to the model.

	Percentage
Related to comfort	15.58
Unrelated to comfort	53.80
Missing	30.62

Table 2: Percentages of adjectives related to comfort, unrelated to comfort and missing, among the responses to the 9 indicators related to public transportation (sample size = 2'265).

Number of adjectives unrelated to comfort	0	1	2	3	4	5	6	7	8	9
Percentage	0.05	2.80	4.90	7.75	8.18	12.21	22.05	18.45	13.61	10.01

Table 3: Percentages of 0, ..., 9 adjectives unrelated to comfort, among the responses to the 9 indicators related to public transportation (sample size = 2'265). Missing values are not considered in the computation of the percentages.

In order to use the nine indicators as a measure of the perception of comfort, they need to be quantified, i.e. we need to find a scale of 'comfort'. Therefore we designed another survey dedicated to the quantification of the adjectives.

2.3 Ratings of 'evaluators'

We conducted an online survey for the quantification of all adjectives which were identified to be related to comfort. We asked 25 students and employees of EPFL to fill it in. Since an important proportion of these individuals were neither speaking German nor French, the adjectives were translated into English. We are aware that the translation of the adjectives might generate a bias and leave this problem for future research. The exhaustive list of adjectives related to comfort is presented in Table 4¹. The students and employees of EPFL were requested to rate how strongly each adjective characterizes the concept of comfort on a five-point scale from -2 to 2. A positive rating means that the evaluator associates the adjective with being *comfortable* and a negative rating corresponds to *discomfort*. The evaluators were asked to report 0 if the adjective is not related to the concept of comfort in transport modes.

In Table 4 we report basic statistics on the reported ratings by the 25 evaluators. On the one hand, some of the adjectives are perceived in a similar way by the evaluators such as *comfortable*

¹Some of the words reported by respondents are actually not adjectives. They were nevertheless included in the analysis.

having a lower standard deviation compared to others. On the other hand, there are some adjectives with a standard deviation higher than 1 such as *hardly full* and *packed*. Such ambiguities are the main driving force of this study for the quantification of indicators of perceptions.

For illustration, we selected two evaluators among the 25, using the following procedure based on the Euclidean distances between the ratings of the evaluators. The first one, called the *central evaluator*, is selected such that the maximum distance from all other evaluators is the smallest. The other evaluator (*outlying evaluator*) is the one that is the furthest from the central evaluator. We report the ratings of these two evaluators in Table 4. It is observed that the central evaluator is closer to the mean value compared to the outlying evaluator as expected. The outlying evaluator has a different perception of the adjectives in terms of their positive/negative characterization of comfort. In Section 4 we present the estimation results for these two evaluators to see the effect of different ratings on the estimated parameters. We also include the estimation results of a model using the median ratings of the adjectives. These median ratings are provided in the third column of Table 4. Furthermore, in Section 6 we provide a sensitivity analysis of demand indicators including market shares, elasticities and VOTs, across all 25 evaluators.

Adjectives	Mean	Median	Standard deviation	Central evaluator	Outlying evaluator
bad air	-1.52	-2	0.714	-1	1
bumpy	-1.08	-1	0.862	-1	1
comfortable	1.72	2	0.542	1	2
difficult	-1.00	-1	0.707	-1	-2
empty	0.880	1	0.726	1	1
expensive	-0.680	0	0.988	-2	-1
fast	1.04	1	0.735	2	1
full	-1.00	-1	1.00	-2	2
hard	-0.920	-1	0.640	-1	-1
hardly full	-0.280	0	1.28	1	2
irritating	-1.44	-2	0.870	-1	1
packed	-0.880	-1	1.20	-2	1
relaxing	1.72	2	0.737	1	1
restful	1.44	2	0.821	1	2
shaking	-1.08	-1	0.997	-1	1
stressful	-1.44	-2	0.917	-1	1
suffocating	-1.60	-2	0.817	-1	1
tiring	-1.16	-1	0.800	-1	1
uncomfortable	-1.56	-2	0.870	-1	2
unsuitable with bags	-0.920	-1	0.812	-1	2
unsuitable with strollers	-0.720	-1	0.891	-1	1
without stress	1.40	2	0.866	2	2

Table 4: Basic statistics on the ratings of evaluators (sample size = 25).

3 Model specification

The choice is defined as the transport mode used by the respondents. It is one of the three following categories:

- *Public transportation (PT) modes*, such as bus, train, etc.

- *Private motorized modes (PMM)*, such as car, motorbike, etc.
- *Soft modes (SM)*, such as walk or bike.

The respondents' mode choices are analyzed on *tours* and not on single trips, that is, one observation corresponds to one sequence of trips starting from each respondent's home and ending at the same place. For example, one simple tour can consist of a sequence of starting points and destinations *home-work-home*. A longer tour such as *home-work-shopping-home* can include an additional trip for shopping.

Hence, the choice variable is defined as the set of modes, i.e. PT, PMM or SM, used on each tour.

In the DCM, the deterministic parts of the utility functions are given as follows:

$$V_{PMM} = ASC_{PMM} + \beta_{cost} \cdot cost_{PMM} + \beta_{time_{PMM}} \cdot time_{PMM} + \beta_{work_{PMM}} \cdot work + \beta_{French_{PMM}} \cdot French \quad (1)$$

$$V_{PT} = ASC_{PT} + \beta_{cost} \cdot cost_{PT} + \beta_{time_{PT}} \cdot time_{PT} + \beta_{work_{PT}} \cdot work + \beta_{French_{PT}} \cdot French + \beta_{comfort} \cdot PCPT \cdot time_{PT} \cdot \frac{1}{1000} \quad (2)$$

$$V_{SM} = \beta_{distance} \cdot distance \quad (3)$$

- We assume that the deterministic utilities V_{PMM} and V_{PT} are influenced by the travel times $time_{PMM}$ and $time_{PT}$, and travel costs $cost_{PMM}$ and $cost_{PT}$. For the SM alternative, a distance term *distance* is included.
- In addition to the characteristics of the transportation alternatives, some socio-economic variables are assumed to have an impact on the transportation mode choice, that is, variable *work* which indicates that the respondent performed *home-work-home* tours and variable *French* which indicates that the respondent resides in a French-speaking region of Switzerland. The latter is introduced into the model since the public transportation offer is slightly better in German-speaking regions than in the French-speaking ones, hence generating differences in the demand for the three types of modes. For example, the average number of daily return trips of suburban busses is 17 in French-speaking regions versus 29 in German-speaking regions (Bierlaire et al., 2011; p. 45).
- In the deterministic utility V_{PT} , a latent explanatory variable *PCPT* is also included and accounts for the image people have of comfort of PT. It is specified as an interaction with the travel time variable in order to analyze the impact of the LV on sensitivity to travel time. The main effect of LV *PCPT* is not included in the utility function, since its integration led

to the non-significance of the interaction term. Though a model including the main effect alone would be the usual selection (due to the parsimony of its specification), the model with the interaction term is adopted for the purposes of illustrating how taste heterogeneity in the sensitivity to travel time can be captured.

To avoid estimation problems that may arise when variables fall into very different ranges, we divide the time by 1000 in the interaction term.

The perception of comfort of PT *PCPT* is described by the following structural equation:

$$\begin{aligned} \text{PCPT} = & \lambda_{\text{mean}} + \lambda_{\text{French}} \cdot \text{French} + \lambda_{\text{age}_{50}} \cdot \text{age}_{50} + \lambda_{\text{active}} \cdot \text{active} \\ & + \lambda_{\text{cars}} \cdot \text{cars} + \omega, \quad \text{with } \omega \sim \mathcal{N}(0, \sigma^2). \end{aligned} \quad (4)$$

In the above equation we specify an intercept λ_{mean} and assume that several socio-economic variables have an effect on an individual's perception of comfort of PT:

- An indicator *French* of a residence in a French-speaking region;
- A variable *age₅₀* indicating that the respondent is younger than 50 years;
- A variable *active* equal to 1 if the respondent has a full-time or part-time job, and 0 for any other working status;
- A variable *cars* indicating that the respondent's household owns at least 2 cars;

A normally distributed random variable ω with mean 0 and standard deviation σ is also added as an error term.

The inclusion of the above variables in the structural equation of the latent variable model (LVM) results from an iterative model building procedure. The presented specification is hence the best which we eventually reached.

As the LV *PCPT* cannot be directly quantified by a survey question, measurement equations that relate it with indicators are specified. The indicators are the values ranging from -2 to 2 which were assigned to each adjective by the evaluators as explained in Section 2.3. The measurement equation for each of the nine indicators I_k , with $k = 1, \dots, 9$ is specified as follows:

$$I_k = \alpha_k + \gamma_k \cdot \text{PCPT} + v_k, \quad \text{with } v_k \sim \mathcal{N}(0, \sigma_k^2), \quad (5)$$

where α_k , γ_k and σ_k are parameters to be estimated.

4 Estimation results

Parameters ASC_h , with $h \in \{\text{PMM}, \text{PT}\}$, β_i , with $i \in \{\text{cost}, \text{time}_{\text{PMM}}, \text{time}_{\text{PT}}, \text{distance}, \text{work}_{\text{PMM}}, \text{work}_{\text{PT}}, \text{French}_{\text{PMM}}, \text{French}_{\text{PT}}, \text{PCPT}\}$, λ_j , with $j \in \{\text{mean}, \text{French}, \text{age}_{50}, \text{active}, \text{cars}\}$, σ , α_k , γ_k and σ_k with $k = 1, \dots, 9$ are estimated using exogenous sample maximum likelihood (ESML) (Ben-Akiva and Lerman, 1985) using the Python version of the software Biogeme (Bierlaire and Fietarison, 2009). The following likelihood function is considered:

$$\mathcal{L} = \prod_{n=1}^N \int_{\text{PCPT}} P(y_n | X_n, \text{PCPT}_n; \beta) \left\{ \prod_{k=1}^9 f(I_{kn} | \text{PCPT}_n; \alpha_k, \gamma_k, \nu_k, \sigma_k) \right\} f(\text{PCPT}_n; \lambda, \sigma) d\text{PCPT}, \quad (6)$$

where N is the total sample size, X_n is a vector of socio-economic attributes of respondent n , y_n is a vector of choice indicators for n , such that $y_{in} = 1$, if individual n selects alternative i , and $y_{in} = 0$ otherwise. In the likelihood formula, $P(y_n | X_n, \text{PCPT}_n; \beta)$ is the choice probability of n , $f(I_{kn} | \text{PCPT}_n; \alpha_k, \gamma_k, \nu_k, \sigma_k)$ is the value of the density function of I_k for n , $f(\text{PCPT}_n; \lambda, \sigma)$ is the value of the density function of PCPT for n .

Let us note that for identification purposes, α_1 is fixed to 0 and the standard deviation σ of the LV is normalized to 1. Moreover, missing values of I_{kn} do not contribute to the likelihood.

The HCMs are estimated separately for all the 25 evaluators (see Tables 10, 11 and 12 in the Appendix for the estimation results of all HCMs). All components of the HCMs are estimated simultaneously. In Table 5 we provide the results for the two evaluators identified in Section 2.3 and the results for the model using the median ratings of the adjectives. These three sets of results were selected for presentation since they give an idea on the impact of the ratings of adjectives on the estimation results taking into account the extreme and median ratings. A first observation is that, although the estimated parameters are close for different evaluators, the results with outlying evaluator differs in a non-negligible way from the other sets of results for some parameters. The parameters that differ the most from the other evaluators are the ones that are closely related to the ratings of adjectives, namely the parameters associated with the explanatory variables of the LV. As explained in Section 2.3 the ratings of the outlying evaluator differ in terms of both the sign and the magnitude from the other evaluators.

In addition to the estimation results of the three selected evaluators, we also display the estimation of a logit model, which is a function of the same variables as the ones contained in the structural equations of the HCMs. More precisely, the utilities of the logit model are the same as in the HCMs, except for the PT alternative, which is expressed as follows:

$$\begin{aligned}
V_{PT} = & ASC_{PT} + \beta_{cost} \cdot cost_{PT} + \beta_{time_{PT}} \cdot time_{PT} + \beta_{work_{PT}} \cdot work \\
& + \beta_{French_{PT}} \cdot French + \lambda_{French} \cdot French \cdot time_{PT} \cdot \frac{1}{1000} + \lambda_{age_{50}} \cdot age_{50} \cdot time_{PT} \cdot \frac{1}{1000} \\
& + \lambda_{active} \cdot active \cdot time_{PT} \cdot \frac{1}{1000} + \lambda_{cars} \cdot cars \cdot time_{PT} \cdot \frac{1}{1000}
\end{aligned} \tag{7}$$

The following conclusions can be drawn from the estimates of the choice model component of the HCMs:

- The negative signs of the parameters β_{cost} , $\beta_{distance}$, $\beta_{time_{PMM}}$ and $\beta_{time_{PT}}$ are consistent with expectations. Since the travel time of PT is interacted with the LV, the time coefficient differs from one evaluator to the other. It also differs from the logit model.
- Some socio-economic variables have a significant effect on the choice of transport modes. First, performing tours that only include a trip from home to work and back from work to home decreases the probability of choosing PMM relative to the base alternative of SM. This can be seen in the negative sign of the coefficient $\beta_{work_{PMM}}$. Second, the positive sign of $\beta_{French_{PMM}}$ shows that being from a French-speaking region increases the probability of choosing PMM relative to SM. For PT we cannot draw clear conclusions since both coefficients have low t -statistics. They were nevertheless kept in the model to distinguish the effect of the trip purpose and the language region on choice.
- The positive sign of coefficient $\beta_{comfort}$ shows that a good perception of comfort of PT increases its utility. Let us recall that variable $PCPT$ was interacted with variable $time_{PT}$. When we look at the net effect of travel time in the utility of PT taking into account the interaction, we have:

$$\beta_{time_{PT}} + \frac{\beta_{comfort}}{1000} \cdot PCPT. \tag{8}$$

Since $\beta_{time_{PT}}$ is negative and $\beta_{comfort}$ is positive, travel time sensitivity decreases with a better perception of comfort of PT. To make the time coefficient positive, the value of the LV should be 13.5 for the central evaluator. This value is very high considering the mean λ_{mean} , that is estimated as 3.33. A similar analysis can be done for the other evaluators. It can then be concluded that the net effect of travel time is hardly ever positive.

From the analysis of the estimated parameters of the structural equation of the LV $PCPT$ we see that the coefficients differ more across the evaluators compared to the parameters of the choice model.

- The negative sign of λ_{active} indicates that individuals with full- or part-time jobs have a more negative perception of comfort of PT compared to those without jobs.
- Age is also a factor affecting the perception of comfort of PT, i.e. people below 50 years have a more negative image of it. We observe that the coefficient for age is not significant for the outlying evaluator.
- Living in a French-speaking region is negatively affecting the perception of comfort of PT.
- Finally, respondents with at least two cars in the household have a more negative image of comfort of PT compared to others.

The estimated parameters regarding the measurement equations of the indicators directly depend on the ratings given by the evaluators. Therefore the estimates of α_k , γ_k and σ_k differ considerably with different set of ratings.

The estimation results of the logit model show that parameters λ_{French} and $\lambda_{\text{age}_{50}}$ are not significant when they are directly included in the utility functions of the choice model. This can be due to the fact that some factors are correlated, namely *French* with *French · time_{PT}*, and *age₅₀* with *active · time_{PT}*. The correlations are respectively 0.75 and 0.61. This however indicates that the introduction of the perception of comfort of PT improves the explanatory power of the choice model. Socio-economic factors *French* and *age₅₀* provide a better understanding of the model when integrated through an LVM rather than through a direct effect on the utility.

Name	Logit		Central evaluator		Outlying evaluator		Median ratings	
	Value	<i>t</i> -test	Value	<i>t</i> -test	Value	<i>t</i> -test	Value	<i>t</i> -test
ASC _{PT}	-0.178	-0.88 *	-0.155	-0.77 *	-0.132	-0.66 *	-0.155	-0.77 *
ASC _{PMM}	0.423	2.30	0.416	2.26	0.410	2.22	0.419	2.27
β_{cost}	-0.0658	-8.67	-0.0637	-8.11	-0.0628	-8.02	-0.0653	-8.08
β_{timePT}	-0.00600	-3.34	-0.0143	-7.71	-0.0293	-4.17	-0.0208	-7.06
β_{timePMM}	-0.0330	-10.27	-0.0313	-9.53	-0.0312	-9.55	-0.0323	-9.43
β_{distance}	-0.236	-11.51	-0.233	-11.4	-0.233	-11.38	-0.235	-11.45
β_{workPT}	0.0987	0.42 *	-0.0602	-0.26 *	-0.0928	-0.40 *	-0.0474	-0.20 *
β_{workPMM}	-0.613	-2.77	-0.572	-2.58	-0.560	-2.53	-0.575	-2.60
β_{FrenchPT}	-0.228	-0.61 *	-0.073	-0.24 *	-0.113	-0.37 *	-0.0808	-0.26 *
$\beta_{\text{FrenchPMM}}$	0.990	3.64	0.966	3.56	0.969	3.57	0.967	3.56
β_{comfort}	-	-	1.06	3.46	1.09	2.65	1.33	4.34
λ_{mean}	-	-	3.33	9.40	15.7	11.46	7.47	9.98
λ_{French}	1.11	0.44 *	-0.559	-1.80	-0.139	-0.48 *	-0.456	-1.58 *
λ_{age50}	1.42	1.25 *	-1.30	-5.53	0.0643	0.30 *	-1.04	-4.62
λ_{active}	-8.34	-6.77	-1.10	-4.37	-0.582	-2.68	-1.12	-4.62
λ_{cars}	-7.81	-6.59	-0.730	-3.06	-0.362	-1.58 *	-0.688	-3.04
α_2	-	-	-0.243	-2.67	0.505	1.61 *	0.00209	0.01 *
α_3	-	-	-0.473	-3.92	1.46	2.91	-0.686	-1.98
α_4	-	-	-1.14	-9.33	-3.33	-7.43	-2.34	-6.55
α_5	-	-	-1.23	-13.33	-0.344	-0.90 *	-2.31	-6.75
α_6	-	-	-1.22	-11.7	0.742	1.68	-2.98	-7.49
α_7	-	-	-0.391	-3.63	-2.30	-5.51	-0.828	-3.03
α_8	-	-	-0.560	-5.03	1.37	5.53	-1.35	-3.97
α_9	-	-	-0.842	-5.83	0.731	2.49	-2.56	-5.90
γ_1	-	-	0.255	10.55	0.116	11.59	0.230	9.00
γ_2	-	-	0.158	4.59	0.0742	3.85	0.151	3.94
γ_3	-	-	0.108	2.23	0.00423	0.13 *	0.203	4.18
γ_4	-	-	0.422	14.89	0.319	30.33	0.433	13.06
γ_5	-	-	0.195	5.69	0.123	5.40	0.321	7.49
γ_6	-	-	0.218	5.67	0.051	1.77	0.453	11.34
γ_7	-	-	0.344	12.71	0.258	16.29	0.324	10.36
γ_8	-	-	0.264	7.88	0.0268	1.69	0.336	8.62
γ_9	-	-	0.237	4.77	0.0613	3.36	0.448	10.38
σ_1	-	-	-0.0585	-1.17 *	-1.06	-25.2	0.0627	1.46 *
σ_2	-	-	0.255	6.09	-0.521	-12.34	0.388	9.85
σ_3	-	-	0.365	7.77	-0.365	-8.29	0.381	7.37
σ_4	-	-	-0.344	-2.81	-1.68	-9.25	0.00596	0.07 *
σ_5	-	-	0.0509	0.92 *	-0.461	-7.56	0.155	2.23
σ_6	-	-	0.00644	0.09 *	-0.394	-7.64	-0.279	-1.74 *
σ_7	-	-	-0.250	-2.92	-1.23	-15.52	0.0178	0.27 *
σ_8	-	-	0.122	1.99	-0.673	-15.2	0.270	4.39
σ_9	-	-	0.180	2.34	-0.675	-11.56	0.0496	0.48 *

(* Statistical significance < 90%)

Table 5: Estimation results for the logit model and the HCMs using the ratings of the central and outlying evaluators and the median ratings (sample size = 2'265).

5 Validation

From the estimation results presented in Section 4, we can conclude that the perception of comfort of PT has a significant impact on mode choice preferences. From Tables 10, 11 and 12 of the Appendix, we see that the associated parameter is indeed significant for all but one of the 25 evaluators (evaluator 25).

As a subsequent step we would like to assess the validation power of each model. In order to obtain indicators of validity of a statistical model, we compute the values of $\bar{\rho}^2$ and final loglikelihood of the choice model. The value of $\bar{\rho}^2$ is computed as follows.

$$\bar{\rho}^2 = 1 - \frac{\mathcal{L}(\hat{\beta}) - J}{\mathcal{L}(0)}, \quad (9)$$

where $\mathcal{L}(\hat{\beta})$ is the final loglikelihood of the choice model component and J is the number of parameters. The null loglikelihood $\mathcal{L}(0)$ is the likelihood of the models where all parameters are set to 0.

The numerical values of loglikelihood and $\bar{\rho}^2$ are reported in Table 6 for the logit model and the DCM components of the three HCMs. We also show the number of parameters J in each choice model component.

The loglikelihood and the values of $\bar{\rho}^2$ of the logit model are higher than those of the DCM components of the HCMs, implying a slightly better fit.

The fit indices of the model using the ratings of the outlying evaluator are slightly lower than the fit indices of the models using the ratings from the central evaluator and the median ratings. However the differences are small. A way to investigate how poor ratings affect the fit of the measurement model, compared to consistent ratings, is calculate the standardized residuals of the measurement equations. As an example, diagnostic plots of the standardized residuals for I_4 are shown for the outlying and central evaluators in Figure 4 of the Appendix. For the outlying evaluator, the standardized residuals for I_4 show a clear departure from normality.

Indicator	Logit	Central evaluator	Outlying evaluator	Median ratings
Loglikelihood	-1153	-1192	-1199	-1190
J	14	11	11	11
$\bar{\rho}^2$	0.443	0.425	0.422	0.427

Table 6: Values of final loglikelihood and $\bar{\rho}^2$ for the logit model and the DCM components of the HCMs based on the ratings of the central and outlying evaluators and on the median ratings (sample size = 2'265).

A thorough validation of the model would require its application on a different data set. As

no other similar data set is available, we first estimate the models described in Section 3 on 80% of the data and in a second phase, we apply them on the remaining 20%. The purpose of this procedure is to see how the fit indicators (final loglikelihood and value of $\bar{\rho}^2$) behave and how often the choice probabilities computed for each observation of the 20% of the data point to the actual choice. By choice probability, we mean the probability predicted by each model that the respondent chooses the transport mode he reported.

For comparison purposes, the logit model presented in Section 4 is also estimated on the same data set with 80% of the observations and validated on the data set with 20% of the observations.

Table 7 shows the loglikelihood, the number of parameters J , the value of $\bar{\rho}^2$ and the frequency of choice probabilities which are higher than 0.5 for the logit model and the three DCM components of the HCMs. All indices show that the prediction power is similar across models and confirm that it is rather high.

Indicator	Logit	Central evaluator	Outlying evaluator	Median ratings
Loglikelihood	-220	-227	-229	-226
J	14	11	11	11
$\bar{\rho}^2$	0.459	0.449	0.444	0.452
Percentage choice probabilities > 0.5	70.5%	69.7%	70.0%	70.5%

Table 7: Values of final loglikelihood, $\bar{\rho}^2$ and proportion of choice probabilities higher than 0.5 for the logit model and the DCM components of the HCMs based on the ratings of the central and outlying evaluators and on the median ratings (sample size = 453).

6 Analysis of demand indicators across evaluators

The estimation results of the mode choice models presented in Section 4 enable us to perform an analysis of demand for PT and PMM. In this section, we will precisely analyze how these indicators vary across models with ratings from different evaluators.

We compute three sorts of aggregate indicators: market shares, values of time (VOT) and elasticities.

6.1 Definition of the indicators

The indicators are computed as follows:

- The market share of an alternative i is given by the weighted sum of the individual probabilities of choosing that particular alternative:

$$MS(i) = \sum_{n=1}^N w_n P_n(i), \quad (10)$$

where w_n is the sample weight introduced in Section 2.1.

- The VOT is an indicator of willingness-to-pay (WTP), measuring the amount of money individuals are ready to spend if their trip can be reduced by one hour. For an individual n , the disaggregate VOT for a decrease of 1 hour in PMM is computed as follows:

$$VOT_{\text{PMM},n} = \frac{\beta_{\text{timePMM}}}{\beta_{\text{cost}} \cdot 60} \quad (11)$$

The disaggregate VOT for a decrease of 1 hour of travel in PT is given by the following formula:

$$VOT_{\text{PT},n} = \frac{\beta_{\text{timePT}} + \beta_{\text{comfort}} \cdot \overline{\text{PCPT}}_n / 1000}{\beta_{\text{cost}} \cdot 60}, \quad (12)$$

where $\overline{\text{PCPT}}_n$ is obtained by considering the mean of PCPT of Equation (4). The aggregate VOTs for PMM and PT are obtained by computing the weighted sum of the above disaggregate indicators. Respectively, we have:

$$VOT_{\text{PMM}} = \sum_{n=1}^N w_n VOT_{\text{PMM},n} \quad (13)$$

$$VOT_{\text{PT}} = \sum_{n=1}^N w_n VOT_{\text{PT},n} \quad (14)$$

- We report two types of aggregate elasticities, that is, *direct elasticities* and *cross elasticities*. We are interested in analyzing the percent change in the market shares of the alternatives of PMM, PT and SM, with respect to changes in variables cost_{PMM} , cost_{PT} , time_{PMM} and time_{PT} .

Aggregate direct elasticities are computed using the following formula, where x_i is any of the above variables:

$$E_{x_i}^i = \frac{\sum_{n=1}^N w_n P_n(i) E_{x_{in}}^i}{\sum_{n=1}^N w_n P_n(i)}. \quad (15)$$

In the above equation, w_n is the sample weight of observation n , $P_n(i)$ is the probability that the individual who performed tour n chooses alternative i and $E_{x_{in}}^i$ is the disaggregate direct elasticity of the demand for observation n for variations in variable x_{in} . This disaggregate elasticity is computed using the following expression:

$$E_{x_{in}}^i = \frac{\partial P_n(i)}{\partial x_{in}} \frac{x_{in}}{P_n(i)}. \quad (16)$$

To summarize, an aggregate direct elasticity denotes the percent change in the market share for alternative i with respect to a change of 1% in the value of an attribute x_i of i .

Aggregate cross elasticities are given by the following expression:

$$E_{x_j}^i = \frac{\sum_{n=1}^N w_n P_n(i) E_{x_{jn}}^i}{\sum_{n=1}^N w_n P_n(i)}, \quad (17)$$

where $E_{x_{jn}}^i$ is the cross elasticity of the demand for observation n for variations in variable x_{jn} . This disaggregate cross elasticity is computed using the following expression:

$$E_{x_{jn}}^i = \frac{\partial P_n(i)}{\partial x_{jn}} \frac{x_{jn}}{P_n(i)}. \quad (18)$$

An aggregate cross elasticity hence represents the percent change in the market share for alternative i with respect to a change of 1% in the value of an attribute x_j of another alternative j .

For the calculation of the elasticities, the partial derivatives are obtained by the built-in function of the software Biogeme. Note that the elasticity formula is not straightforward to derive, as $time_{PT}$ appears in two terms of the utility V_{PT} .

6.2 Aggregate indicators across evaluators

To obtain a quantitative assessment of the demand for the transport modes, it is also interesting to look how indicators of demand differ across evaluators. For this purpose, the market shares, aggregate VOTs and elasticities for the models with different ratings of the evaluators are computed.

Market shares, VOTs and elasticities are presented in Table 13 in the Appendix. The results show that aggregate indicators of demand do not show important changes with different ratings of the adjectives. Small differences can be observed between the indicators computed for the logit model and the ones computed for the HCMs.

As expected the market share of PMM is rather high compared to PT. The use of SM is rather low. The VOTs show that individuals are willing to pay about 30 CHF to gain one hour in a trip in PMM, whereas they are only ready to spend about 12 CHF to gain the same time in PT. We moreover note that these values are comparable with the ones reported in a study on Swiss values of travel time savings (Axhausen et al., 2008). We also conclude that the demand for the three types of transportation modes is rather inelastic with respect to changes in the travel fare or duration. The absolute value of the aggregate elasticities is indeed lower than 1 (Arnold, 2008). Nevertheless, the demand is more elastic with respect to changes in travel time than with respect to changes in travel cost. By considering the results of the central evaluator as an example, we can observe that an increase of 1% in the travel time in PT leads to a decrease of 0.62% of its market share and that the same percent increase in the time in PMM leads to an increase of 0.61% in the market share of PT. Similar conclusions can be drawn for the other evaluators.

6.3 Disaggregate indicators across evaluators

After investigating the market shares, VOTs and elasticities of cost or time for each evaluator, we are interested in analyzing the distribution of such indicators over all evaluators. In particular, we want to analyze the distribution of (i) disaggregate demand indicators for three respondents of the initial RP survey and (ii) aggregate demand indicators computed over the whole sample.

To analyze the distributions of the disaggregate demand indicators over all evaluators, we select three examples of respondents: a respondent whose probability to select the alternative he actually chose (PT) is poorly predicted by the model, a respondent with an average predicted probability ($\cong 0.5$) for the chosen alternative (PT) and a respondent whose predicted probability to select the chosen alternative (PMM) is high. Means and standard deviations (SD) for these disaggregate demand indicators are computed over the sample of evaluators and reported in columns ‘Low’, ‘Medium’ and ‘High’ of Table 8. Aggregate demand indicators are also obtained from the application of each HCM on the sample of all 2’265 respondents of the RP survey. Their means and SDs over all evaluators are reported in column ‘All observations’ of Table 8.

From Table 8 we observe a general trend: the distribution of disaggregate indicators across evaluators has a larger variance than the distribution of aggregate indicators. This holds for every indicator, except VOT.

For the three chosen respondents, the probability of choosing an alternative has a larger standard deviation than the market share of the same alternative. Figure 1 provides a visual example of this phenomenon for the probability of choosing the PT alternative. For each of the four graphs, we considered the same scale for the horizontal axes, in order to make the comparison between the four distributions easier². We mean that the difference between the extreme points of the axis is the same across graphs.

Similarly, Figure 2 shows that the disaggregate elasticities of the cost of PMM are more spread than the aggregate cost elasticities.

The aggregate VOT for PT has a standard deviation which is smaller than the disaggregate VOTs for PT for two out of the three examples of respondents (see Table 8 and Figure 3). In addition, we have a different VOT for PT for each individual due to the interaction with the perception of comfort of PT, while the VOT for PMM is constant across respondents.

6.4 Impact of an increase of the comfort level

We present here an example demonstrating an important property of HCMs, that is, we can measure in a quantitative way the impact of a change in the perceptual variable on the choice. In

²The same remark holds for Figures 2 and 3.

		Examples of respondents with different probabilities for the chosen alternative						All observations	
Indicator	Mode	Low		Medium		High		Mean	SD
		Mean	SD	Mean	SD	Mean	SD		
Probability of choice / Market share	PT	0.047	0.008	0.543	0.013	0.067	0.005	0.278	0.001
	PMM	0.953	0.008	0.457	0.013	0.933	0.005	0.659	0.001
	SM	0.000	0.000	0.000	0.000	0.000	0.000	0.063	0.000
VOT	PT	10.417	0.603	11.617	0.320	13.584	0.538	12.311	0.323
	PMM	29.635	0.108	29.635	0.108	29.635	0.108	29.635	0.108
Elasticity of cost	PT	-2.681	0.036	-0.247	0.008	-1.493	0.022	-0.282	0.001
	PMM	-0.026	0.005	-0.246	0.006	-0.036	0.002	-0.073	0.000
Elasticity of time	PT	-2.209	0.379	-0.914	0.046	-2.122	0.039	-0.626	0.005
	PMM	-0.120	0.021	-1.239	0.027	-0.097	0.007	-0.276	0.001
Cross-elasticity of cost of PMM	SM	0.562	0.005	0.241	0.005	0.512	0.011	0.040	0.001
	PT	0.531	0.007	0.207	0.007	0.499	0.007	0.163	0.000
Cross-elasticity of cost of PT	PMM	0.131	0.023	0.293	0.007	0.107	0.007	0.115	0.001
	SM	0.131	0.023	0.293	0.007	0.107	0.007	0.048	0.001
Cross-elasticity of time in PMM	SM	2.591	0.024	1.216	0.027	1.402	0.029	0.188	0.003
	PT	2.446	0.032	1.044	0.034	1.365	0.020	0.610	0.002
Cross-elasticity of time in PT	PMM	0.105	0.008	1.084	0.008	0.151	0.011	0.254	0.002
	SM	0.105	0.008	1.084	0.008	0.151	0.011	0.110	0.001

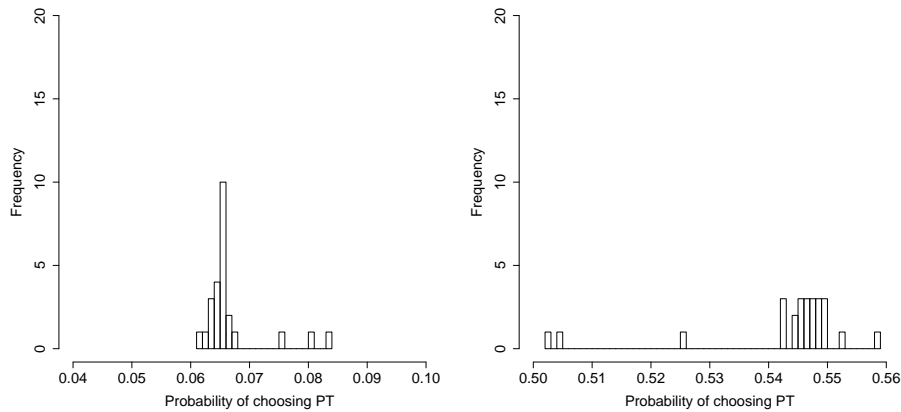
Table 8: Probabilities of choice, disaggregate VOTs, direct and cross-elasticities for three respondents of the RP survey (Mean and SD over all 25 evaluators); market shares, aggregate VOTs, direct and cross-elasticities resulting from the application of each HCM on the sample of all 2'265 respondents of the RP survey (Mean and SD over all 25 evaluators).

the present case study, we can indeed quantify the effect of an increase in the comfort level on the mode choice. By construction, changes in the variables characterizing a certain level of comfort of PT lead to an increase or decrease of the perception of comfort of PT, which results in variations in the market shares.

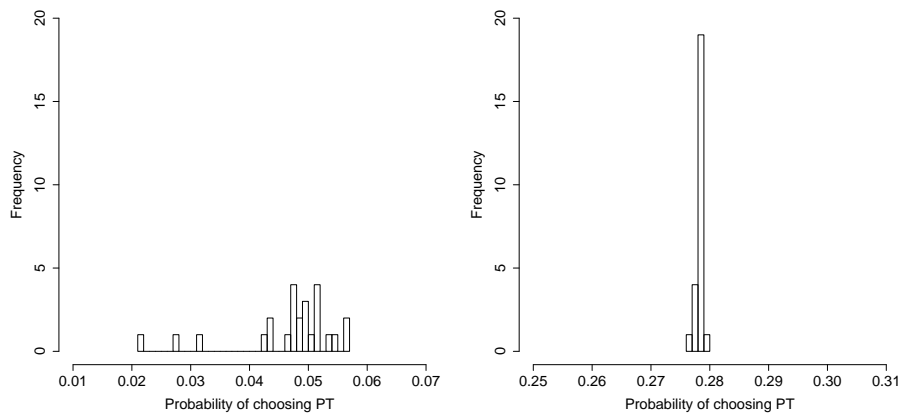
The scenario presented in this section is based on the estimation sample, where the number of cars in each household has been decreased by 1. Table 9 shows that the market share of PT has slightly increased as a consequence of this potential policy change. A better perception of comfort of PT can hence drive individuals to make an increased use of such transportation modes.

7 Some recommendations for questionnaire design

The outcomes of this research have shown that more qualitative data such as adjectives can be used to measure a perception. However the adjective data set we are using was not initially designed



(a) Respondent with a high probability for the chosen alternative (PMM) (b) Respondent with an average probability for the chosen alternative (PT)



(c) Respondent with a low probability for the chosen alternative (PT) (d) Aggregation over all respondents

Figure 1: Histograms of the probabilities of choosing PT for three respondents and histogram of the market share of PT (sample size = 25).

to be used for quantitative research. The whole data collection procedure could therefore benefit from some improvements in order to be faster and comparable with existing approaches. In this section, we report some recommendations for future survey design.

The two surveys in one step. The RP data collection and the collection of ratings of the adjectives could be performed in a single step, where the respondents of the RP survey would be asked to evaluate the adjectives. The detailed survey methodology would consist of (1) asking them to report freely adjectives for the different transportation modes, (2) asking them to which theme they associate which reported adjectives and (3) rate each adjective on the scale of the given theme.

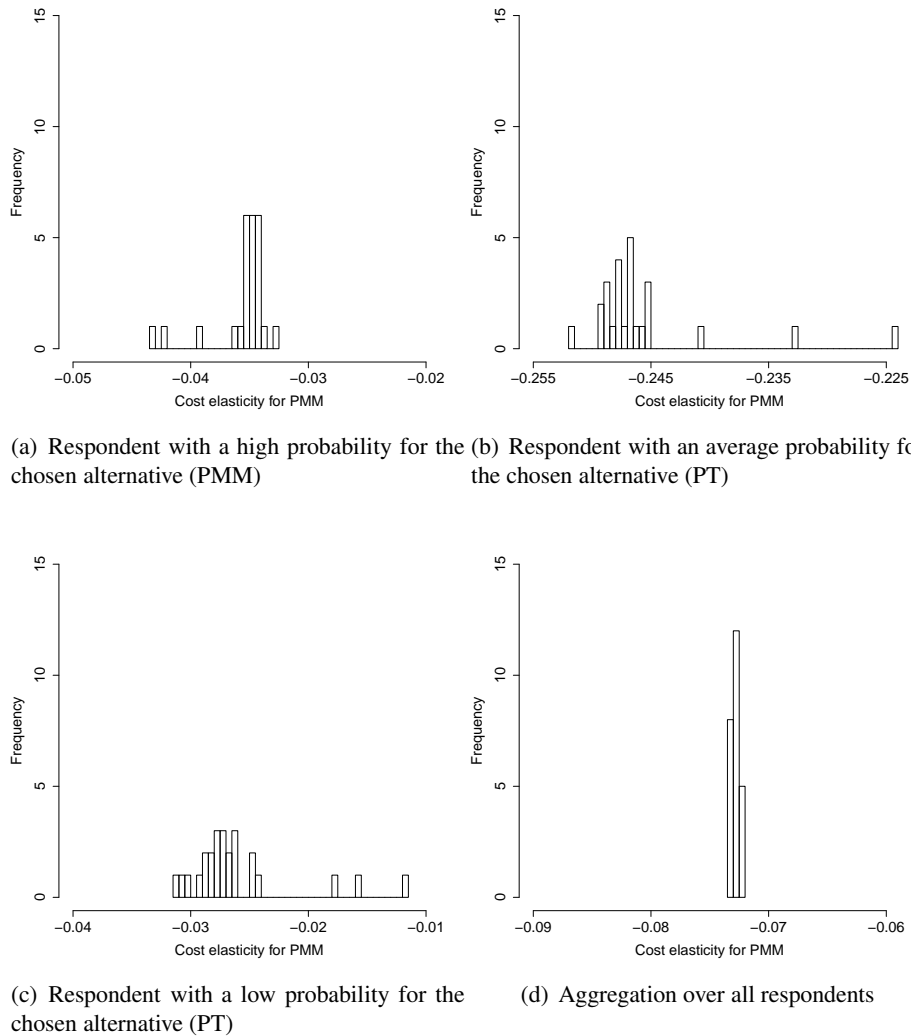
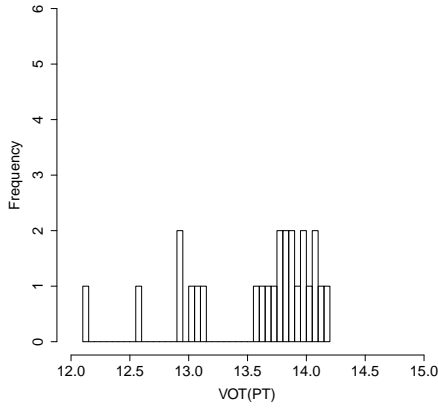


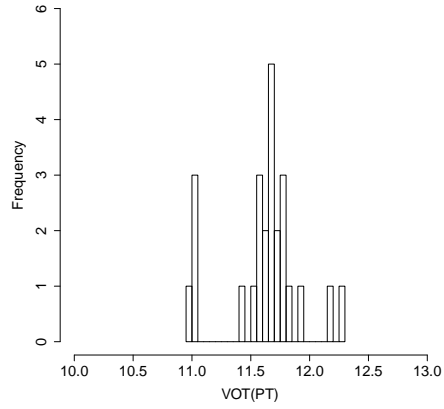
Figure 2: Histograms of the disaggregate cost elasticities for PMM for three respondents and histogram of the aggregate cost elasticities for PMM (sample size = 25).

A comparative approach with the classical method. A standard way to measure perceptions is to ask respondents to rate their agreement to opinion statements on a five-point Likert scale. As highlighted earlier, such a method depends on the conception of the perception by the survey designer. However, as raised by a reviewer, the methodology presented in this paper also involves some subjectivity, in the sense that the adjectives are associated to a theme by an analyst.

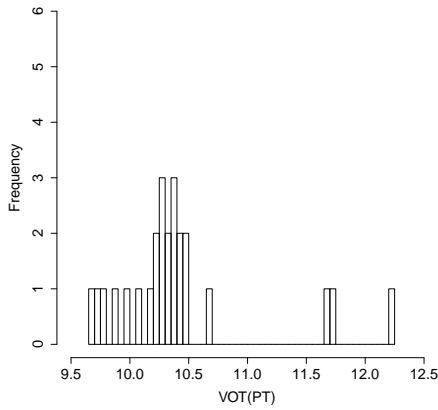
Since multiple evaluators are involved, the bias linked to the adjective rating is reduced, but it could be argued that in the case of the design of opinion statements, some bias is also reduced since several individuals are usually involved in the survey design.



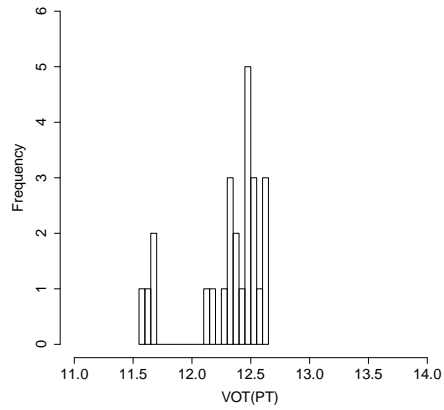
(a) Respondent with a high probability for the chosen alternative (PMM)



(b) Respondent with an average probability for the chosen alternative (PT)



(c) Respondent with a low probability for the chosen alternative (PT)



(d) Aggregation over all respondents

Figure 3: Histograms of the VOTs for three respondents for PT and histogram of the aggregate VOT for PT (sample size = 25).

Therefore, an interesting research path to investigate would be to separate the RP respondents into two groups and ask the first one to answer opinion statements related to a number of themes of interest (e.g. including comfort) and the second one to report adjectives related to the same themes.

8 Conclusion and further works

This research presents a methodology to measure and integrate perceptual information into a DCM. The use of semi-open questions provides spontaneous information about respondents' perceptions, and the estimation and application of an HCM have shown that the latter affect the choice in a significant way.

Indicator	Mode	Base case		Less cars per households	
		Mean	SD	Mean	SD
Market share	PT	0.278	0.001	0.283	0.002
	PMM	0.659	0.001	0.654	0.002
	SM	0.063	0.000	0.062	0.000

Table 9: Market shares for the whole sample of the RP survey under a scenario where the number of cars per household is decreased (sample size = 2'265).

This paper contributes to several achievements. First, to the best of our knowledge, this is the first attempt to integrate qualitative information, such as semi-open questions, into DCMs. We achieve a meaningful characterization of the perception of comfort of public transportation by using the adjectives as measurements of it. Second, we provide a methodology to rate the adjectives on a scale of comfort which minimizes the subjectivity resulting when using a single modeler's ratings. Third, we highlight the impact of ratings from different evaluators on the demand by exploring several indicators, such as market shares, values of time and elasticities.

By considering responses to semi-open questions, we aim at proposing an alternative way to measure perceptions. The obtained adjectives contain rich information about the individuals' perceptions since they are freely reported. This way we wish to overcome the bias inherent to the collection of responses to opinion statements, which are reflecting the survey designer's representation of a perception. The methodology moreover highlights the importance of considering mixed qualitative/quantitative approach to conduct revealed or stated preferences surveys.

Future works involve the integration of the ratings of all evaluators into the HCM framework, in order to benefit from the global judgement to value adjectives. In addition, we will also include ratings from a finer-grained scale, ranging from -1000 to 1000 , which were reported by additional evaluators. Since differences across evaluators were observed in disaggregate indicators, a finer characterization of perceptions will be investigated as a future research. For this purpose we plan to integrate the socio-economic information of the evaluators as explanatory variables of the reported ratings.

The effect of the perception of comfort of other transport modes might also affect the choices of individuals. Moreover its interaction with other variables could be tested, since the perception of comfort might differ among population segments. The list of reported adjectives revealed eleven themes (see Section 2.2) which might also be important factors of transportation mode choices. In future research, all these aspects should be investigated, in order to have a broader view of all perceptual variables affecting transportation mode decisions. However, it is important to remark

that in practice, rating adjectives relative to eleven themes can be burdensome for evaluators. Such an exercise should be designed to minimize the impact of the fatigue on the quality of the evaluation. Clearly, using many evaluators would ease the burden for each of them. In addition, it is worth noting that the presented modeling methodology shares the same limitations with HCMs. First, when the number of indicators for the latent construct is high, the maximum likelihood estimation becomes cumbersome, and other estimation techniques (e.g. Bayesian) should be considered. Second, more work needs to be done to validate how well the structural equation is able to characterize an LV. In particular, a good measure of the fit would be useful to support this analysis.

Regarding survey design, further work is needed to assess whether all the measures of a psychological construct which is an important explanatory factor of a respondent's behavior are actually recalled. The understanding of this cognitive process will allow to highlight better the strengths and weaknesses of semi-open questions versus opinion statements as measurements of a perception.

In this paper, we have presented a method that exploits a new type of data, and illustrated it on a real case study. Note that the additional complexity can be a curse or a blessing depending on the application, the data and the context in general. For example, the use of semi-open questions for similar concepts (like 'comfort' and 'convenience') may complicate the task of the evaluators, as the same adjectives may be reported for both concepts. The associated modeling issues are yet to be investigated.

9 Acknowledgements

The authors would like to thank PostBus for their advice and feedback on the collaborative project which was set up and for funding it. They also would like to acknowledge the Urban Sociology Laboratory (LaSUR) and the Urban and Regional Planning Laboratory (CEAT) who co-conducted the qualitative and the RP surveys which serve as a basis to this research. They are also grateful to Alberto Monticone, who established the classification of the adjectives into themes. This paper has also benefited from the comments from three anonymous reviewers as well as two anonymous reviewers of an earlier version (Glerum et al., 2011) which was part of the conference proceedings of the ICMC conference in Leeds in 2011.

References

Abou-Zeid, M., Ben-Akiva, M., Bierlaire, M., Choudhury, C. and Hess, S. (2010). Attitudes and value of time heterogeneity, in E. V. de Voorde and T. Vanelslander (eds), *Applied Transport*

Economics A Management and Policy Perspective, de boeck, pp. 523–545.

Abou-Zeid, M., Witter, R., Bierlaire, M., Kaufmann, V. and Ben-Akiva, M. (2012). Happiness and travel mode switching: Findings from a swiss public transportation experiment, *Transport Policy* **19**(1): 93–104.

Amazon's Mechanical Turk (2005-2012).

URL: <https://www.mturk.com/>

Arnold, R. A. (2008). *Economics*, ninth edn, Cengage Learning, Mason.

Atasoy, B., Glerum, A. and Bierlaire, M. (2013). Attitudes towards mode choice in Switzerland, *disP - The Planning Review* **49**(2): 101–117.

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.

Banerjee, M., Capozzoli, M., McSweeney, L. and Sinha, D. (1999). Beyond kappa: A review of interrater agreement measures, *The Canadian Journal of Statistics* **27**(1): 3–23.

Bearden, W. and Netemeyer, R. (1999). *Handbook of marketing scales: multi-item measures for marketing and consumer behavior research*, Association for Consumer Research, Sage Publications.

Beirao, G. and Cabral, J. S. (2007). Understanding attitudes towards public transport and private car: A qualitative study, *Transport Policy* **14**(6): 478–489.

Ben-Akiva, M. and Lerman, S. (1985). *Discrete choice analysis: theory and application to travel demand*, Vol. 9, The MIT Press.

Ben-Akiva, M., McFadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D., Daly, A., de Palma, A., Gopinath, D., Karlstrom, A. and Munizaga, M. A. (2002). Hybrid choice models: Progress and challenges, *Marketing Letters* **13**(3): 163–175.

Bierlaire, M., Curchod, A., Danalet, A., Doyen, E., Faure, P., Glerum, A., Kaufmann, V., Tabaka, K. and Schuler, M. (2011). Projet de recherche sur la mobilité combinée, rapport définitif de l'enquête de préférences révélées, *Technical Report TRANSP-OR 110704*, Transport and Mobility Laboratory, Ecole Polytechnique Fédérale de Lausanne.

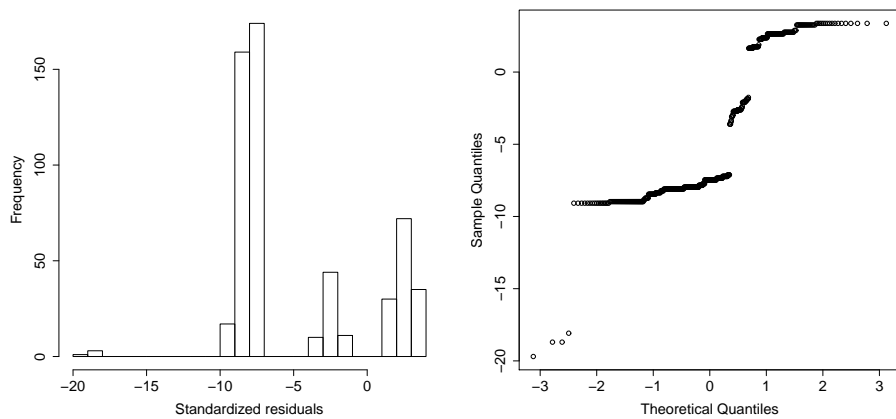
- Bierlaire, M. and Feterison, M. (2009). Estimation of discrete choice models: extending BIOGEME, *Swiss Transport Research Conference (STRC)*.
- Bolduc, D. and Daziano, R. A. (2011). Incorporating pro-environmental preferences toward green automobile technologies through a Bayesian Hybrid Choice Model, *Transportmetrica* pp. 1–33.
- Brownstone, D. (1998). Multiple imputation methodology for missing data, non-random response, and panel attrition, *Theoretical foundations of travel choice modelling*, Elsevier, New York, Amsterdam, pp. 421–450.
- Cramer, P. (1968). *Word association*, Academic Press.
- Daly, A., Hess, S., Patrini, B., Potoglou, D. and Rohr, C. (2012). Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour, *Transportation* **39**: 267–297.
- Espino, R., Román, C. and Ortúzar, J. d. D. (2006). Analyzing demand for suburban trips: A mixed RP/SP model with latent variables and interaction effects, *Transportation* **33**(3): 241–261.
- Franklin, M. J., Kossmann, D., Kraska, T., Ramesh, S. and Xin, R. (2011). Crowddb: answering queries with crowdsourcing, *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, SIGMOD '11, ACM, New York, NY, USA, pp. 61–72.
- Glerum, A., Atasoy, B., Monticone, A. and Bierlaire, M. (2011). Adjectives qualifying individuals' perceptions impacting on transport mode preferences, *Proceedings of the Second International Choice Modeling Conference*, Leeds, UK.
- Greene, W. H. and Hensher, D. A. (2003). A latent class model for discrete choice analysis: contrasts with mixed logit, *Transportation Research Part B: Methodological* **37**(8): 681–698.
- Hess, S. and Train, K. (2011). Recovery of inter- and intra-personal heterogeneity using mixed logit models, *Transportation Research Part B: Methodological* **45**(7): 973–990.
- Kaufmann, V., Jemelin, C. and Guidez, J.-M. (2001). *Automobile et modes de vie urbains quel degré de liberté?*, Documentation Française.
- Kaufmann, V., Tabaka, K., Louvet, N. and Guidez, J.-M. (2010). *Et si les Français n'avaient plus seulement une voiture dans la tête?*, Certu.
- Likert, R. (1932). A technique for the measurement of attitudes, *Archives of Psychology* **22**(140).

- Looker, E., Denton, M. A. and Davis, C. K. (1989). Bridging the gap: Incorporating qualitative data into quantitative analyses, *Social Science Research* **18**(4): 313–330.
- McFadden, D. and Train, K. (2000). Mixed MNL models for discrete response, *Journal of Applied Econometrics* **15**(5): 447–470.
- Mossholder, K. W., Settoon, R. P., Harris, S. G. and Armenakis, A. A. (1995). Measuring emotion in open-ended survey responses: An application of textual data analysis, *Journal of Management* **21**(2): 335–355.
- Potkay, C. R. and Allen, B. P. (1973). The adjective generation technique: an alternative to adjective check lists, *Psychological Reports* **32**: 457–458.
- Robin, T., Bierlaire, M. and Cruz, J. (2011). Dynamic facial expression recognition with a discrete choice model, *Journal of Choice Modelling* **4**(2): 95–148.
- Schuman, H. and Presser, S. (1996). *Questions and answers in attitude surveys: experiments on question form, wording, and context*, Quantitative Studies in Social Relations, Sage Publications.
- Snow, R., O'Connor, B., Jurafsky, D. and Ng, A. Y. (2008). Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks, *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '08, Association for Computational Linguistics, Stroudsburg, PA, USA, pp. 254–263.
- Sorci, M., Antonini, G., Cruz, J., Robin, T., Bierlaire, M. and Thiran, J.-P. (2010). Modelling human perception of static facial expressions, *Image and Vision Computing* **28**(5): 790–806.
- Thorndike, E. L. (1920). A constant error in psychological ratings, *Journal of Applied Psychology* **4**: 469–477.
- Venetis, P., Garcia-Molina, H., Huang, K. and Polyzotis, N. (2012). Max algorithms in crowdsourcing environments, *Proceedings of the 21st international conference on World Wide Web*, WWW '12, ACM, New York, NY, USA, pp. 989–998.
- Vredin Johansson, M., Heldt, T. and Johansson, P. (2006). The effects of attitudes and personality traits on mode choice, *Transportation Research Part A: Policy and Practice* **40**(6): 507–525.
- Walker, J. and Ben-Akiva, M. (2002). Generalized random utility model, *Mathematical Social Sciences* **43**(3): 303–343.

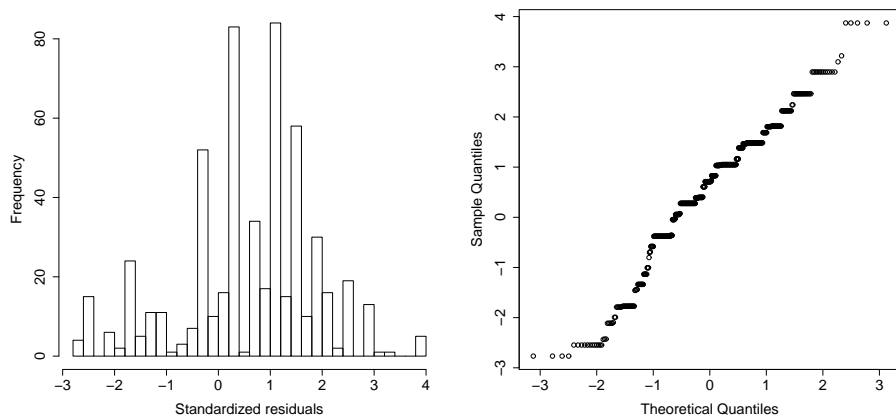
Walker, J. L. (2001). *Extended Discrete Choice Models: Integrated Framework, Flexible Error Structures, and Latent Variables*, PhD thesis, Massachusetts Institute of Technology.

Yáñez, M., Raveau, S. and Ortúzar, J. d. D. (2010). Inclusion of latent variables in mixed logit models: Modelling and forecasting, *Transportation Research Part A: Policy and Practice* **44**(9): 744–753.

Appendix



(a) Histogram of standardized residuals (Outlying evaluator) (b) Normal Q-Q of standardized residuals (Outlying evaluator)



(c) Histogram of standardized residuals (Central evaluator) (d) Normal Q-Q of standardized residuals (Central evaluator)

Figure 4: Diagnostic plots of the standardized residuals for measurement equation relative to I_4 (sample size = 2'265).

Name	Evaluator 1		Evaluator 2		Evaluator 3		Evaluator 4		Evaluator 5		Evaluator 6		Evaluator 7 (central evaluator)		Evaluator 8		Evaluator 9 (outlying evaluator)		Evaluator 10		Evaluator 11	
	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test
ASC _{PT}	-0.152	-0.76	-0.145	-0.72	-0.159	-0.79	-0.165	-0.82	-0.151	-0.75	-0.168	-0.83	-0.155	-0.77	-0.161	-0.8	-0.132	-0.66	-0.157	-0.78	-0.148	-0.74
ASC _{PMM}	0.415	2.25	0.415	2.25	0.419	2.28	0.421	2.28	0.415	2.25	0.422	2.29	0.416	2.26	0.419	2.27	0.41	2.22	0.419	2.27	0.417	2.26
β_{cost}	-0.0644	-8.05	-0.065	-8.16	-0.0648	-8.08	-0.0651	-8.07	-0.0644	-8.06	-0.0648	-8.15	-0.0637	-8.11	-0.0649	-8.07	-0.0628	-8.02	-0.0653	-8.14	-0.0653	-8.15
β_{timePT}	-0.0156	-7.49	-0.0181	-7.11	-0.0199	-7.07	-0.0183	-7.45	-0.0156	-7.47	-0.0164	-7.79	-0.0143	-7.71	-0.0182	-7.44	-0.0293	-4.17	-0.0179	-7.46	-0.0193	-6.87
$\beta_{timePMM}$	-0.0316	-9.45	-0.0319	-9.48	-0.0321	-9.46	-0.0323	-9.45	-0.0317	-9.45	-0.032	-9.5	-0.0313	-9.53	-0.0321	-9.46	-0.0312	-9.55	-0.0321	-9.47	-0.0321	-9.46
$\beta_{distance}$	-0.234	-11.41	-0.234	-11.43	-0.234	-11.43	-0.235	-11.44	-0.234	-11.41	-0.234	-11.43	-0.233	-11.4	-0.235	-11.43	-0.233	-11.38	-0.235	-11.44	-0.235	-11.44
β_{workPT}	-0.0555	-0.24	-0.0618	-0.26	-0.0492	-0.21	-0.0399	-0.17	-0.0555	-0.24	-0.0355	-0.15	-0.0602	-0.26	-0.0476	-0.2	-0.0928	-0.4	-0.0487	-0.21	-0.0552	-0.24
$\beta_{timePMM}$	-0.573	-2.59	-0.571	-2.58	-0.575	-2.6	-0.577	-2.6	-0.573	-2.59	-0.576	-2.6	-0.572	-2.58	-0.575	-2.6	-0.56	-2.53	-0.574	-2.59	-0.573	-2.59
$\beta_{FrenchPT}$	-0.0855	-0.28	-0.133	-0.43	-0.0428	-0.14	-0.0088	-0.03	-0.0897	-0.29	-0.0067	-0.02	-0.073	-0.24	-0.035	-0.11	-0.113	-0.37	-0.0736	-0.24	-0.114	-0.37
$\beta_{FrenchPMM}$	0.968	3.56	0.972	3.58	0.963	3.55	0.961	3.54	0.968	3.57	0.96	3.53	0.966	3.56	0.963	3.55	0.969	3.57	0.966	3.56	0.97	3.57
$\beta_{comfort}$	-1.21**	-3.69	-1.21**	-3.89	1.27	4.1	1.36	4.3	-1.2**	-3.67	1.29	4.3	1.06	3.46	1.29	4.21	1.09	2.65	1.29	4.25	-1.25**	-4.05
λ_{mean}	-3.97	-8.93	-5.71	-8.59	7.14	10.64	5.7	10.64	-4.01	-8.82	4.73	11.28	3.33	9.4	5.78	10.53	15.7	11.46	5.56	9.84	-6.62	-7.79
λ_{French}	0.447	1.41	-0.0646	-0.22	-0.83	-2.83	-1.12	-3.58	0.401	1.26	-1.14	-3.64	-0.559	-1.8	-0.89	-3.02	-0.139	-0.48	-0.512	-1.76	0.143	0.47
λ_{age50}	1.04	4.34	0.984	4.31	-1.07	-4.71	-1.09	-4.59	1.06	4.41	-1.5	-6.29	-1.3	-5.53	-1.03	-4.52	0.0643	0.3	-1.15	-5.03	1.09	4.68
λ_{active}	1.18	4.59	0.953	3.89	-1.12	-4.56	-1.29	-5.08	1.16	4.51	-1.28	-5.01	-1.1	-4.37	-1.18	-4.83	-0.582	-2.68	-1.13	-4.65	1.06	4.22
λ_{cars}	0.785	3.25	0.693	3.04	-0.718	-3.11	-0.848	-3.52	0.804	3.32	-0.8	-3.32	-0.73	-3.06	-0.796	-3.48	-0.362	-1.58	-0.708	-3.12	0.691	2.96
α_2	-0.0721	-0.6	-0.0381	-0.25	0.0237	0.11	-0.127	-0.64	-0.0924	-0.75	-0.201	-1.37	-0.243	-2.67	-0.218	-1.06	0.505	1.61	-0.189	-0.86	-0.0451	-0.25
α_3	-0.47	-2.88	-0.616	-2.69	-0.43	-1.44	-0.388	-1.54	-0.475	-2.87	-0.49	-2.73	-0.473	-3.92	-0.591	-2.12	1.46	2.91	-0.806	-2.62	-0.68	-2.63
α_4	-1.29	-8.33	-1.41	-7.12	-2.1	-6.6	-2.03	-7.22	-1.32	-8.23	-1.58	-7.72	-1.14	-9.33	-2.19	-7.45	-3.33	-7.43	-2.41	-7.55	-1.42	-6.3
α_5	-1.3	-10.79	-1.29	-7.89	-2.12	-6.84	-1.97	-7.97	-1.32	-10.58	-1.42	-9.01	-1.23	-13.33	-2.09	-8.17	-0.344	-0.9	-2.29	-8.44	-1.34	-6.96
α_6	-1.34	-9.67	-1.74	-7.38	-2.63	-7.69	-1.99	-7.56	-1.37	-9.57	-1.73	-8.69	-1.22	-11.7	-2.35	-7.74	0.742	1.68	-2.74	-8.22	-2.02	-6.8
α_7	-0.534	-3.93	-0.488	-3.04	-0.75	-3.02	-0.758	-3.35	-0.55	-3.96	-0.463	-2.67	-0.391	-3.63	-0.789	-3.31	-2.3	-5.51	-0.781	-3.1	-0.542	-2.94
α_8	-0.506	-3.86	-0.801	-4.26	-1.14	-3.74	-0.839	-3.34	-0.545	-3.99	-0.648	-3.33	-0.56	-5.03	-1.04	-3.89	1.37	5.53	-1.29	-4.36	-0.984	-4.21
α_9	-0.989	-5.29	-1.46	-5.6	-2.27	-5.65	-1.46	-4.22	-1	-5.14	-1.34	-5.22	-0.842	-5.83	-2.01	-5.19	0.731	2.49	-2.21	-5.69	-1.84	-5.75
γ_1	-0.231	-8.74	-0.16	-8.09	0.239	9.53	0.3	9.47	-0.227	-8.54	0.347	10.26	0.255	10.55	0.29	9.43	0.116	11.59	0.297	8.84	-0.134	-6.99
γ_2	-0.134	-3.3	-0.112	-3.58	0.139	3.81	0.187	4.17	-0.141	-3.47	0.181	4.04	0.158	4.59	0.193	4.36	0.0742	3.85	0.201	4.11	-0.0979	-3.05
γ_3	-0.165	-3.22	-0.187	-4.8	0.164	3.54	0.146	2.58	-0.163	-3.17	0.102	1.92	0.108	2.23	0.191	3.4	0.00423	0.13	0.233	3.77	-0.169	-4.4
γ_4	-0.386	-13.55	-0.29	-11.83	0.427	12.97	0.516	13.4	-0.385	-13.2	0.528	13.62	0.422	14.89	0.526	14	0.319	30.33	0.558	13.29	-0.249	-10.75
γ_5	-0.205	-5.43	-0.195	-7.92	0.311	7.13	0.32	6.36	-0.214	-5.66	0.222	4.66	0.195	5.69	0.35	7.59	0.123	5.4	0.361	7.19	-0.181	-7.19
γ_6	-0.24	-6.17	-0.305	-9.5	0.421	10.59	0.352	7.03	-0.239	-6.02	0.357	7.11	0.218	5.67	0.421	8.31	0.051	1.77	0.495	9.48	-0.319	-10.79
γ_7	-0.334	-11.54	-0.216	-9.1	0.333	10.89	0.419	11.14	-0.332	-11.48	0.447	11.43	0.344	12.71	0.407	10.76	0.258	16.29	0.41	9.83	-0.196	-9.31
γ_8	-0.231	-6.86	-0.225	-8.6	0.319	8.11	0.335	7.34	-0.238	-6.99	0.334	6.63	0.264	7.88	0.366	8.41	0.0268	1.69	0.411	8.66	-0.227	-8.4
γ_9	-0.27	-5.55	-0.31	-10.03	0.426	9.16	0.351	5.38	-0.271	-5.42	0.373	6.15	0.237	4.77	0.453	7.42	0.0613	3.36	0.492	8.51	-0.33	-11.93
σ_1	0.00192	0.04	-0.204	-4.84	0.0325	0.72	0.206	4.25	0.00488	0.1	0.226	4.09	-0.0585	-1.17	0.214	4.67	-1.06	-25.2	0.312	7.12	-0.188	-4.81
σ_2	0.366	9.2	0.144	3.66	0.347	8.84	0.516	12.65	0.36	8.93	0.524	12.88	0.255	6.09	0.523	12.96	-0.521	-12.34	0.612	15.29	0.146	3.75
σ_3	0.38	7.45	0.127	2.31	0.366	7.43	0.565	11.9	0.384	7.56	0.551	12.04	0.365	7.77	0.556	11.3	-0.365	-8.29	0.631	12.49	0.0995	1.83
σ_4	-0.262	-2.58	-0.338	-4.04	0.00059	0.01	0.0575	0.57	-0.242	-2.4	0.0482	0.47	-0.344	-2.81	0.0388	0.38	-1.68	-9.25	0.116	1.12	-0.285	-4.19
σ_5	0.025	0.4	-0.313	-4.71	0.159	2.29	0.288	4.21	0.0105	0.16	0.399	7.75	0.0509	0.92	0.226	3.15	-0.461	-7.56	0.3	4.3	-0.33	-5.05
σ_6	-0.041	-0.52	-0.699	-3.8	-0.166	-1.29	0.235	2.74	-0.0263	-0.34	0.233	2.59	0.00644	0.09	0.0856	0.77	-0.394	-7.64	-0.0244	-0.16	-0.805	-3.84
σ_7	-0.298	-2.96	-0.288	-4.45	-0.0353	-0.5	0.0831	1.01	-0.272	-2.79	0.105	1.2	-0.25	-2.92	0.134	1.79	-1.23	-15.52	0.251	3.63	-0.313	-5.32
σ_8	0.0532	0.87	-0.148	-2.35	0.28	4.63	0.397	6.54	0.0596	0.96	0.498	8.44	0.122	1.99	0.364	5.83	-0.673	-15.2	0.451	7.22	-0.168	-2.53
σ_9	0.109	1.27	-0.416	-3.29	0.105	1.04	0.427	5.15	0.101	1.16	0.4	4.57	0.18	2.34	0.283	2.61	-0.675	-11.56	0.29	2.68	-0.532	-4.01

(** Even though the coefficient of comfort is negative, the impact of *PCPT* on the utility function for PT is still positive since the γ_k 's are negative.)

Table 10: Estimation results (1st part) for the HCMs using the ratings of all evaluators and the median ratings (sample size = 2/265).

Name	Evaluator 12		Evaluator 13		Evaluator 14		Evaluator 15		Evaluator 16		Evaluator 17		Evaluator 18		Evaluator 19		Evaluator 20		Evaluator 21		Evaluator 22	
	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test	Value	t-test
ASC _{PT}	-0.161	-0.8	-0.157	-0.78	-0.158	-0.78	-0.15	-0.75	-0.155	-0.77	-0.159	-0.79	-0.151	-0.75	-0.155	-0.77	-0.164	-0.82	-0.158	-0.78	-0.15	-0.74
ASC _{PMM}	0.421	2.28	0.418	2.27	0.418	2.27	0.418	2.27	0.418	2.27	0.419	2.28	0.417	2.27	0.419	2.27	0.421	2.29	0.42	2.28	0.418	2.27
β_{cost}	-0.0653	-8.08	-0.0648	-8.13	-0.0652	-8.07	-0.0648	-8.1	-0.0652	-8.16	-0.0648	-8.08	-0.0643	-8.1	-0.0653	-8.08	-0.0649	-8.09	-0.0658	-8.07	-0.0644	-8.13
β_{timePT}	-0.0204	-7.17	-0.0174	-7.44	-0.0187	-7.37	-0.0226	-6.43	-0.0205	-6.91	-0.02	-7.04	-0.0219	-6.39	-0.0208	-7.06	-0.0195	-7.21	-0.0215	-7.04	-0.0202	-6.79
β_{timePMM}	-0.0323	-9.43	-0.0319	-9.48	-0.0322	-9.45	-0.0322	-9.43	-0.0321	-9.48	-0.032	-9.46	-0.0319	-9.45	-0.0323	-9.43	-0.0322	-9.46	-0.0326	-9.41	-0.0318	-9.48
β_{distance}	-0.235	-11.44	-0.234	-11.42	-0.235	-11.44	-0.235	-11.43	-0.235	-11.43	-0.234	-11.43	-0.234	-11.42	-0.235	-11.45	-0.235	-11.44	-0.235	-11.46	-0.234	-11.42
β_{workPT}	-0.0427	-0.18	-0.0522	-0.22	-0.047	-0.2	-0.0454	-0.19	-0.0518	-0.22	-0.0492	-0.21	-0.049	-0.21	-0.0474	-0.2	-0.0456	-0.19	-0.0396	-0.17	-0.0586	-0.25
β_{timePMM}	-0.575	-2.6	-0.573	-2.59	-0.575	-2.6	-0.575	-2.6	-0.573	-2.59	-0.574	-2.59	-0.574	-2.59	-0.575	-2.6	-0.574	-2.59	-0.576	-2.6	-0.571	-2.58
β_{FrenchPT}	-0.0429	-0.14	-0.0713	-0.23	-0.0595	-0.19	-0.106	-0.35	-0.0671	-0.22	-0.0457	-0.15	-0.1	-0.33	-0.0808	-0.26	-0.0087	-0.03	-0.0715	-0.23	-0.102	-0.33
$\beta_{\text{FrenchPMM}}$	0.963	3.55	0.966	3.56	0.965	3.56	0.969	3.57	0.966	3.56	0.963	3.55	0.968	3.57	0.967	3.56	0.96	3.54	0.966	3.56	0.967	3.56
β_{comfort}	1.35	4.39	1.22	4.05	1.33	4.3	1.3	4.04	1.28	4.1	1.25	4.09	1.22	3.83	1.33	4.34	1.29	4.2	1.41	4.55	1.15	3.79
λ_{mean}	7.21	10.37	5.46	10.02	5.91	9.99	9	9.66	7.54	9.93	7.38	10.53	9.03	10.02	7.47	9.98	6.83	10.78	7.63	9.56	7.91	10.38
λ_{French}	-0.806	-2.74	-0.549	-1.89	-0.658	-2.24	-0.217	-0.72	-0.595	-2.01	-0.795	-2.72	-0.268	-0.89	-0.456	-1.58	-1.15	-3.83	-0.538	-1.82	-0.227	-0.79
λ_{age50}	-1.1	-4.8	-1.16	-5.12	-0.978	-4.31	-1.24	-5.25	-1.12	-4.86	-1.11	-4.89	-1.29	-5.44	-1.04	-4.62	-1.14	-4.92	-1.08	-4.7	-1.29	-5.67
λ_{active}	-1.17	-4.76	-1.12	-4.6	-1.2	-4.92	-1.09	-4.26	-1.11	-4.49	-1.11	-4.52	-1.1	-4.32	-1.12	-4.62	-1.11	-4.47	-1.2	-4.83	-0.9	-3.68
λ_{cars}	-0.771	-3.33	-0.701	-3.09	-0.738	-3.23	-0.68	-2.84	-0.654	-2.82	-0.737	-3.22	-0.65	-2.72	-0.688	-3.04	-0.804	-3.44	-0.739	-3.2	-0.598	-2.61
α_2	0.0153	0.07	-0.182	-0.85	-0.171	-0.75	0.322	1.2	0.0663	0.27	-0.0066	-0.03	0.385	1.5	0.00209	0.01	-0.0374	-0.18	0.0427	0.17	0.158	0.69
α_3	-0.492	-1.64	-0.753	-2.53	-0.699	-2.26	-0.46	-1.21	-0.49	-1.45	-0.435	-1.43	-0.352	-0.98	-0.686	-1.98	-0.365	-1.34	-0.664	-1.89	-0.534	-1.52
α_4	-2.14	-6.56	-2.35	-7.59	-2.33	-7.2	-2.02	-5.33	-2.19	-6.27	-2.15	-6.63	-1.97	-5.41	-2.34	-6.55	-1.96	-6.75	-2.32	-6.28	-2.3	-6.33
α_5	-2.04	-6.61	-2.25	-8.57	-2.28	-7.86	-1.88	-4.75	-1.89	-6	-2.09	-6.81	-1.82	-4.71	-2.31	-6.75	-1.86	-6.88	-2.22	-6.27	-2.02	-5.7
α_6	-2.67	-7.44	-2.65	-8.3	-2.5	-7.62	-3.24	-6.93	-2.46	-6.71	-2.7	-7.51	-3.04	-6.98	-2.98	-7.49	-2.43	-7.26	-2.95	-6.98	-3.32	-7.97
α_7	-0.762	-3.02	-0.862	-3.45	-0.856	-3.36	-0.879	-2.95	-0.758	-2.81	-0.778	-3.03	-0.907	-3.15	-0.828	-3.03	-0.633	-2.7	-0.816	-2.91	-0.57	-2.19
α_8	-1.15	-3.63	-1.24	-4.33	-1.09	-3.79	-1.35	-3.25	-1.05	-3.31	-1.28	-4.03	-1.27	-3.19	-1.35	-3.97	-1.06	-3.66	-1.34	-3.72	-1.66	-4.4
α_9	-2.2	-5.57	-2.13	-5.68	-2.1	-5.16	-2.86	-5.65	-1.94	-4.72	-2.47	-5.81	-2.8	-5.67	-2.56	-5.9	-2.02	-5.24	-2.46	-5.48	-2.83	-6.61
γ_1	0.234	9.14	0.305	9.03	0.286	8.93	0.198	8.74	0.23	8.91	0.231	9.42	0.199	9.06	0.23	9	0.244	9.51	0.223	8.44	0.217	9.49
γ_2	0.141	3.81	0.204	4.2	0.2	4.31	0.105	3.01	0.145	3.81	0.139	3.83	0.0987	2.94	0.151	3.94	0.139	3.66	0.143	3.66	0.121	3.59
γ_3	0.167	3.6	0.226	3.65	0.22	3.77	0.165	3.77	0.178	3.72	0.161	3.52	0.154	3.64	0.203	4.18	0.145	3.14	0.193	3.97	0.18	3.87
γ_4	0.426	12.81	0.562	13.64	0.531	13.69	0.335	10.03	0.416	12.38	0.421	13.08	0.336	10.72	0.433	13.06	0.421	12.43	0.42	12.25	0.424	12.74
γ_5	0.294	6.72	0.36	7.23	0.366	7.46	0.236	5.26	0.283	6.99	0.305	7.42	0.231	5.17	0.321	7.49	0.291	7.03	0.301	6.72	0.273	5.99
γ_6	0.428	10.8	0.481	9.2	0.435	8.65	0.435	11.43	0.388	9.72	0.423	10.57	0.412	11.23	0.453	11.34	0.414	9.48	0.445	10.93	0.499	13.88
γ_7	0.327	10.2	0.433	10.22	0.408	10.55	0.288	10.03	0.316	10.15	0.326	10.66	0.293	10.79	0.324	10.36	0.326	9.84	0.316	9.66	0.28	10.06
γ_8	0.319	7.91	0.407	8.58	0.365	8.29	0.287	7.01	0.298	8.04	0.333	8.56	0.279	6.95	0.336	8.62	0.321	7.8	0.331	8.16	0.359	8.86
γ_9	0.415	9.45	0.482	8.28	0.456	7.58	0.43	10.79	0.37	8.26	0.443	9.72	0.423	10.52	0.448	10.38	0.412	8.54	0.428	9.87	0.471	12.48
σ_1	0.0425	0.95	0.303	6.74	0.246	5.46	-0.142	-3.02	0.0413	0.93	0.0239	0.54	-0.161	-3.4	0.0627	1.46	0.0423	0.91	0.0753	1.75	-0.0269	-0.63
σ_2	0.346	8.79	0.61	15.2	0.556	13.74	0.251	6.49	0.347	8.72	0.347	8.86	0.228	5.93	0.388	9.85	0.365	9.29	0.39	9.92	0.285	7.38
σ_3	0.358	7.21	0.634	12.65	0.574	11.33	0.222	4.24	0.357	7.02	0.358	7.31	0.209	4.07	0.381	7.37	0.36	7.45	0.373	7.24	0.322	6.37
σ_4	-0.0123	-0.14	0.103	1	0.0625	0.62	0.102	1.6	-0.0571	-0.63	-0.011	-0.13	0.0453	0.7	0.00596	0.07	0.0091	0.11	0.0293	0.36	0.00377	0.05
σ_5	0.174	2.64	0.298	4.27	0.239	3.14	0.254	4.51	0.0589	0.82	0.123	1.8	0.245	4.37	0.155	2.23	0.131	1.96	0.183	2.74	0.298	5.28
σ_6	-0.234	-1.58	0.0163	0.12	0.0793	0.7	-0.239	-1.64	-0.259	-1.76	-0.2	-1.48	-0.2	-1.59	-0.279	-1.74	-0.176	-1.18	-0.28	-1.67	-0.439	-2.09
σ_7	0.00595	0.09	0.217	2.84	0.147	1.97	-0.161	-2.16	-0.0247	-0.36	-0.0243	-0.35	-0.235	-2.94	0.0178	0.27	0.0231	0.33	0.0335	0.51	-0.0248	-0.44
σ_8	0.283	4.65	0.455	7.32	0.377	6.1	0.261	4.21	0.203	3.31	0.259	4.14	0.248	4.03	0.27	4.39	0.296	4.82	0.276	4.43	0.304	4.83
σ_9	0.0786	0.79	0.308	2.93	0.294	2.82	-0.0804	-0.7	0.0619	0.61	0.0463	0.42	-0.0756	-0.66	0.0496	0.48	0.0887	0.8	0.0679	0.68	-0.0345	-0.34

Table 11: Estimation results (2nd part) for the HCMs using the ratings of all evaluators and the median ratings (sample size = 2'265).

Name	Evaluator 23		Evaluator 24		Evaluator 25		Median ratings	
	Value	t-test	Value	t-test	Value	t-test	Value	t-test
ASC _{PT}	-0.152	-0.76	-0.126	-0.63	-0.123	-0.61	-0.155	-0.77
ASC _{PMM}	0.419	2.27	0.41	2.23	0.408	2.22	0.419	2.27
β_{cost}	-0.0653	-8.1	-0.0624	-8.3	-0.0613	-8.33	-0.0653	-8.08
β_{timePT}	-0.0235	-6.54	-0.0179	-5.18	-0.0125	-7.43	-0.0208	-7.06
β_{timePMM}	-0.0323	-9.42	-0.0308	-9.78	-0.0304	-9.97	-0.0323	-9.43
β_{distance}	-0.235	-11.44	-0.232	-11.37	-0.232	-11.36	-0.235	-11.45
β_{workPT}	-0.0438	-0.19	-0.0905	-0.39	-0.112	-0.48	-0.0474	-0.2
β_{timePMM}	-0.576	-2.6	-0.562	-2.54	-0.556	-2.51	-0.575	-2.6
β_{FrenchPT}	-0.11	-0.36	-0.154	-0.5	-0.126	-0.41	-0.0808	-0.26
$\beta_{\text{FrenchPMM}}$	0.969	3.57	0.971	3.58	0.967	3.56	0.967	3.56
β_{comfort}	1.36	4.26	0.687	1.97	0.0757	0.23	1.33	4.34
λ_{mean}	9.25	9.93	8.6	9.28	0.444	1.26	7.47	9.98
λ_{French}	-0.182	-0.61	0.652	2.17	0.832	2.41	-0.456	-1.58
$\lambda_{\text{age}_{50}}$	-1.08	-4.67	-1.21	-5.08	0.746	2.83	-1.04	-4.62
λ_{active}	-1.14	-4.55	-0.348	-1.32	0.476	1.69	-1.12	-4.62
λ_{cars}	-0.658	-2.8	-0.256	-1.02	0.217	0.82	-0.688	-3.04
α_2	0.31	1.18	0.354	1.19	0.295	3.48	0.00209	0.01
α_3	-0.48	-1.24	-0.343	-0.71	0.721	6.3	-0.686	-1.98
α_4	-2.12	-5.49	-2.8	-5.23	0.131	1.23	-2.34	-6.55
α_5	-2.11	-5.2	-1.67	-2.99	0.784	6.63	-2.31	-6.75
α_6	-3.29	-6.97	-3.42	-5.78	1.05	7.53	-2.98	-7.49
α_7	-0.699	-2.44	-0.684	-1.96	-0.0106	-0.13	-0.828	-3.03
α_8	-1.49	-3.59	-1.63	-3.25	0.457	4.63	-1.35	-3.97
α_9	-3.14	-5.94	-2.67	-4.56	0.908	6.89	-2.56	-5.9
γ_1	0.195	9.02	0.218	9.21	0.193	9.08	0.23	9
γ_2	0.106	3.24	0.146	4.16	0.0271	0.67	0.151	3.94
γ_3	0.169	3.94	0.203	3.83	-0.102	-2.28	0.203	4.18
γ_4	0.347	11.01	0.486	10.72	0.371	13.11	0.433	13.06
γ_5	0.258	6.04	0.257	4.05	-0.0351	-0.6	0.321	7.49
γ_6	0.44	11.64	0.497	9.41	-0.193	-3.81	0.453	11.34
γ_7	0.262	9.85	0.289	8.53	0.311	13.68	0.324	10.36
γ_8	0.302	7.97	0.329	6.73	-0.0028	-0.06	0.336	8.62
γ_9	0.454	11.41	0.453	8.71	-0.216	-4.89	0.448	10.38
σ_1	-0.151	-3.31	-0.18	-3.57	-0.291	-5.71	0.0627	1.46
σ_2	0.222	5.75	0.243	5.94	0.143	3.88	0.388	9.85
σ_3	0.209	3.99	0.356	6.67	0.208	4.34	0.381	7.37
σ_4	0.0361	0.53	0.146	1.46	-0.427	-2.78	0.00596	0.07
σ_5	0.228	3.93	0.61	11.89	0.272	6.25	0.155	2.23
σ_6	-0.211	-1.55	0.258	2.57	0.138	2.03	-0.279	-1.74
σ_7	-0.173	-2.62	-0.058	-0.86	-0.781	-4.13	0.0178	0.27
σ_8	0.208	3.27	0.425	6.86	0.103	2.36	0.27	4.39
σ_9	-0.144	-1.12	0.183	1.75	0.0229	0.3	0.0496	0.48

Table 12: Estimation results (3rd part) for the HCMs using the ratings of all evaluators and the median ratings (sample size = 2'265).

Indicator	Mode	Logit	Evaluator 1	Evaluator 2	Evaluator 3	Evaluator 4	Evaluator 5	Evaluator 6	Evaluator 7 (central evaluator)	Evaluator 8	Evaluator 9 (outlying evaluator)	Evaluator 10	Evaluator 11	Evaluator 12	Evaluator 13
Market shares	PT	0.285	0.279	0.279	0.279	0.279	0.279	0.278	0.278	0.279	0.278	0.279	0.279	0.279	0.279
	PMM	0.653	0.659	0.659	0.659	0.659	0.659	0.660	0.659	0.659	0.659	0.659	0.659	0.659	0.659
	SM	0.062	0.063	0.063	0.063	0.062	0.063	0.062	0.063	0.062	0.063	0.063	0.063	0.062	0.063
VOT (CHF/hour)	PT	12.191	11.683	11.578	12.519	12.643	11.698	12.613	12.306	12.493	12.183	12.346	11.618	12.551	12.336
	PMM	30.114	29.483	29.499	29.689	29.707	29.500	29.667	29.485	29.633	29.762	29.492	29.522	29.723	29.492
Elasticities of cost	PT	-0.275	-0.282	-0.284	-0.283	-0.282	-0.281	-0.281	-0.281	-0.283	-0.281	-0.284	-0.284	-0.283	-0.283
	PMM	-0.071	-0.073	-0.073	-0.073	-0.073	-0.073	-0.072	-0.073	-0.073	-0.072	-0.073	-0.073	-0.073	-0.073
Elasticities of time	PT	-0.575	-0.621	-0.624	-0.625	-0.623	-0.622	-0.624	-0.624	-0.622	-0.629	-0.621	-0.625	-0.624	-0.622
	PMM	-0.276	-0.274	-0.276	-0.276	-0.276	-0.274	-0.274	-0.273	-0.276	-0.275	-0.276	-0.277	-0.277	-0.275
Cross-elasticity of cost of PMM	SM	0.040	0.040	0.041	0.040	0.041	0.040	0.040	0.040	0.041	0.039	0.041	0.041	0.041	0.040
	PT	0.154	0.163	0.164	0.163	0.163	0.163	0.162	0.163	0.163	0.163	0.164	0.164	0.163	0.164
Cross-elasticity of cost of PT	PMM	0.115	0.115	0.115	0.115	0.115	0.114	0.114	0.114	0.115	0.114	0.115	0.116	0.115	0.115
	SM	0.050	0.048	0.048	0.048	0.048	0.048	0.048	0.047	0.048	0.047	0.048	0.048	0.048	0.048
Cross-elasticity of time of PMM	SM	0.191	0.186	0.188	0.189	0.190	0.186	0.189	0.184	0.189	0.184	0.189	0.189	0.191	0.188
	PT	0.590	0.606	0.611	0.611	0.610	0.606	0.608	0.606	0.610	0.610	0.610	0.612	0.612	0.609
Cross-elasticity of time of PT	PMM	0.241	0.252	0.254	0.254	0.253	0.253	0.253	0.253	0.253	0.255	0.252	0.254	0.253	0.253
	SM	0.111	0.109	0.110	0.110	0.110	0.110	0.109	0.109	0.110	0.110	0.110	0.110	0.110	0.109
Indicator	Mode	Evaluator 14	Evaluator 15	Evaluator 16	Evaluator 17	Evaluator 18	Evaluator 19	Evaluator 20	Evaluator 21	Evaluator 22	Evaluator 23	Evaluator 24	Evaluator 25	Median ratings	
Market shares	PT	0.279	0.278	0.279	0.279	0.278	0.279	0.278	0.279	0.278	0.279	0.277	0.277	0.279	
	PMM	0.658	0.659	0.659	0.659	0.660	0.659	0.659	0.659	0.660	0.659	0.661	0.660	0.659	
	SM	0.063	0.062	0.063	0.063	0.062	0.063	0.062	0.062	0.063	0.062	0.062	0.062	0.063	
VOT (CHF/hour)	PT	12.462	12.542	12.391	12.494	12.490	12.446	12.627	12.517	12.364	12.466	12.268	12.140	12.446	
	PMM	29.608	29.794	29.512	29.690	29.723	29.655	29.799	29.704	29.620	29.695	29.622	29.792	29.655	
Elasticities of cost	PT	-0.283	-0.282	-0.283	-0.283	-0.281	-0.283	-0.282	-0.284	-0.282	-0.283	-0.279	-0.280	-0.283	
	PMM	-0.073	-0.073	-0.073	-0.073	-0.072	-0.073	-0.073	-0.073	-0.073	-0.073	-0.072	-0.072	-0.073	
Elasticities of time	PT	-0.621	-0.628	-0.623	-0.625	-0.628	-0.623	-0.627	-0.623	-0.629	-0.624	-0.640	-0.642	-0.623	
	PMM	-0.277	-0.276	-0.276	-0.276	-0.275	-0.277	-0.277	-0.278	-0.275	-0.277	-0.272	-0.274	-0.277	
Cross-elasticity of cost of PMM	SM	0.041	0.040	0.041	0.040	0.041	0.040	0.041	0.041	0.040	0.041	0.039	0.038	0.041	
	PT	0.163	0.163	0.164	0.163	0.163	0.164	0.163	0.163	0.164	0.163	0.163	0.163	0.164	
Cross-elasticity of cost of PT	PMM	0.115	0.114	0.115	0.115	0.114	0.115	0.115	0.115	0.114	0.115	0.113	0.113	0.115	
	SM	0.048	0.048	0.048	0.048	0.048	0.048	0.048	0.049	0.048	0.048	0.047	0.046	0.048	
Cross-elasticity of time of PMM	SM	0.189	0.190	0.189	0.189	0.188	0.190	0.190	0.192	0.188	0.190	0.181	0.179	0.190	
	PT	0.610	0.613	0.610	0.611	0.610	0.612	0.613	0.613	0.612	0.612	0.609	0.611	0.612	
Cross-elasticity of time of PT	PMM	0.253	0.254	0.253	0.254	0.254	0.253	0.254	0.253	0.254	0.253	0.258	0.260	0.253	
	SM	0.110	0.111	0.110	0.110	0.110	0.110	0.110	0.111	0.110	0.111	0.111	0.110	0.110	

Table 13: Market shares, VOTs, direct and cross-elasticities for the logit model, the models using the ratings of the evaluators and the model using the median ratings (sample size = 2'265).