An empirical analysis of the valuation of travel time savings within consumer choices among modern urban transportation modes

Master Thesis by Daniel Rohbeck

7/21/2015

Program: Master in Economics: Empirical Applications and Policies

Faculty: Faculty of Economics and Business Studies, UPV/EHU, Bilbao

Advisor: Petr Mariel

I. CONTENT

II. FIGURES

III. ABBREVIATIONS

ASC	Alternative specific constant
AVC	Asymptotic Variance-Covariance Matrix
C2G	Car2Go, Car Sharing Brand
CLM	Conditional Logit Model
CS	Carsharing
DCE	Discrete Choice Experiment
ECM	Error Component Model
GoF	Goodness-of-fit indicator
Iid	Independently identically distributed
LCM	Latent Class Model
LM	Lagrange Multiplier
MNL	Multinomial Logit Model
Moovel	Moovel GmbH, Company behind C2G and subsidiary of Daimler AG
MXL	Mixed Logit Model
NLM	Nested Logit Model
OLS	Ordinary least square regression
RCE	Revealed Choice Experiments
RPM	Random Parameter Logit Model
SCE	Stated Choice Experiments
VoT	Valuation of Travel-Time Savings

1. Introduction

As recent years have seen an accelerated growth of new urban transportation modes like car sharing (Car2Go, DriveNow, Blablacar, etc.) and new taxi formats (uber, lyft, etc.), the aim of this study will be to assess the value of travel time for urban transportation choices. While there is a wide range of empirical studies for intercity travel, the valuation within urban transportation is not yet researched in depth (Hensher , 2011). The growth of these new mobility modes, often also called shared mobility, is driven by the availability of live, local information of transport options as supported by the broad availability of smartphones. The smartphone has become the ultimate navigation tool by combining maps, options and timetables in an easy to use fashion.

As the choice of transport mode becomes on the one hand more easy to assess due to the transparent listing of different activities with local specific comparison of alternatives, but on the other side at the same time more complex due to the increasing number of options, the motivation of choosing one alternative will probably differ per individual and in certain situations. The first aim of this study is therefore to identify, which factors influence the probability of choosing one option under different circumstances in general and what kind of incentive must be offered to increase the usage for a certain individual. The second step will become increasingly important as neither pricing nor availability will remain statically in the future. Already today mobile payment systems allow for individual and situation specific pricing and with autonomous vehicles the regional distribution of car sharing fleets will become dependent on the demand, thereby making access and egress travel time flexible under certain circumstances. A second aim of this study is therefore to provide decision maker of public transport services and providers of car sharing services with an assessment of the valuation of travel time savings in an urban transportation context.

To answer these two research questions, the data of a stated discrete choice experiment (DCE) from a market entry study of Car2Go (C2G) will be used to estimate the Valuation of Travel-Time Savings (VoT) for different activities, situations and individual characteristics. The VoT will be estimated based on conditional logit (CLM) and mixed logit models (MXL) using random parameters based on just a random term, a random term plus situational factors and a random term plus individual parameters. The DCE has been run with 800 individuals and each individual had to answer three choice situations. The design of this experiment was based on actual regional parameters of the used alternatives (prices, typical distances and travel times etc.) and while only public transport modes were explicitly named, a “none of the above” was supposed to capture other alternatives. The study has been conducted in cooperation with moovel GmbH (*moovel*), the German based Daimler AG subsidiary behind C2G. As the raw data for this analysis has been surveyed by *moovel* and all rights remain with the company, all alternative specific constants (ASCs) cannot be shown in the thesis. These censored results however will already give a good impression on the influences of situational and individual factors on urban travel mode decision processes as the general VoTs are not affected.

The thesis is structured into five parts: First the above introduction, followed by a short literature review on the use of DCE in transportation studies and the methodology behind the estimation of discrete choice models in general. A third part explains the choice model used and survey setup in specific. The fourth part will then review the estimation process and the following fifth part presents the results of the DCE with respect to the proposed research questions. Finally the conclusion will summarize the analysis and give an outlook on further questions for research.

2. Methodology

2.1. Empirical studies on choice of transportation modes and the value of travel time

The importance of the total travel time for the decision of transport modes has been analyzed in various studies based on revealed and stated choices, used for example to calculate cost-benefits for different transportation planning scenarios or to estimate optimal prices for transportation services from the 1970s until now (i.e. De Donnea (1972), Beesley (1973), Accent (1994), Kamga (2014)).

In his summary of the current status of literature on transportation studies, Hensher (2011) states that travel time savings are still the single most relevant user benefit in all transportation studies and gives a synthesis of empirical findings from various studies within this field as shown in Figure 1.

Country	Year	Geo-graphical Loen	Distance	Mode	Purpose	Per hr	To USD2000
Sweden	1994	various	> 50 km	air	business	141	18.682
Norway	1995	various	> 50 km	air	business	343	55.381
Norway	1995	various	100–300 km	air	business	258	41.657
Norway	1995	various	> 300 km	air	business	324	52.314
USA	1995	intercity	Nationwide	air	business	34.5	37.715
Australia	1999	Syd–Can	300 km	All	business	46.71	31.352
Finland	1995	intercity	varies	car	business	124.6	28.604
Norway	1996	intercity	< 50 km	car	business	253	40.850
Norway	1995	intercity	50–100 km	car	business	377	60.871
Norway	1995	intercity	100–300 km	car	business	207	33.423
Norway	1995	intercity	> 300 km	car	business	137	22.120
Sweden	1994	intercity	> 50 km	car	business	167	22.126
Sweden	1994	intercity	> 50 km	IC-Train	business	129	17.092
Sweden	1994	intercity	> 50 km	X2000-Train	business	134	17.754
Sweden	1994	various	> 50 km	air	non-business	88	11.659
Norway	1995	various	> 30 km	air	non-business	155	25.027
Norway	1995	various	50–100 km	air	non-business	172	27.771
Norway	1995	various	100–300 km	air	non-business	170	27.449
Norway	1995	various	> 300 km	air	non-business	151	24.381
USA	1995	intercity	Nationwide	air	non-business	19.5	21.317
Chile	1993	intercity	520 km	air	non-business	13312	37.311
Spain	1992	intercity	40 min	air	non-business	1360	15.158
Australia	1999	Syd–Can	300 km	All	non-business	10.05	6.746
Finland	1996	intercity	varies	car driver	non-business	16	3.673
Norway	1995	intercity	> 50 km	car driver	non-business	86	13.886
Norway	1995	intercity	50–100 km	car driver	non-business	101	16.308
Norway	1995	intercity	100–300 km	car driver	non-business	97	15.662
Norway	1995	intercity	> 300 km	car driver	non-business	77	12.433
Sweden	1994	intercity	> 50 km	car driver	non-business	81	10.732
New Zealand	1999	intercity	30–540 mins	car driver	non-business	7.86	4.588
New Zealand	1999	intercity	< 100 mins	car driver	non-business	6	3.503
New Zealand	1999	intercity	100–200 mins	car driver	non-business	7.6	4.437
New Zealand	1999	intercity	200–300 mins	car driver	non-business	9	5.254
New Zealand	1999	intercity	300–400 mins	car driver	non-business	10.5	6.129
New Zealand	1999	intercity	> 400 mins	car driver	non-business	12.1	7.063
New Zealand	2000	intercity	> 3 h	car driver	non-business	6.97	4.069
Sweden	1994	intercity	> 50 km	IC-Train	non-business	74	9.804
Sweden	1994	intercity	> 50 km	X2000-Train	non-business	102	13.514

Figure 1: Synthesis of empirical findings on the Valuation of Travel Time Savings (from Hensher, 2011, in US2000-$)

Comparing the VoT for one hour of saved time the estimations range from 4.33€ (adjusted to 2014 prices[1]) for private intercity car travel up to 74.25€ for business intercity car travel and average at 29.26€ for non-business, non-air travel. But as all of these studies focus on intercity travel and

[1] All inflation adjustments were done using the Retail Prices Index as published by the Office For National Statistics (UK, 2015) for pounds, the Consumer Price Index as published by the Bureau of Labor Statistics (USA, 2015) for US-$ and the Consumer price index as published by EuroStat (2015), the currency conversion was done using the actual market trading rate from 17.06.2015 (finanzen.net, 2015).

distances greater 30km, the VoTs cannot be directly applied for urban transportation, where the distances and total travel-times are in most cases shorter. Also while most studies focused before 2000 on the total differences in time, current research analyzes the heterogeneity of trip time such as walk, wait, transfer and in-vehicle time or traveling at different speeds for road-based travel modes (i.e. Hensher (2001a), Hensher (2001b), Amador et al. (2005), Gunn (2007)) and there is an ongoing debate on the influence of the type of transportation mode and underlying demographics on the VoT (i.e. Wardman (2004), Shires and de Jong (2009)).

Differentiating by mode and type of activities Wardman (2011) applied a meta-analysis on 226 studies covering local, regional and national travel carried out between 1960 and 2008 in the UK and stated that the average VoT is from 0.075€ per Minute (adjusted to 2014 prices) for bus commuters over 0.115€/Min for car commuters up to 0.137€/Min for train commuters when looking at a travel distance of 2 miles (= 3.2 km). He also found that the valuation of walking time is on average between 1.25 (bus), 1.42 (car) and 1.45 (train) times the VoT of in-mode travel time, but as the estimation and embedding of walking distances is in most cases only a secondary application, Wardman also states the need of further research in this area. Wardman's results vary a little from the results of a similar meta-analysis by Shires and de Jong (2009), who find an average VoT of between 0.095€/Min and 0.270€/Min (adjusted to 2015 prices) for bus commuters in 25 European countries and 0.183€/Min in the UK as shown in Figure 2.

Country	Bus	Other modes (car, train)
Austria	8.52	10.30
Belgium	8.12	9.82
Cyprus	6.95	8.40
Czech Republic	5.91	7.14
Denmark	9.89	11.96
Estonia	5.17	6.26
Finland	8.52	10.30
France	10.17	12.30
Germany	8.20	9.91
Greece	6.43	7.77
Hungary	5.36	6.48
Ireland	9.37	11.33
Italy	9.05	10.94
Latvia	4.64	5.61
Lithuania	4.63	5.60
Luxembourg	13.11	15.85
Malta	5.64	6.82
Netherlands	8.22	9.94
Poland	5.02	6.08
Portugal	5.98	7.23
Slovakia	4.74	5.74
Slovenia	8.63	10.44
Spain	7.59	9.18
Sweden	9.39	11.35
UK	8.87	10.73
EU average	8.84	10.69
Switzerland	10.64	12.87

Figure 2: Empirical meta-analysis on the Valuation of Travel Time Savings within 25 European Countries (from de Jong, 2009, in 2003-€)

2.2. Discrete Choice Modelling

2.2.1 Multinomial logit and conditional logit models

The method of Discrete Choice Experiments is based on the Random Utility Theory as developed by McFadden (1974) and Manski (1977). Following the underlying theory of utility maximization as first proposed by Thurstone (1927) and further developed into the binary probit model by Marschak (1960), the decision maker will maximize his utility by choosing the alternative with the highest utility score in a discrete choice situation. In Random Utility Models the utility for each individual n of each alternative i is composed by a deterministic part $V_{ni}(\beta) = \beta' X_{ni}$, where X_{ni} is a vector of characteristics of the alternative and β a corresponding vector of weights[2], and a stochastic part ε_{ni}:

$$U_{ni} = \beta' X_{ni} + \varepsilon_{ni} \quad .$$

While McFadden argues in his original paper that the deterministic part can be contributed to "… the 'representative' taste of the population, and ε […] reflects the idiosyncrasies of this individual in tastes for the alternative with attributes x …" (McFadden, 1974, p. 108), Manski increased the weighed of the random part by letting it account also for unobserved attributes of the alternative affecting the decision process, measurement errors and functional misspecification (Manski, 1977). Today an even wider argumentation for this utility form is based on prospect theory by Kahneman and Tversky (1979), who argue that individuals base their decisions not only on certain but also on probabilistic characteristics, which might or might not be accessed correctly, and therefore choose the alternative with the highest expected utility. Therefore the random error term does also account for the uncertainty in the decision process for each individual (Manrai, 1995).

No matter which argumentation is followed the unobserved probability of choosing alternative i in a discrete choice situation, meaning that there is a finite set of alternatives, can then be stated as follows:

$$\begin{aligned} P_{ni}^{*} &= Prob\big(U_{ni} > U_{nj}, \forall\, j \neq i\big) \\ &= Prob\big(\beta' X_{ni} + \varepsilon_{ni} > \beta' X_{nj} + \varepsilon_{nj}, \forall\, j \neq i\big) \\ &= Prob\big(\varepsilon_{nj} - \varepsilon_{ni} < \beta' X_{ni} - \beta' X_{nj}, \forall\, j \neq i\big) \\ &= \int_{\varepsilon} I\big(\varepsilon_{nj} - \varepsilon_{ni} < \beta' X_{ni} - \beta' X_{nj}, \forall\, j \neq i\big) f(\varepsilon_n) d\varepsilon_n, \end{aligned}$$

where $I(\cdot)$ is an indicator function, equaling one if alternative i was chosen and zero otherwise, and $f(\varepsilon_n)$ is the density function of the assumed random term. As U_{ni} is a latent variable and therefore not observed, the estimation depends on the assumptions on the distribution of ε_n and different assumptions for $f(\varepsilon_n)$ will lead to different discrete choice models. The most commonly used specifications are the multinomial logit, where the unobserved part of the utility is independently identically distributed (iid) extreme value, and the probit, where the density of the random term follows a multivariate normal distribution (Train, 2009).

Using the multinomial logit specification the latent probability of alternative i can then be rewritten as

$$P_{ni}^{*} = Prob\big(U_{ni} > U_{nj}, \forall\, j \neq i\big) = \frac{\exp\left(\frac{\beta' X_{ni}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta' X_{nj}}{\mu}\right)} \quad ,$$

[2] It is assumed throughout this thesis that the functional form of the utility is linear in β. In fact $V_{ni}(\beta)$ could also take other functional forms, but the linear form is most common (Train, 2009).

where μ is a positive scale parameter and normally adjusted to fit one benchmark alternative z:

$$P_{nz}^{*} = \frac{1}{\sum_{j=1}^{J} \exp\left(\frac{\beta' X_{nj}}{\mu}\right)} \text{ and } P_{ni}^{*} = \frac{\exp\left(\frac{\beta' X_{ni}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta' X_{nj}}{\mu}\right)} \text{ for } \forall\, i \neq z \text{ (McFadden D. , 1974).}$$

This model specification is also known as multinomial logit model (MNL) and can be estimated maximizing the likelihood function for $P_n^{*} = Prob(y_n = m|\beta, X_{ni})$ when observing the outcome y_n to be alternative m:

$$L(\beta|y_n, X_{ni}) = \prod_{n=1}^{N} P_n^{*}$$

$$= \prod_{m=1}^{J} \prod_{y_n=m} \frac{\exp\left(\frac{\beta' X_{nm}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta' X_{nj}}{\mu}\right)} \quad ,$$

where $\prod_{y_n=m}$ is the product over all cases for which $y_n = m$ (Long, 1997, S. 157).

There are two possible setups for use of multinomial logit models (MNL) in economics. The first aims to identify individual characteristics, which affect the decision making process. In this case the value of the alternative specific characteristics are fixed over all individuals and individual characteristics can be weighted differently for each alternative:

$$P_{ni}^{*} = \frac{\exp\left(\frac{\beta_i' X_n}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta_j' X_n}{\mu}\right)} \quad .$$

The second as used in the equations above keeps the individual characteristics within each choice situation fixed and allows the alternative specific characteristics to fluctuate in value. This second setup estimates probabilities of choosing alternative i conditional on a certain vector of alternative specific characteristics and is therefore also called CLM:

$$P_{ni}^{*} = \frac{\exp\left(\frac{\beta' X_{ni}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta' X_{nj}}{\mu}\right)} \quad .$$

As the above described model implies that only differences between alternatives enter the choice process, this differentiation is critical since CLM cannot include variables, which are constant for one individual like i.e. age or gender. A combination of both models is possible by using the standard MNL specification and including alternative specific dummy variables and dummy interactions (Long, 1997, S. 181).

2.2.2 Restrictions of multinomial and conditional logit models

Both the CLM and MNL in their basic configuration are subject to a number of assumptions due to the use of IID random components. First and most important this requires that the probability of each alternative over another alternative is independent of all other possible alternatives for one individual. This also called Independence of Irrelevant Alternatives Axiom (IIA) means that if the probability of alternative A is 25% in a set of alternatives {A,B} and a third alternative C is added, the ratio $Prob(A)/\,Prob(B)$ is kept constant i.e. $Prob(A)$ is 12.5%, $Prob(B)$ is now 37.5% and $Prob(C)$ is 50%. This property therefore becomes critically if some of the alternatives are close substitutes (McFadden D. , 1974) or if the substitution is not symmetric between all alternatives for changes in properties. If for example a change in price for taxi has a different impact on the probability of bus then on choosing the car as car and taxi are closer substitutes then bus and taxi (Ben-Akiva M. L., 1985; Baltas, 2001).

Similar the preferences between alternatives must not be independent and could be correlated due to unaccounted similarities between the alternatives or shared unobserved characteristics. Again a change in price for taxi might imply a difference in substitution by car and bus as car and taxi share the same travel time variation while the travel time variation with bus might differ. This would also imply that the random components are not independent (McFadden D. , 1981).

Following the argumentation of Manski and Manrai to justify the existence of the random term also the question of differences in the scale of the random part between alternatives rises. The uncertainty in estimating the utility of choosing bus might be higher than for the use of car. Similar the unobserved characteristics might influence the choice of bus more than the choice of taxi, therefore not justifying the assumption of identically distributed utilities between alternatives (Baltas, 1998).

A final caveat of CLM and MNL models comes into play when looking at panel data or in general data with more than one observation per individual. As the random term captures both interpersonal and intrapersonal dynamics, it can be correlated for one individual either due to static unobserved factors between individuals or preference inertia / habit-persistence within one individual (Baltas, 2001).

Summing up the stochastic restrictions, MNL and CLM should not be used in cases, where there is:

1. Differences in substitution pattern between alternatives
2. Similar or correlated unobserved attributes between alternatives
3. Unobserved taste heterogeneity between alternative specific attributes or between individuals
4. Inter- or Intrapersonal dynamics when looking at more than one observation per individual

2.2.3 Mixed logit models

The standard CLM has therefore undergone many extensions in current years. Different extensions of the model like the multinomial probit, the nested multinomial logit or the generalized extreme value model allow also for unobserved heterogeneity of alternatives and individuals either in panel or cross-sectional data (Baltas, 2001). But the most general and flexible extension commonly used is the MXL, which allows by extending the utility function to

$$U_{ni} = (\beta + \eta_{ni})' X_{ni} + \varepsilon_{ni},$$

where η_{ni} is the vector of heterogeneity, unlike the MNL for random taste variations, unsimple substitution patterns and correlation in unobserved factors over (Train, 2009).

The mixed logit takes its name from the mixed functional form of its choice probabilities, which can be expressed as the standard logit probabilities but as a multidimensional integral weighted by a density function dependent on the coefficients:

$$P_{ni}^{*} = \int \frac{\exp\left(\frac{\beta' X_{ni}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta' X_{nj}}{\mu}\right)} f(\beta)\, d\beta \quad .$$

In the special case of $f(\beta = b) = 1$ and $f(\beta \neq b) = 0$ this collapses back to the simple CLM

$$P_{ni}^{*} = \frac{\exp\left(\frac{b' X_{ni}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{b' X_{nj}}{\mu}\right)}$$

and if the mixing distribution $f(\beta)$ is discrete and β therefore only taking a finite number of values M,

the MXL becomes the latent class model (LCM)

$$P_{ni}^{*} = \sum_{m=1}^{M} \frac{\exp\left(\frac{\beta_m' X_{ni}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta_m' X_{nj}}{\mu}\right)} \quad .$$

But the most common application of MXL is assuming $f(\beta)$ as continuous, where $f(\beta)$ for example follows a certain common continuous probability distribution like i.e. normal, lognormal or triangular, which are defined by a mean b and a covariance σ^2 (Train, 2009). By allowing $f(\beta|\theta)$ to be a function conditional on individual and choice situational characteristics plus a random term, the researcher can therefore integrate individual taste heterogeneity, inter- and intrapersonal time dynamics and by letting the random term be correlated over alternatives or individuals also different substitution patterns or unobserved taste heterogeneity (Train, 2009). Defining β_{nit}, where t is the sub-index for the choice occasion, for example as

$$\beta_{nit} = \gamma_0 + \gamma_i{}'\eta_{ni} + \gamma_n{}'\eta_{ni} + \rho_{nit}{}'y_{nt-1} + \omega_{nit}$$

incorporates alternative specific heterogeneity ($\gamma_i{}'\eta_{ni}$), individual specific heterogeneity ($\gamma_n{}'\eta_{ni}$), intrapersonal dynamics ($\rho_{nit}{}'y_{nt-1}$) and depending on the distribution and covariance of ω_{nit} further unobserved characteristics.

As a result the utility of individual n for alternative i is now not given anymore by a scalar, but by a distribution of the random parameter ω_{nit}, which mean is shifted by the individual and alternative specific parameters. The coefficient in the utility function can therefore be composed into its mean ($\beta = \gamma_0$) and deviations ($\eta_{nit} = \gamma_i{}'\eta_{ni} + \gamma_n{}'\eta_{ni} + \rho_{nit}{}'y_{nt-1} + \omega_{nit}$), which leaves the above stated utility function

$$U_{nit} = (\beta + \eta_{nit})'X_{nit} + \varepsilon_{nit} \quad ,$$

with $\beta_{nit} = (\beta + \eta_{nit})$ being the random coefficient in a random parameter model (RPM). Mathematically the same, but with a slightly different focus, the model can also be expressed as an error component model (ECM)

$$U_{nit} = \beta' X_{nit} + \eta_{nit}{}'Z_{nit} + \varepsilon_{nit},$$

where X_{nit} would be the fixed and Z_{nit} the random parameters of the RPM. While technically the same the random parameter model aims at modelling and analyzing underlying heterogeneity, the error component model is used to emphasis certain substitution patterns and predict outcomes. This functional form allows to nest a wide range of models: Adding i.e. dummy variables for each cluster of individuals k as random parameters with $\omega_{nit} \sim N(0, \sigma_k)$ will give an analog to the nested logit model (NLM) with correlation within and non-correlation between the clusters and σ_k being analog to the inclusive value coefficient (Train, 2009). The MXL can therefore be seen as the generalization of most discrete choice models and nests among others CLM, LCM, NLM, RPM and ECM. McFadden and Train (2000) actually show that any random utility model can be approximated with a MXL to any degree of accuracy if the appropriate variables and mixing distributions are chosen.

2.2.4 Estimation of mixed logit models

As the multidimensional integral of choice probabilities normally does not take a closed form, the estimation of the MXL uses simulation technique to simulate the distribution of the random coefficients. Using the choice probability of the MXL

$$P_{ni}^{*} = \int L_{ni}(\beta) f(\beta|\theta)\, d\beta$$

$$\text{with } L_{ni}(\beta) = \frac{\exp\left(\frac{\beta_n{}' x_{ni}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta_n{}' x_{nj}}{\mu}\right)}$$

can be approximated by drawing values of β for any value of θ as

$$\hat{P}_{ni}^{*} = \frac{1}{R} \sum_{r=1}^{R} L_{ni}(\beta^{r}) \quad , \qquad (1)$$

where *R* is the number of repetitions and β^{r} is the rth draw of β. $\hat{P}_{ni}^{*}$ is an unbiased estimator of P_{ni}^{*}, which variance decreases with increasing *R* (Train, 2009). There are various methods to determine the drawing algorithm, but most commonly used and sufficiently enough efficient are Halton draws, which use a quasi-random, but more uniform distribution of draws by splitting a unit interval into a prime number of parts and then continue splitting each of those intervals again by the same prime number until the needed amount of draws is reached (Train, 1999; Bhat, 2001). The simulated probabilities from (1) are then inserted into the log-likelihood function to estimate the coefficients and distribution parameters

$$SLL = \sum_{n=1}^{N} \sum_{i=1}^{I} d_{ni} \ln \hat{P}_{ni}^{*} \quad ,$$

where d_{ni} is a dummy equaling one if the alternative was chosen. The estimation can be extended to panel data by replacing the choice probability by the probability of that series of choices for one individual (Train, 2009):

$$L_{ni}(\beta) = \prod_{t=1}^{T} \frac{\exp\left(\frac{\beta_n{}' X_{nit}}{\mu}\right)}{\sum_{j=1}^{J} \exp\left(\frac{\beta_n{}' X_{njt}}{\mu}\right)} \quad .$$

While the possibilities of MXL are very wide, its heavy dependence on (Quasi-) Monte Carlo Simulations to approximate the complex error structure and therefore necessary computational resources have simple the use in the past. Therefore these models have only been increasingly popular in recent years with the availability of computational power (Revelt and Train 1998, Train 1998, McFadden and Train 2000, Ben-Akiva et al. 2001, Hensher and Greene 2003, Train 2009 and Hoyos 2009). The transportation literature on this topic is vast and the following references can be mentioned: Viton (2004), Small et al. (2005), Han et al. (2001), Bekhor et al. (2002), Frejinger and Bierlaire (2007), Hess and Polak (2005), Carrier (2003), Brownstone et al. (2000), Hess et al. (2006), Bhat and Castelar (2002). MXL was used to capture the value of time in transportation choice experiments by i.e. Algers et al. (1998), Brownstone and Small (2005), Hess et al. (2005) and Greene et al. (2006).

2.3 Discrete Choice Experiment Design

2.3.1 Discrete Choice Experiments

The purpose behind conducting DCEs is to determine the independent influence of different variables on observed choices and to establish a causal relationship. As in all experiments there is a trade-off between control for unobserved influences and external validity of the results. While stated choice experiments (SCE) are more easily controlled and allow to restrict the number of alternatives, revealed choice experiments (RCE) have a higher external validity. But as DCEs are often conducted for policy or enterprise research to predict hypothetical compared to existing alternatives, SCE using simulated choice situations were introduced by Louviere & Woodworth (1983) / Louviere & Hensher (1983) and are common in marketing, transportation, health and environmental economics (Louviere, 1992). The difference to most typical survey data analysis is that in DCEs only the dependent variable is captured while the independent variables are apart from some individual specific covariats predetermined by the researcher (Rose, 2014).

SCE normally consist of numerous respondents being asked to choose alternatives from a finite set of possibilities in a hypothetical situation. The alternatives can be representations of existing or theoretical options, equal a status quo and be labeled or unlabeled. Each alternative is defined in a number of attribute parameters and a choice situation can be set in a specific scenario context. The level of the attribute parameters and the scenario are drawn from an underlying experiment design (Rose, 2014).

The created choice situations are presented to the observed individual in a paper-based or digital survey and might be assisted with explanations by the interviewer. As causal inference can only be conducted, if there is no external influences and all parameters are understood and taken into account in the decision process, the visual design of the choice situation presentation should be tested in a pre-study and taken into account when analyzing the results.

2.3.2 Choice set formation and parameter selection

Conceptual the experimental design is just a matrix of values assigning levels to attributes over alternatives within choice situations. While some researcher set up the design in a wide format, where one row equals one choice situation (i.e. Bliemer, 2009), others use a long format, where each line equals one alternative and a number of rows are grouped to form a choice situation (i.e. Sándor and Wedel, 2001, 2002, 2005).

Either way the researcher needs to decide on the number and form of alternatives that enter the design. While labeled alternatives have the benefit of an higher external validity, unlabeled alternatives offer the possibility of a wider and more independent distribution of alternative attributes as they are not forced to stay within a given and at least marginal realistic framework. Therefore unlabeled alternatives are in most cases generic, while labeled alternatives are simple and often involve estimation of alternative specific parameters (Rose, 2014). Similar the inclusion of a status quo alternative is in line with the prospect theory and can be seen as a realistic capture of anchoring preferences, with all its benefits and caveats. Offering a “none of the shown alternative” will increase external validity as it captures all not represented possibilities, but is at the same time hard to interpret and might induce a status quo bias (i.e. Tversky and Shafir 1992, Dhar and Simonson 2001, Kontoleon and Yabe 2003).

Also the decision on the number and spread of attribute levels and their functional form will greatly influence outcomes. If for example a non-linear effect of an attribute is to be included in the model, then more than two levels need to be inserted for this attribute. Typically this would be implemented using dummy or effect coding, so the number of parameters needed for each attribute would be the

number of levels minus one and therefore increase the minimum number of different choice tasks. Similar the researcher has to decide, if the design should be balanced, meaning that each attribute level appears the same number of time over all choice tasks or if there are restrictions for certain combinations of levels over attributes or alternatives, which are common for labeled alternatives and often oppose balance to the design (Rose, 2014).

The range of the attribute levels needs to be balanced between too narrow for statistically relevant results as alternatives become indistinguishable and too wide for realistic representation of labeled alternatives or utility balance for unlabeled and labeled alternatives. While literature suggest that a wide attribute level design will decrease standard errors and therefore increase the statistical relevance of the estimates (Bliemer, 2010), a too wide design might result in utilities and therefore choice probabilities that create dominant choices (Rose, 2014). Further the range of the attribute levels must again match an at least marginal realistic framework, if labeled alternatives are analyzed or there are logical restrictions in the alternative formation (Rose, 2014).

Most DCEs are only capturing one chosen alternative, but there are some current studies using rankings as response mechanism (Flynn et al., 2008). While there are some psychological downsides to this, an increasing number of stated ranks (i.e. only best, best and worse, full) will decrease the number of observations necessary to retrieve statistically relevant results (Rose, 2014).

The minimum number of necessary different choice tasks is therefore a function of the number and properties of alternatives, parameters and parameter levels conditional on the response mechanism and the type of model that should be used to estimate. While most literature focuses on MNL the number and optimal composition of choice tasks might also differ for a MXL or other alternative models (Rose, 2014). Similar the number of needed surveyed individuals depends on the chosen design and the number of choice tasks per individual (Rose, 2014).

2.3.3 Orthogonal and efficient designs

A common caveat of SCE is that unless the number of individual specific observations is very high, it is necessary to pool the responses received from various individuals to obtain statistically reliable results. At the same time only for very limited designs with very few alternatives, attributes and levels it will be possible to observe all possible combinations also called a full design. For more complex models the number of possible choice sets grows exponential with the number of alternatives, parameters and parameter levels. Therefore not only the underlying model parameters but also the decision, which of the possible choice sets should be included into the survey, becomes of major importance (ChoiceMetric Pty Ltd., 2014). Louviere et al. (2008) as well as Bliemer & Rose (2011) show empirical evidence suggesting that the attribute level distribution will effect estimation results, while Burgess & Street (2005) and Sándor & Wedel (2001, 2002, 2005) address the same problem from a theoretical view point. The distribution of attribute levels a researcher chooses for example might determine whether or not an independent assessment of each attribute's contribution can be estimated and the statistical power of the experiment to detect statistical relationships. The statistical power is linked to the sample size and might not be as important for big samples of the population, but in most common used DCEs the sample sizes are simple and therefore a poor design might diminish an experiments results (Rose, 2014).

To diminish this risk instead of just randomly assigning the attribute levels there have been different methods established to finding a more balanced and optimal design. The aim of the optimization is to find a design, which allows for a higher statistical efficiency and increase the ability to detect independently the effects of multiple variables on some observed outcome. Statistical efficiency means in this context that given a fixed sample size the models estimations are more precise and

standard errors are reduced. The focus is therefore on the relation between the experiment design and the variance-covariance matrix (Rose, 2014).

The early work on SCE was focused on the differences and advantages of DCE over the up to this date common conjoint analysis, which used regression on linear parameter designs, and therefore first optimization methods also looked at linear parameter models.

The variance-covariance matrix in linear regression models can be written as

$$V(X) = \sigma^2 (X'X)^{-1},$$

where σ^2 is the model variance and X as before the vector of alternative parameters. Assuming the model variance to be fixed, as it is only a scaling factor, the elements of $V(X)$ can be minimized by setting the elements of X to be orthogonal. This property also guarantees non-multi-collinearity in linear models and is therefore under these assumptions a desirable property, which allows for both mentioned criteria of small variances and independently estimates for multiple variables. While σ^2 can still have an effect as it is also not independent of the coefficients in the model and therefore some orthogonal designs might optimize the right part of the equation at the cost of the left side, orthogonal models tend to perform well overall for linear models (Louviere et al., 2000).

The problem with these type of models arises from its assumptions as most models used in SCE are including non-linear parameters and only in unlabeled SCE the setting of parameters is completely free, while for labeled (and some unlabeled) experiments there might be need of a certain part of multi-collinearity due to logic restrains. In the late 1980s and the subsequent decade a group of researcher at the University of Leeds around Fowkes, Toner and Wardman tried to solve the second caveat of orthogonal designs and began to develop binary designs based on non-zero priors and boundary values to include constraints on parameter level combination for both generic and alternative specific parameters. As the researchers mainly focused on transportation models, parameters are assumed to be continuous in these models. Those designs were optimized precisely on the assumed to be known priors and lose efficiency, if the true coefficients differ. But the focus of the group was less to optimize statistical efficiency directly than to generate robust estimates for WTP estimates by optimizing the variances of the ratio of parameter coefficients. While not guaranteeing to find the optimal design due to the constraints and priors, the designs can still be robust against miss-specification of one parameter prior as the efficiency of the model is based on all parameters (Fowkes and Wardman 1988; Fowkes et al. 1993; Toner et al. 1998 and 1999; Watson et al. 2000).

While the aim of Wardman and his team was to generate more realistic choice tasks, Bunch et al. (1996) tried to find more efficient strategies in design generation for MNL. Using zero and non-zero priors together with the designs parameter levels utilities for each alternative and therefore choice probabilities can be calculated. Based on these the asymptotic variance-covariance matrix (AVC) Ω_N can be determined as the inverse of the Fisher information matrix I_N, which is the negative second derivate of the log-likelihood function, where N is the number of considered respondents (Train, 2009). By changing the parameter levels within the choice sets the AVC can then be minimized. Bunch et al. (1996) proposed to use the D-error statistic, which is the determinant of the AVC for N set to one and normalized by the number of parameters K, to find the most efficient design.

$$D\text{-}error = \det(\Omega_1)^{\frac{1}{K}} \quad .$$

At the same time Huber & Zwerina (1996) in a very similar paper showed, using the D-error optimization method but letting go of the orthogonality constraint and relaxing attribute level balance, that orthogonal designs not only restrict choice task formation, but also can generate non-efficient designs for non-zero local priors, if non-linear parameters are included. A second more controversially

discussed finding of this paper was that designs with roughly balanced choice probabilities will result in statistically more efficient models. Rose and Bliemer (2014) and Louviere (2010) for example argue that this is not in line with earlier findings of the Leeds group and later findings by Kanninen (2002), which suggest that optimal designs will result from parameter level combinations that produce certain unbalanced choice probability combinations and proof these for binary SCE with generic alternatives. Using the D-error statistic Kanninen also extends the Leeds group approach of using value boundaries and restrictions to models with more than two alternatives by optimizing the statistically significance for the variance of the parameter ratios instead of the variance of the parameters.

A further improvement to design generation was added by Sándor and Wedel (2001), who relaxed the assumption of perfect knowledge of the priors by using a Bayesian like approach to models with effect coding and fixed attribute levels. Assuming a specific distribution depending on the uncertainty and given information for each parameter prior instead a fixed value and calculating the efficiency of each design over a number of random draws from those distributions. In subsequent research they also showed that these results can be extended to MXL, if one adopts the AVS (Sándor and Wedel, 2002; 2003). Also Rose and Bliemer (2006) adopted the approach to include demographic data and Ferrini and Scarpa (2007) to panel data incorporating preference correlations. Bliemer and Rose (2010) analyzed the differences in designs needed for MNL and MXL in panel and cross-sectional data. While designs optimized for MNL and panel MXL are similar efficient when used for the other, MNL and cross-sectional MXL optimized designs differ widely in efficiency and needed number of observations when used for the other. Parallel Rose et al. (2009) proposed a model averaging process, if researcher need to estimate different models from one design.

Even as the discussion in literature to the priority of balancing choice probabilities and optimal design methods continuous (Louviere, 2010; Rose, 2014), Devarassety et al. (2012) show that Bayesian efficient D-error designs are up to date the most efficient for VoT models in a transportation experiment, testing a Bayesian efficient design compared to a random level design by optimizing the attribute levels to get balanced probabilities and an adaptive random designs using the results from previous choice tasks of the individual as priors for balancing the following choice probabilities.

2.3.4 Optimal sample size

Bliemer and Rose (Bliemer, 2005; Rose, 2005) show that the AVC is inversely related to the number of times one choice task is repeated in the sample N_R and therefore to the number of respondents N. As an AVC can be obtained for any level of N by calculating the AVC for N equal to one and then dividing the resulting matrix by N, this means that the standard errors decrease by $1/\sqrt{N}$. Therefore the impact of one more observation decreases with sample size and as a consequence often better results can be achieved by increasing the efficiency of the design than by only increasing the number of respondents.

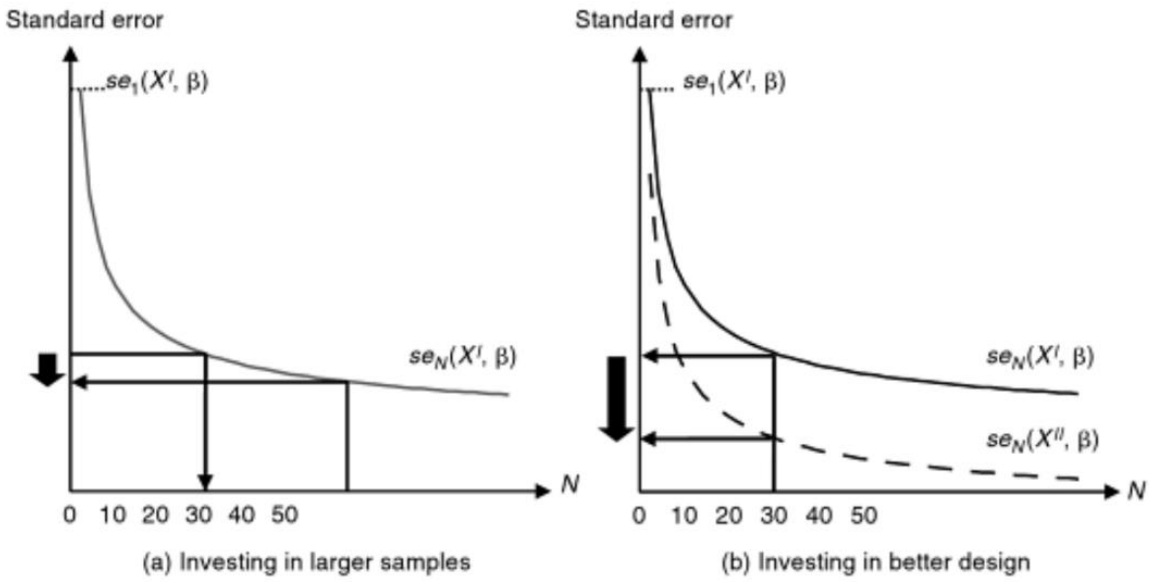


Figure 3: Comparison of investing in a bigger sample size versus more efficient design (Rose, 2014)

Using the relationship between the statistical efficiency measured in the t-statistic and the number of respondents, Rose and Bliemer than showed that the minimum necessary number of observations to obtain statistical relevant results for each parameter can be calculated using the priors:

$$t_k = \frac{\beta}{\sigma/\sqrt{N}} \rightarrow N = \frac{t_k^2 \sigma^2}{\beta^2} \quad .$$

The authors note that the resulting minimum sample size should be taken as an absolute theoretical minimum as this method assumes certain asymptotic properties that might not hold in small samples and does not consider the stability of the parameter estimates or sample size at which stability is obtained. Designs optimized to reduce the necessary sample size for expecting statistically relevant results in all parameters are called S-efficient designs (Rose, 2014).

2.4 Measurements of fit for discrete choice models

In ordinary least square regression (OLS) the ratio of explained deviation over total deviation from the sample mean is taken as the measurement of explanatory value for a model and denoted as R^2:

$$R^2 = \frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \quad ,$$

where y_i is the actual observed outcome, $\hat{y}_i$ the predicted outcome and $\bar{y}$ the average outcome. As the MNL can be seen as a simultaneously estimation of multiple binary logits and identification therefore similarly depends on the assumption concerning the error term, there is no straight forward assessment of the goodness of fit as in standard OLS possible. For binary logit and probit a number of pseudo-R^2 have been proposed, trying to access the share of explained deviations. Elfron (1978) for example defines $\hat{y}_i$ as $\hat{\pi}_i = \Pr(y_i|x_i)$ and states

$$R^2_{El} = \frac{\sum_{i=1}^{N}(y_i - \hat{\pi}_i)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2} \quad .$$

That is McFadden (1974) takes a similar analogy, but concentrates on the log-likelihood of the model with explanatories (= explained deviation) $\hat{L}(M_\beta)$ over the model without explanatories (= total deviation) $\hat{L}(M_\alpha)$:

$$R^2_{McF} = 1 - \frac{\ln \hat{L}(M_\beta)}{\ln \hat{L}(M_\alpha)} \quad .$$

Similar to the adjusted R^2 in OLS, Ben-Akiva and Lerman (1985) added an adjustment for the number of explanatories K to the measure:

$$R^2_{aMcF} = 1 - \frac{\ln \hat{L}(M_\beta) - K}{\ln \hat{L}(M_\alpha)} \quad .$$

Further adoptions were made by Maddala (1983)

$$R^2_{ML} = 1 - \left[\frac{\hat{L}(M_\beta)}{\hat{L}(M_\alpha)}\right]^{2/N}$$

and Cragg and Uhler (1970), who normed R^2_{ML} to be between zero and one:

$$R^2_{nML} = \frac{1 - \left[\frac{\hat{L}(M_\beta)}{\hat{L}(M_\alpha)}\right]^{2/N}}{1 - \left[\hat{L}(M_\alpha)\right]^{2/N}} \quad .$$

While the adjusted McFadden R^2 is probably the most used Goodness of Fit indicator (GoF) for MNL and is also included in most standard statistical software packages, Hagle and Michell (1992) and Windmeijer (1995) showed in simulation studies, that there are better fitting indicators like McKelvey and Zavoina's Ratio of Variance (1975):

$$R^2_{M\&Z} = \frac{\widehat{Var}\,(\widehat{y^*})}{\widehat{Var}\,(y^*)} = \frac{\widehat{Var}\,(\widehat{y^*})}{\widehat{Var}\,(\widehat{y^*}) + Var\,(\varepsilon)} \quad .$$

McKelvey and Zavoina's measure takes advantage of the fact that the variance of the error term in binary and multinomial choice models is fixed by the researchers choice of model (i.e. $Var\,(\varepsilon) = \frac{\pi^2}{3}$

for logit), therefore only leaving the predicted variance of the latent variable, which depends only on the predicted coefficients and the variance of the explanatories: $\widehat{Var}\ (\widehat{y^*}) = \hat{\beta}'\widehat{Var}\ (x)\hat{\beta}$

Instead of focusing on the share of explained deviation a second stream of literature started to look at the predictive value of the included covariates. Maddala (1992) proposed to measure the number of right predictions n_{jj} of the outcome over the number of total observations:

$$R^2_{Count} = \frac{1}{N}\sum_j n_{jj} \quad .$$

As this measure can overstate the predictive value of the model, if there is one very dominant alternative by simply always predicting that alternative, the R^2_{Count} should be adjusted by subtracting the percentage of the most dominant option (Bishop et al. (1975)):

$$R^2_{aCount} = \frac{\sum_j n_{jj} - \max_r(n_{r+})}{N - \max_r(n_{r+})} \quad .$$

There is some discussion to these measurements as some researchers argue, that the aim of DCE's is to predict the probability of alternatives and by reducing the outcome to the alternative with the highest probability, information is lost. If for example $\Pr(\widehat{alt_1}|x_\iota) = 51\%$ and $\Pr(\widehat{alt_2}|x_\iota) = 49\%$ than the R^2_{Count} should maximal reach 51% as it will predict always the first alternative with the higher probability (Train, 2009). To account for this caveat the predicted R^2_{Count} was developed, which does not define $n_{jj} = 1$ if actual alternative and the alternative with the highest predicted probability are equal but splits a unit interval by the probabilities of all alternatives and declares $n_{jj} = 1$ if a draw from a uniform distribution between zero and one will fall into the area of the actually chosen alternative.

Next to the explained deviation or predictive value, Akaike (1973) promoted one further point of origin for GoF indicators by looking at the information deducted from the model:

$$AIC = \frac{-2\ln\hat{L}(M_\beta) + 2P}{N} \quad ,$$

where P equals the number of parameters in the model. As not normed this measure can only be used to compare different models, where a smaller AIC means a better fit.

All these measures-of-fit were developed for binary choice models and then adapted to more complex designs. While there is use of these measures in MXL, there is few work on the theoretical application for these measures yet done.

3. Market Entry Study Car2Go

3.1 Car2Go and public transportation in survey city

The concept of Car2Go was first tested in 2008 as a project of the automobile manufacturer Daimler AG. Instead of owning a private car, carsharing (CS) proposes to use cars out of a centrally owned fleet on demand. In difference to classical rental companies CS focuses on urban travel and offers often also short term or hourly rentals and 24 hour self-service pickups. While the idea of sharing a car can be traced back to the Swiss "Selbstfahrergenossenschaft" in 1948 and there already existed many CS or rental companies in different cities, C2G added an innovative further element to the mix as cars could be rented and parked everywhere in the city region. Instead of going to a fixed spot to get a car, the fleet is distributed throughout the city and a user finds his car by locally searching with his smartphone for the closest match (Wikipedia, 2015a; Wikipedia, 2015b). This instant availability changed the CS market as it competes even closer with short distance urban public transportation then with long distance transportation modes and increased flexibility and usability greatly. Consequently it was fast copied by a row of other CS companies like for example BMW's DriveNow, the Deutsche Bundesbahn subsidiary Flinkster and many more (Wikipedia, 2015a). In beginning of 2015 C2G has been introduced in 30 European and American cities and is with over one million customers and round about 12.500 Smarts whereof 1.200 are electric vehicles the biggest car sharing provider worldwide (Wikipedia, 2015c; Car2Go, 2015).

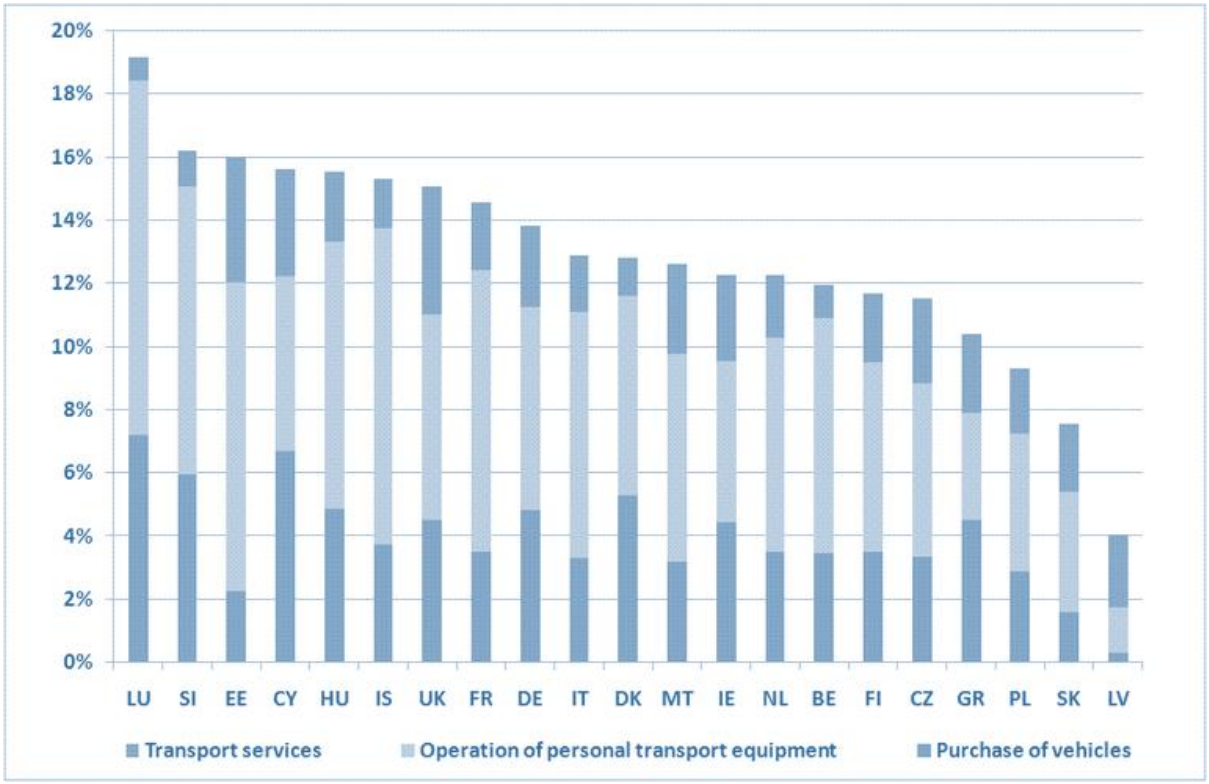


Figure 4: Average household expenditures in Europe on transportation (2008 in percent of total expenditures, European Environment Agency, 2012)

The city under study is a typical central European metropolis with a little more than one million inhabitants and a high population density in the city's downtown area. The city has one of the highest GDP per capita in Europe with above 60.000€ in 2011, which is boosted by a high rate of commuters to neighboring regions (about half the workforce or 350.000 commuters) but still leaves a disposable average income of round about 17.500€ per inhabitant. The average household spent roughly 3.000€

(10% of total expenditures) in 2011 on transportation, which is below the European average as seen in Figure 4.

The city's public transportation infrastructure is based on a network of bus, tram and metro services with round about 2 000 stops in total and thereof 50 to 100 metro stations, which provides about 350 million trips in 2014. Further the city registers round about 500 000 private cars and there is a fast row of taxi service companies. There are already two established classical car sharing companies operating in the city area with almost 10.000 subscribers and further a bike sharing company with 30.000 users exists, so mobility sharing concepts are already well known. [3]

3.2 Survey Design

3.2.1 Alternatives, Attributes and Attribute Levels

Based on the existing public transportation infrastructure the DCE and the expected primary substitution pattern for C2G, the DCE included four labeled alternatives bus, metro, taxi and C2G. To also capture further preferences and obtain realistic results a "none of the described" option (from now on labeled as "none") was added. To assess the effect of different activities on the VoT the total travel time was separated into access, in-mode and egress time. The access time was converted into walking distance using a factor of 1.2 minutes per 100m equaling a walking speed of 1.4m/s (Browning, 2005; Mohler, 2007; Levine, 1999) and the fixed parameter levels were set to between zero meter and one kilometer. For taxi the parameter level was fixed to zero meters to account for pick-up service. As the primary focus of the study was to access urban travel modes the in-mode travel time was picked to match typical intracity travel distances with 5, 15 and 30 minutes per trip. For the two public transport modes a transition or indirection top up between zero and 30 minutes was added. As using the same routes C2G and taxi were simple to the same in-mode time per choice situation and for bus the minimum time was set to match. Similar the metro in-mode time was simple to vary maximal 40 minutes from the C2G parameters. The egress time was added to account for either the search for a parking space with C2G (max. 5 Min.) or the walking time from the next station to the destination for bus and metro (max. 10 Min.). Again no time was added for taxi as delivery directly to the destination was assumed.

The pricing was adopted to the regional parameters for bus, metro and taxi. The prices for bus and metro were linked and set to 2.00€ or to account for users of an abo 0.00€. For taxi the price includes the three components: Pick-up fee 2.40€ (4.40€ at night), 1.80€/km for driving times and 0.50€/Min. for waiting times. Using the average traveling speed at 13:00 (equaling the base scenario) for some typical origin-destination combinations within the analyzed city as stated by google maps (2015), an average price of 0.8€/Min. was then used. For C2G the base price was set to 0.30€/Min. matching the current pricing in other European cities. From this starting point four levels were introduced by adding or subtracting one third or one sixth of the base price, leaving the five price levels 0.20, 0.25, 0.30, 0.35 and 0.40 €/Min. multiplied by in-mode and egress time.

To further distinguish VoT three situational variables were added, which take the same value over all alternatives per choice task: (1) to account for weather influences a dummy coded variable equaling one for rain and zero for sun, (2) for the point of origin another dummy equaling one for being at home or zero for being in the city and (3) to assess the influences of the time of day an effect coded level variable for morning (08:00), midday (13:00) or evening (18:00). For the city scenario no pick-up charge was included in the taxi price, while for home and evening the night charge was added. The morning

[3] All figures are taken from the official Regional Institute for Statistics & Analysis for 2013 and the 2014 report of the main Regional Public Transportation Provider.

scenario was linked to a rush hour situation adding an extra 10 Min. to the 5 Min. in-mode travel time of bus, C2G and taxi and 30 Min. to the 15 and 30 Min. distances. As the additional time represents waiting time it was added at 0.50€/Min. to the taxi price. While this setup insured a high external validity as it is very realistic, it also introduced a correlation between the morning dummy with travel time and therefore also price. A further caveat of this design choice were a high number of restrictions, which reduced the possible choice set combinations and thereby the statistical efficiency of the design as discussed in 2.3.3 shows a summary of the different attributes, levels and situational variables entering the design.

Travel Time							
Access***		in m	0	150	300	500	1000
		in Min.	0	2	4	6	12
In-Mode		in Min.	5	15	30		
Adder Public Transport*		in Min.	0 or 10	0 or 30	0 or 30		
Adder Rushhour / Morning**		in Min.	10	30	30		
Egress***		in Min.	0	5	10*		
Price							
Bus		in €/trip	0.00€ (Abo)	2.00 €			
Metro		in €/trip	0.00€ (Abo)	2.00 €			
Taxi	Pick-Up	in €/trip		2.40 € at day		4.40 € at night	
	Driving	in €/km	1.80 €	0.80€/Min x In-Mode		0.50€/Min x Adder Rushhour	
	Waiting	in €/Min.	0.50 €				
C2G		in €/Min.	0.20 €/Min	0.25€/Min	0.30€/Min	0.35€/Min	0.40€/Min
Situation							
Time of Day		Morning (8:00)	Midday (13:00)	Night (22:00)			
Weather		Sun	Rain				
Origin		Home	City				

* only bus and metro | ** only bus, taxi and C2G | *** only bus, metro and C2G

Figure 5: Summary of alternative attributes and levels

3.2.2 Design and Model

As the aim of the study was to gain an insight into the base utility of different travel modes, the effect of access, in-mode and egress travel time and pricing conditional on certain parameter variables. The basic model can be separated into three parts: (1) the intercepts per alternative compared to the base alternative "none", (2) the main effects of access, in-mode and egress travel time plus price and (3) the interactions between these variables and the situational variables. While a conditional logit should not be able to estimate coefficients for variables, which are fixed over alternatives, the situation variables were interacted also with the intercepts of the alternatives to capture their base effects against the "none"-alternative. Equation 2 shows the complete model:

$$
\begin{aligned}
U_{nit}(X_{nit}) = {} & \beta_{11}BUS_{nit} + \beta_{12}METRO_{nit} + \beta_{13}TAXI_{nit} + \beta_{14}C2G_{nit} \qquad (2) \\
& + (\beta_{111}CITY_{nt} + \beta_{112}RAIN_{nt} + \beta_{113}MORNING_{nt} + \beta_{114}NIGHT_{nt})\, BUS_{nit} \\
& + (\beta_{121}CITY_{nt} + \beta_{122}RAIN_{nt} + \beta_{123}MORNING_{nt} + \beta_{124}Night_{nt})\, metro_{nit} \\
& + (\beta_{131}CITY_{nt} + \beta_{132}RAIN_{nt} + \beta_{133}MORNING_{nt} + \beta_{134}Night_{nt})\, taxi_{nit} \\
& + (\beta_{141}CITY_{nt} + \beta_{142}RAIN_{nt} + \beta_{143}MORNING_{nt} + \beta_{144}Night_{nt})\, C2G_{nit} \\
& + \beta_2 ACCESS_{nit} + \beta_{21}ACCESS_{nit}CITY_{nt} + \beta_{22}ACCESS_{nit}RAIN_{nt} \\
& \qquad + \beta_{23}ACCESS_{nit}MORNING_{nt} + \beta_{24}ACCESS_{nit}NIGHT_{nt} \\
& + \beta_3 INMODE_{nit} + \beta_{31}INMODE_{nit}CITY_{nt} + \beta_{32}INMODE_{nit}RAIN_{nt} \\
& \qquad + \beta_{33}INMODE_{nit}MORNING_{nt} + \beta_{34}INMODE_{nit}NIGHT_{nt} \\
& + \beta_4 EGRESS_{nit} + \beta_{41}EGRESS_{nit}CITY_{nt} + \beta_{42}EGRESS_{nit}RAIN_{nt} \\
& \qquad + \beta_{43}EGRESS_{nit}MORNING_{nt} + \beta_{44}EGRESS_{nit}NIGHT_{nt} \\
& + \beta_5 PRICE_{nit} + \beta_{51}PRICE_{nit}CITY_{nt} + \beta_{52}PRICE_{nit}RAIN_{nt} \\
& \qquad + \beta_{53}PRICE_{nit}MORNING_{nt} + \beta_{54}PRICE_{nit}NIGHT_{nt} \\
& + \varepsilon_{nit} \quad ,
\end{aligned}
$$

where BUS_{nit}, $METRO_{nit}$, $TAXI_{nit}$ and $C2G_{nit}$ are dummy coded variables equaling one if the alternative corresponds to this mode, $CITY_{nt}$, $RAIN_{nt}$, $MORNING_{nt}$ and $NIGHT_{nt}$ are dummy variables signaling the scenario if set to one and $ACCESS_{nit}$, $INMODE_{nit}$, $EGRESS_{nit}$ and $PRICE_{nit}$ are the attribute variables for each alternative. Based on the size of the survey population and the number of parameters within the model, the number of different choice situations in the experiment design was set to 60 and parted into three blocks. The design will be replicated 40 times generating 2.400 rows. This corresponds to a sample of 800 respondents as each individual faces three different choice cards. While also interested in extend the model to a MXL with travel time and price as random parameters and linked to demographics, the main focus was first to get stable results for a MNL prediction including the scenario variables. Therefore the design was developed for MNL and not for MXL or a mixture of both models as proposed by Rose et al. (2009). As optimization method the D-error minimization was used and as the design included constraints the default swaping algorithmus was utilized (ChoiceMetric, 2014). Based on the conclusion Devarassety et al. (2012) as presented in 2.3.3 a Bayesian efficient design was generated using NGENE 1.1.2, a designated software developed by Rose, Collins, Bliemer and Hensher to create SCE designs. Appendix 1 includes the code and Figure 3 shows the mean of the priors used in the designing process. The distribution was assumed to be normal with the standard error being equal to half the mean: $\beta_j \sim N\left(\mu_j, \frac{\mu_j^2}{4}\right)$.

Intercept	bus		1.45	Interactions	city	0.5	rain	0.50	morning	-0.50	night	-0.20
	metro		1.45			0.5		0.80		-0.20		-0.12
	taxi		1.40			0.5		1.00		-0.50		0.30
	C2G		1.30			0.5		0.80		-1.00		-0.30
Main Effects	access	(in 100m)	-0.10	Interactions	city	0.05	rain	-0.20	morning	-0.05	night	-0.10
	in-mode	(in Min.)	-0.10			0.01		-0.005		-0.05		-0.05
	egress	(in Min.)	-0.10			0.01		-0.20		-0.05		-0.10
	price	(in EUR)	-0.20			0.01		0.05		-0.01		0.05

Figure 6: Priors used in designing process

The priors were set using previous market studies of C2G and expert approximations by the pricing team of C2G. As some of the interactions were seen as not relevant, there are some very low priors for example for the interaction between in-mode travel time and rain. This and also the fact that the design suffers from correlation between morning and price / in-mode travel time increases the minimum necessary observations to retrieve all parameters as defined in the S-statistic (see also chapter 2.3.4, Bliemer, 2005; Rose, 2005), which can be seen in Figure 7 per parameter and in maximum would be 26.798 observations of all rows to retrieve the interaction between price and morning. Appendix 4 includes an overview of the finally used choice situations and parameter levels.

D-Error		bus		Interactions	city	38	rain	117	morning	163	night	1,025
0.040943		metro				38		37		1,541		2,992
		taxi				38		18		1,104		371
		C2G				38		19		19		232
Main Effects	access		9	Interactions	city	24	rain	9	morning	164	night	64
	in-mode		7			191		1,481		100		36
	egress		23			5,496		10		256		83
	price		20			288		826		26,798		66

Figure 7: D- and S-Statistic for final design

At the same time the D-Error is with 0.040943 still quite large in relation to other designs based on the same parameters. Both effects are due to the high number of restriction and total number of parameters, as the design without restrictions can easily be optimized to half the D-Error size and with only the main effects without the scenarios also to 1/10 of the D-Error and less.

To further check the validity of the generated design, the best of each iteration was also simulated using 1000 iterations of Monte Carlo Simulation with the above priors and distributions and checking these with a number of measures of fit and comparing the estimated coefficients to the “true” coefficients. In this stage also MXL models were considered and similar performance for MNL and MXL found. (See Appendix 3 for detailed results and Appendix 2 for the code base)

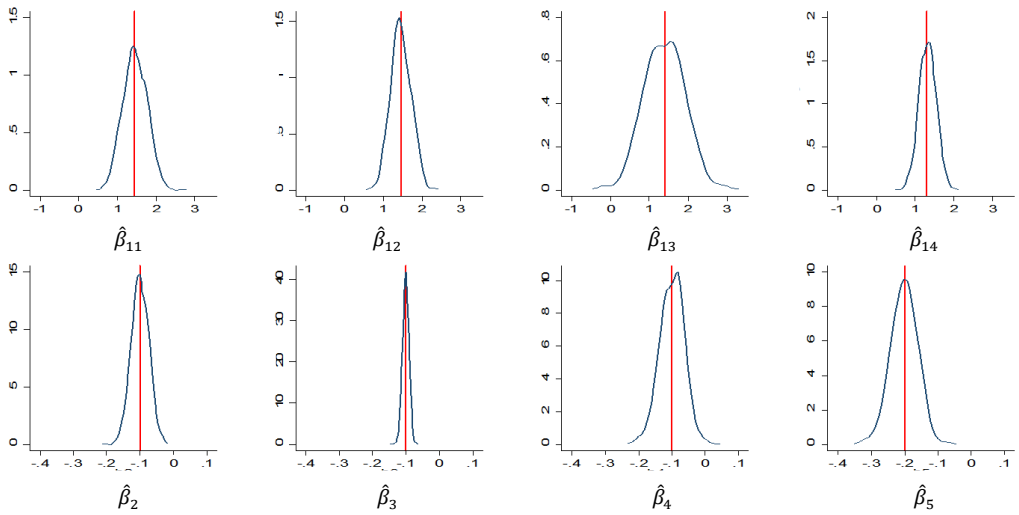


Figure 8: Spread of estimated main coefficients in Monte Carlo simulation exercise for dedicated design and 1.000 draws

As seen in Figure 8 the model is able to retrieve all main effects within a very narrow spread. Similar the intercepts of *BUS*, *METRO* and *C2G* are close to the real values. For *TAXI* the spread is slightly wider, but the mean of the simulated estimates still converts to the true coefficient. While only dependent on the priors Figure 9 shows the GoF indicators for the simulated models (the solid red line equals the mean and the dotted red line show the standard deviations), which range from 45% for the adjusted McFadden measure up to 60% for the Ratio of Variances. Due to a high number of *NONE*-predictions the adjusted count R^2 is relatively low and varies strongly between iterations. The MXL (grey line) performs only slightly less well based on the indicators.

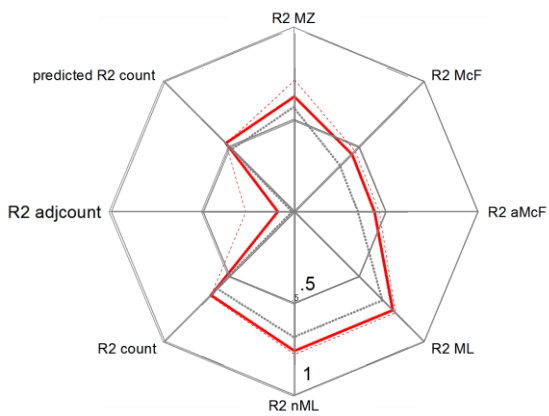


Figure 9: Measures of fit for simulation exercise

3.3 Data collection and Basic statistics

3.3.1 Survey Design and Population

The choice cards were structured like shown in Figure 10 and stated the five alternatives, attribute values per alternative and the choice situation. The situation was further described using small icons to help comprehension. The design was tested upfront with the help of round about 40 students at the University of the Basque Country and enhanced to give a fast and aggregated overview over the choice situation. In difference to the design as described in 3.2.2 the egress time was not stated separated but included with the in-mode and access time into the variable *TRAVELTIME*. The access time was stated as "thereof" in meters and minutes. As egress and in-mode coefficients shared the same priors in the designing process the design should be robust to this simplification and a second simulation experiment estimating only one coefficient for egress and in-mode travel time seemed to confirm this hypothesis. For the morning scenario also an icon for rush hour was added to explain the longer travel times.

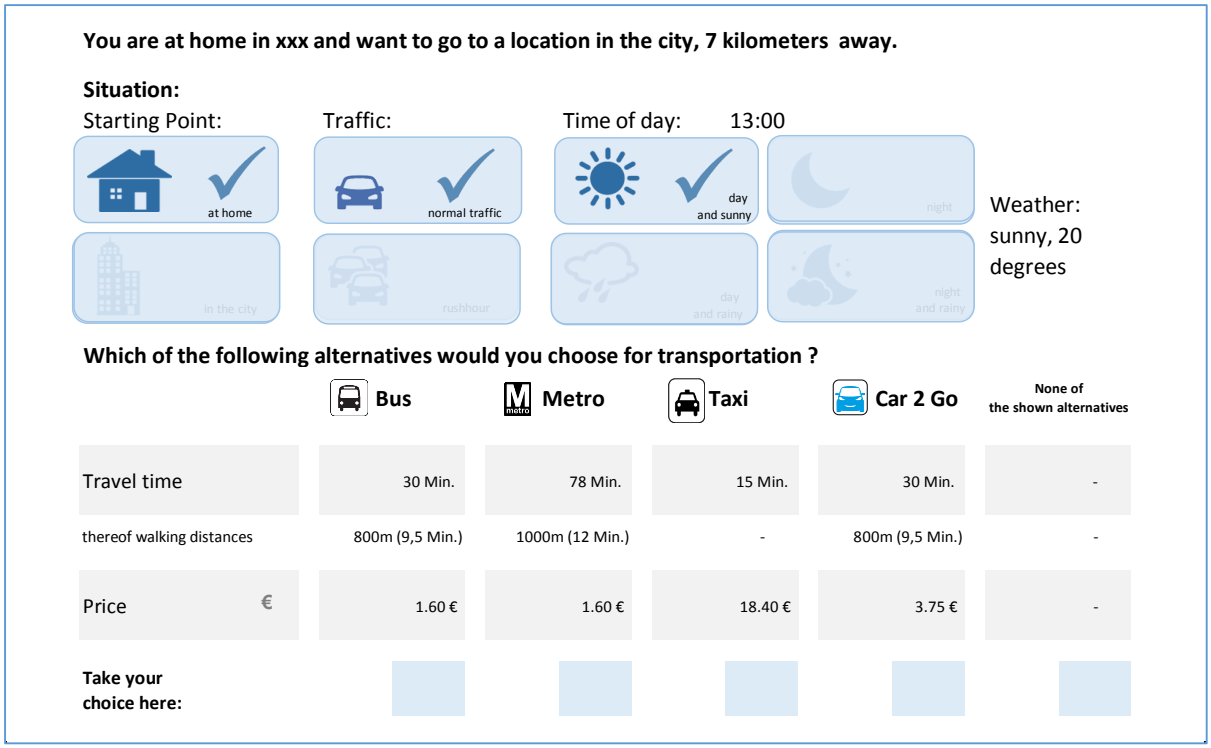


Figure 10: Example Choice Card

The survey was collected using computer-assisted personal interviews in a laboratory environment. The questionnaire included 32 questions and was answered completely by each of the 800 respondents. Each individual was first asked a number of questions concerning their demographics, opinion and usage of different transport modes and finally the three choice cards. The sample population is representative for the city's population between 18 and 60 with a driving license and choice cards were randomly assigned therefore no negative impact on causal inference should obscure results due to the survey design.

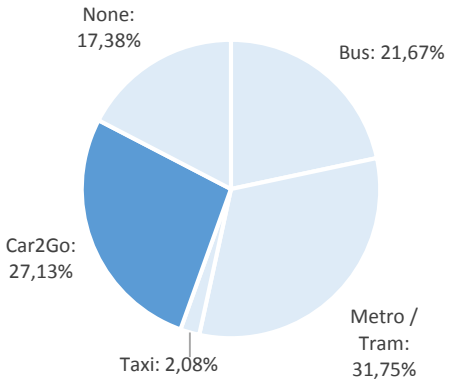


Figure 11: Distribution of choices

As seen in Figure 11 the alternative *METRO* was chosen most often with 31.75% followed by *C2G* with 27.13% and *BUS* with 21.67%. *TAXI* only played a minor role in the DCE as it was only 50 times taken. The *NONE*-option was chosen 417 times or in 17.38% of all cases.

On average each respondent chose 2.06 (standard error 0.76) different transport modes in his or her three choice tasks and only 25.75% choose three times the same mode. This behavior was most common for respondents choosing the *NONE*-option with in total 66 respondents. This could be interpreted as refusal, but also be explained by actual preferences for not stated alternatives like car or bike. Similar those choosing three times *C2G* (51 respondents) could be seen as compliers. As the reasons for this behavior could also be rational both groups were after discussion with C2G not excluded from the analysis but several analysis with and without those respondents were run.

The average number of different chosen transport modes per choice situation were 4.35 (standard error 0.65) out of five and only in one choice situation only two alternatives were chosen. Mostly missing in half of all choice situations was the taxi alternative matching the low overall number of *TAXI* choices.

3.3.2 Individual characteristics

Figure 16 gives an overview of demographics and stated opinion of the surveyed population. The average age of the respondents is 38.5 years with a standard deviation of 11 years. Figure 12 shows a relative equal distribution among age categories and the differences among choices in the DCE for each group. There is a significant decrease in the probability of choosing *C2G* and *BUS* for older age group, while the *NONE* option increases with age. Both gender are relatively equal represented in the survey with a slightly higher percentage of male (50.48%) than female (49.52%). There are no big differences in choices between the genders with male choosing *TAXI* a little bit more often (3.27 versus 0.91%) and *METRO* / *C2G* slightly less often (19.96% versus 23.33% and 25.86% versus 28.37%). Interestingly but to be expected the percentage of choices of *C2G* are with 28.64% much higher among those respondents having a smartphone (87.88%) compared to 16.15% for those without. There is no difference in percentage between iOS (41.54%) and Android (42.25%) phones with 29.34% and 28.62% for *C2G*, but for Windows phones (10.53%) the percentage is lower at 22.52%.

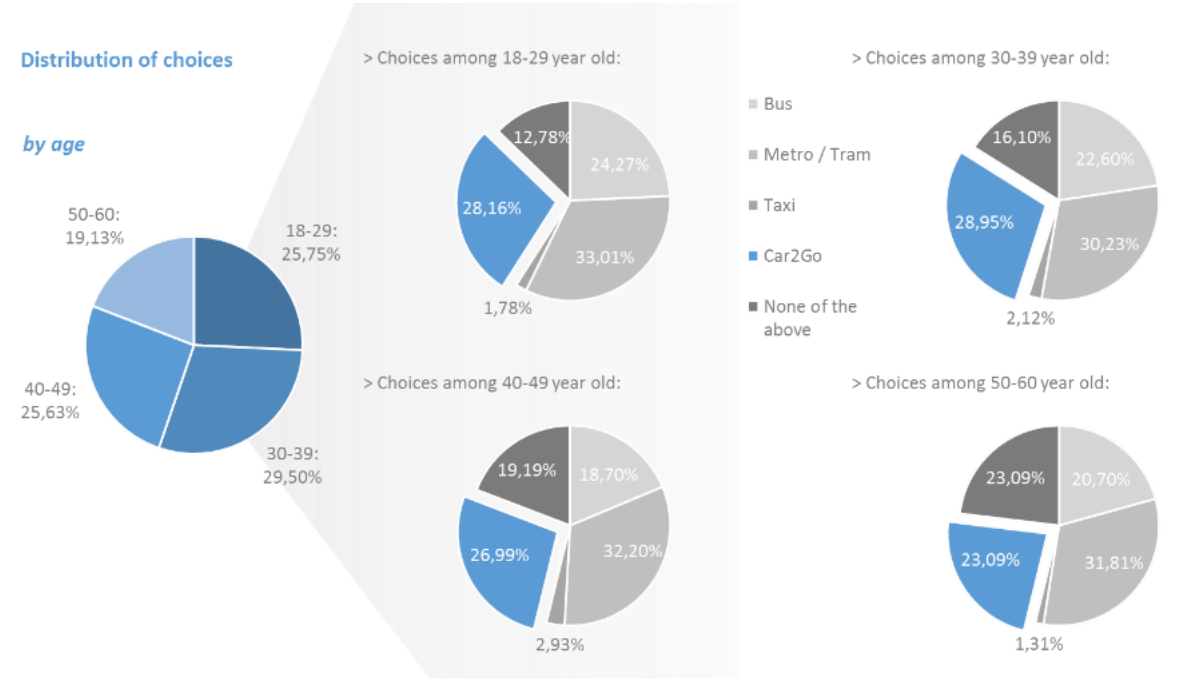


Figure 12: Distribution of choices among age groups

Looking at the income distribution in Figure 13 most of the respondents who answered this question (66.12%) earn between 1.000 and 1.999€ per month. While lowest in this income class the probability of choosing *C2G* increases with income and is with 47% highest for incomes above 3.000€ per month.

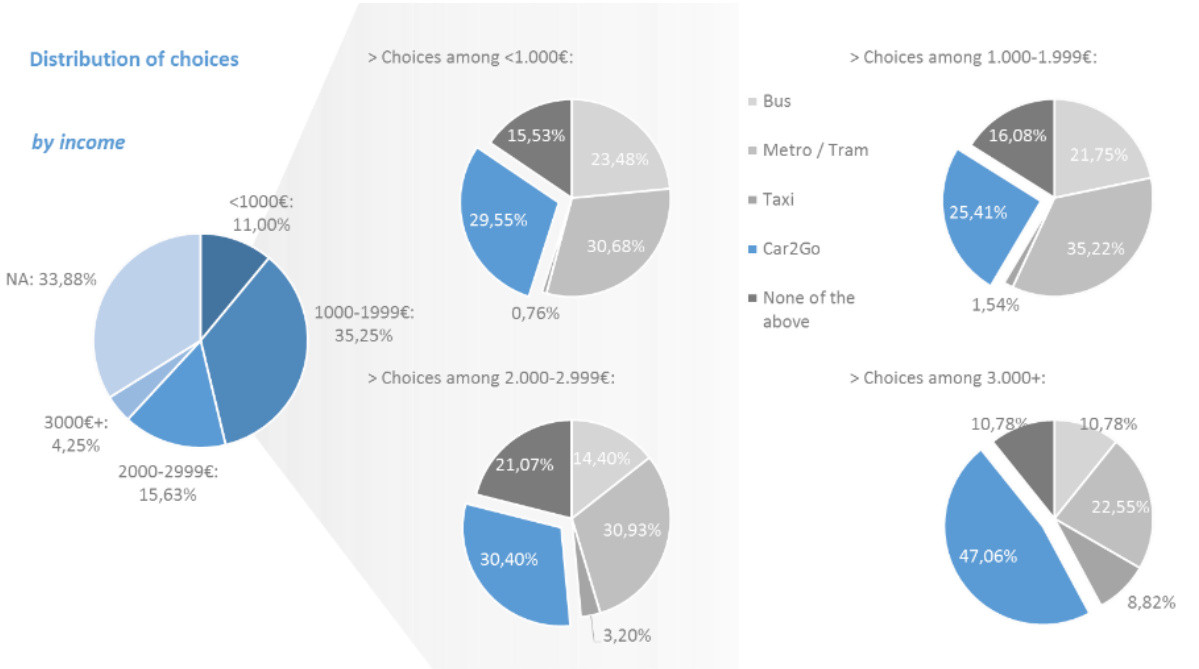


Figure 13: Distribution of choices among income groups

The average number of persons per household is 2.75 (Standard deviation 1.13) and there are 1.79 cars per household in the mean with a standard error of 1.29. Around 39.75% of the surveyed population state that they use their own car seven days a week and another 16% fall into the four to six days and 19.25% into the two to three days category. Only 20.38% never use a privately owned car. Of all individuals 17.12% use their bike at least once per month, but only 7.62% more than once a

week. 25.38% of the population walk each day and 68.50% at least once a week somewhere. Of the respondents 16.75% use the bus more than four days a week, but 48.75% do not use buses at all. Similar 16.63% use the tram to commute, but 51.88% use the tram not even once a month. The metro is more popular with 21.75% using it regularly more than four times while only 37.38% do not the metro ever. Combining these three alternatives 27.13% of the respondents use public transportation at least four days a week, 52.38% at least once a week and 32.38% do not use public transportation. Only 2.88% of the respondents use in contrast taxies at least once a week, while 90.25% do not use taxi services ever. Similar only 4.5% of the survey population use car sharing services at least once a week, while 88.88% state that they never use car sharing.

Interestingly this does not translate into a low interest in the concept of C2G. A high number of 26.75% of the respondents state that they are very interested in the concept and another 43.00% are at least partially in agreement that C2G represents urban lifestyle, solves parking problems and should be implemented in the city. Only 3.38% do not agree with these statements. As to be expected the percentage of choices of *C2G* increases drastically with agreement to the concept and 40.19% of the respondents who agree strongly chose at least once *C2G* in the DCE while with decreasing agreement the *NONE*-option becomes more likely.

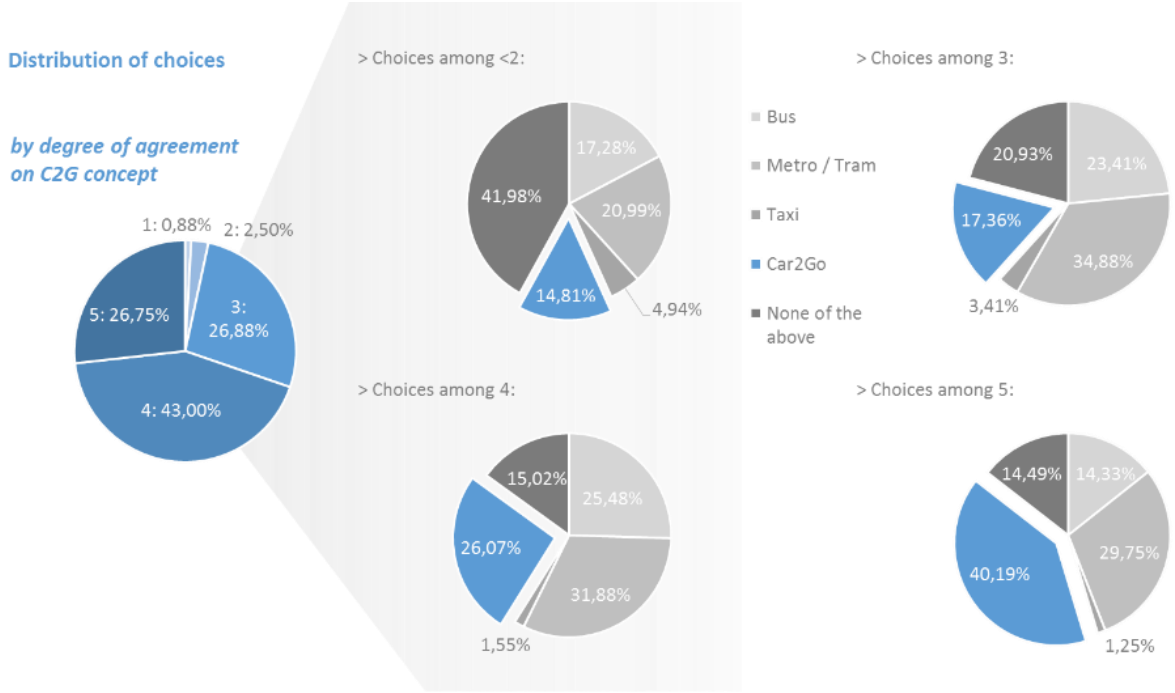


Figure 14: Distribution of choices among interest in C2G concept

Asked for the likeliness to use C2G, if the concept should be implemented, the agreement is lower and only 9.50% are sure to use the concept and 20.75% think they probably would use the concept, 47.38% do not think they will use C2G and the remaining 22.38% are undecided. Interestingly 15.15% of the sure not to use and 19.70% of the probably not to use C2G respondents still choose *C2G* at least once in the DCE. Again the increase in level of stated usage also increases the likeliness to observe at least one *C2G* choice for an individual in the DCE.

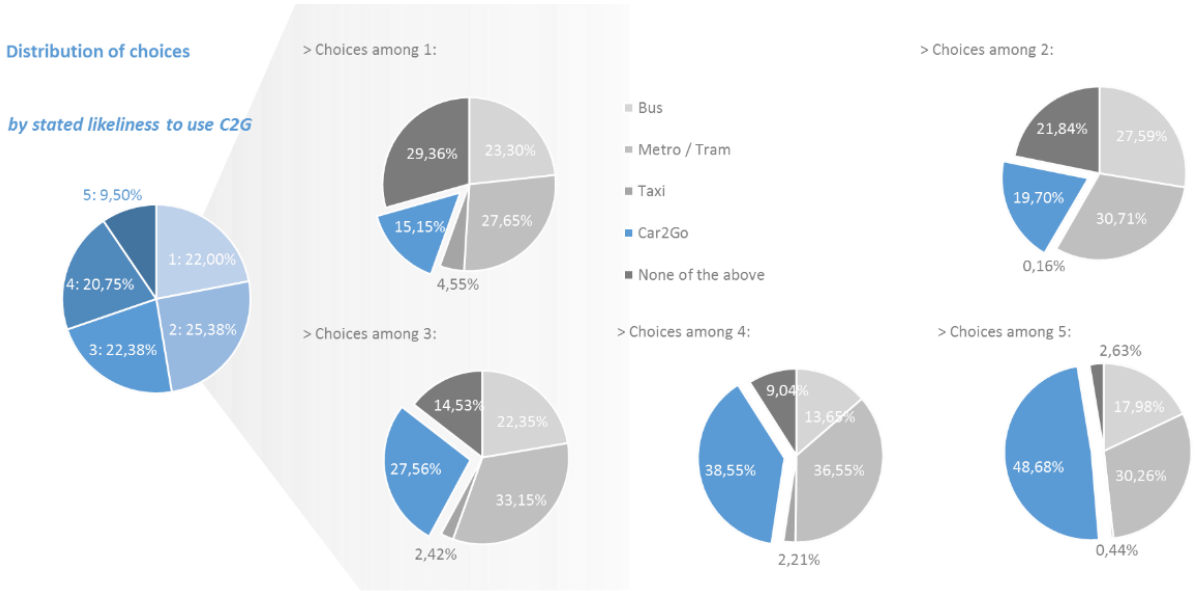


Figure 15: Distribution of choices among stated likeliness to use C2G concept

Variable	Type	Mean	Std. Dev.	Min	Max
AGE	in years	38.4675	11.1338	18.00	60.00
GENDER	Dummy with "male" = 1	0.4963	0.5003	0.00	1.00
HOUSEHOLD SIZE	in individuals	2.7513	1.1296	1.00	4.00
INCOME	Level coded				
	< 1 000 €	0.1100	0.3131	0.00	1.00
	1 000 to 1 999 €	0.3525	0.4778	0.00	1.00
	2 000 to 2 999 €	0.1563	0.3631	0.00	1.00
	3 000 to 3 999 €	0.0250	0.1561	0.00	1.00
	4 000 to 4 999 €	0.0125	0.1111	0.00	1.00
	> 5 000 €	0.0050	0.0705	0.00	1.00
	No Answer	0.3388	0.4733	0.00	1.00
SMARTPHONE	Dummy with "owning" = 1	0.8788	0.3266	0.00	1.00
CARS	in number per household	1.7938	1.2926	1.00	5.00
AGREEMENT	in % per level				
Average Agreement on C2G over 11 questions	Highly positive	0.4009	0.3559	0.00	1.00
	Positive	0.2523	0.2542	0.00	1.00
	Medium	0.2391	0.2701	0.00	1.00
	Low	0.0606	0.1179	0.00	1.00
	Very low	0.0472	0.1316	0.00	1.00
LIKELINESS	Level coded				
Stated likeliness to use C2G	1: No use	0.0950	0.2934	0.00	1.00
	2: Probably no use	0.2075	0.4058	0.00	1.00
	3: Uncertain	0.2238	0.4170	0.00	1.00
	4: Probably use	0.2538	0.4354	0.00	1.00
	5: Use	0.2200	0.4145	0.00	1.00
KNOWLEDGE	Dummy with "know" = 1	0.1050	0.5517	0.00	1.00
PRIOR USE OF CS	Dummy with "yes" = 1	0.0175	0.1312	0.00	1.00
PREFERENCE	Dummy per mode				
	Public Transport to CS	0.0700	0.2553	0.00	1.00
	Taxi to CS	0.0075	0.0863	0.00	1.00
	Other to CS	0.3575	0.4796	0.00	1.00
	No preference for CS	0.2275	0.4195	0.00	1.00
OPINION CARS	in % per level				
Average over 14 questions	Highly positive	0.2819	0.1610	0.00	0.71
	Positive	0.1473	0.1347	0.00	0.79
	Medium	0.1805	0.1720	0.00	1.00
	Low	0.1580	0.1351	0.00	0.64
	Very low	0.2322	0.1712	0.00	0.71
OPINION CS	Level coded				
	Highly positive	0.0175	0.1312	0.00	1.00
	Positive	0.2725	0.4455	0.00	1.00
	Medium	0.3800	0.4857	0.00	1.00
	Low	0.2563	0.4368	0.00	1.00
	Very low	0.0738	0.2615	0.00	1.00

Figure 16: Summary of demographics and stated opinion

3.3.3 Alternative attributes

Keeping in mind that the access time only applies for the alternatives *BUS*, *METRO* and *C2G* 42.5% of all alternatives had zero meters or minutes access time. Thereof 20% are the *TAXI* alternatives, the *NONE*-option is excluded from this and the following comparisons. The rest of the distances is relatively similar distributed with a slight lower number of alternatives with the maximum distance 1.000m. The

distribution of *DISTANCE* among the chosen alternatives is very close to the original distribution, therefore issuing the question of the relevance of access time within the DCE. 52.04% of all times the closest option is taken, but in 14.73% of the cases alternatives with 500 to 1.000m more walking distance compared to the closest alternative were still chosen. Similar the difference in distribution between overall in-mode time and chosen in-mode time are less different than expected, but there is a slightly higher number of 5 and 20 minute chosen alternatives (63.49%) compared to the general distribution (54.58%) and in 74.03% the fastest alternative was taken.

For *PRICE* there is as expected also a higher chance of lower prices per minute in the chosen alternatives than compared to the complete set of choice situations. For example 20.57% of the chosen alternatives had a price of zero and 28.34% between zero and ten cents while only 16.67% of all alternatives had a price of zero and 24.58% a price between zero and ten cents. As *C2G* only starts at prices of 0.20€/min. and *TAXI* at 0.80€/min. the first three categories combined belong to *BUS* and *METRO*, which share the same price, and therefore 62.32% of choices can also be interpreted as taking the lowest price.

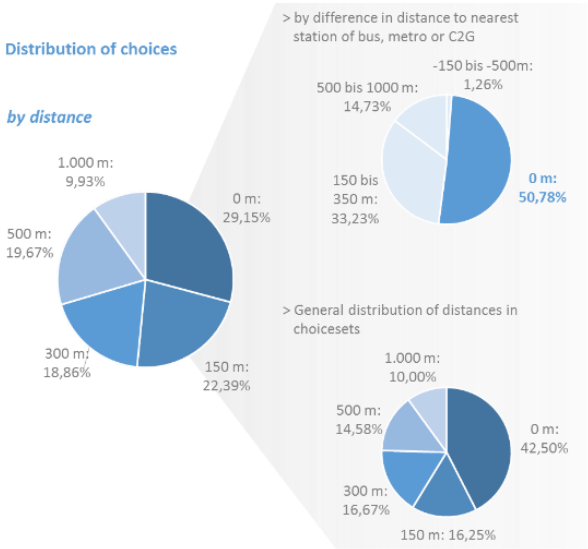


Figure 18: Distribution of choices by access time

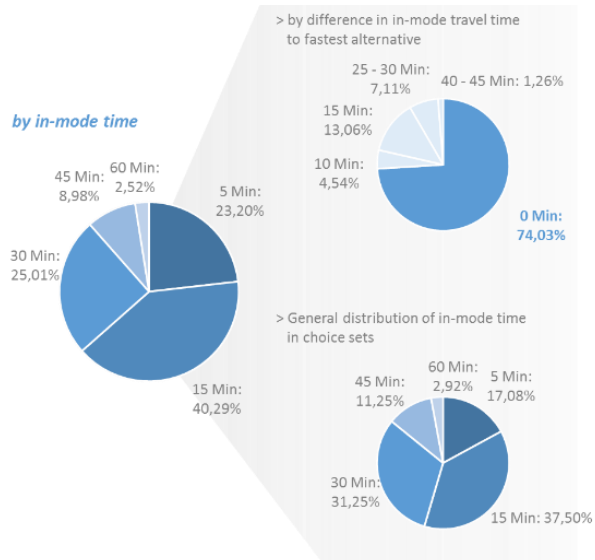


Figure 19: Distribution of choices by in-mode time

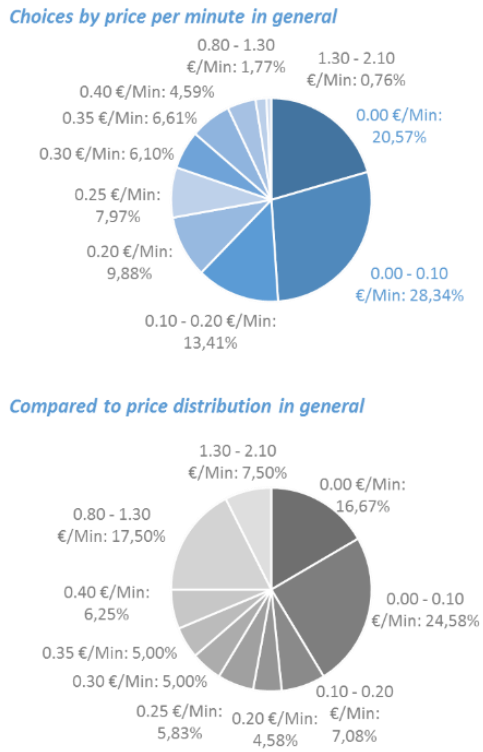


Figure 17: Distribution of choices by price

3.3.4 Scenarios

Comparing the scenarios as seen in Figure 18 there seems to be an influence of the *MORNING* scenario on the probability of choosing *BUS* as it drops down to 7.78% compared to around 22% in the other scenarios. The missing 14% seem to be substituted by the *METRO* alternative as it goes up comparable. Similar the *NONE* option is significantly less often chosen in the *CITY* scenario, which could be interpreted as the non-availability of some not included alternatives (i.e. taking the bike and similar). The *RAIN* scenario seems to negatively impact the *C2G* option as it goes down to 23.42% compared to round about 28% in the other scenarios.

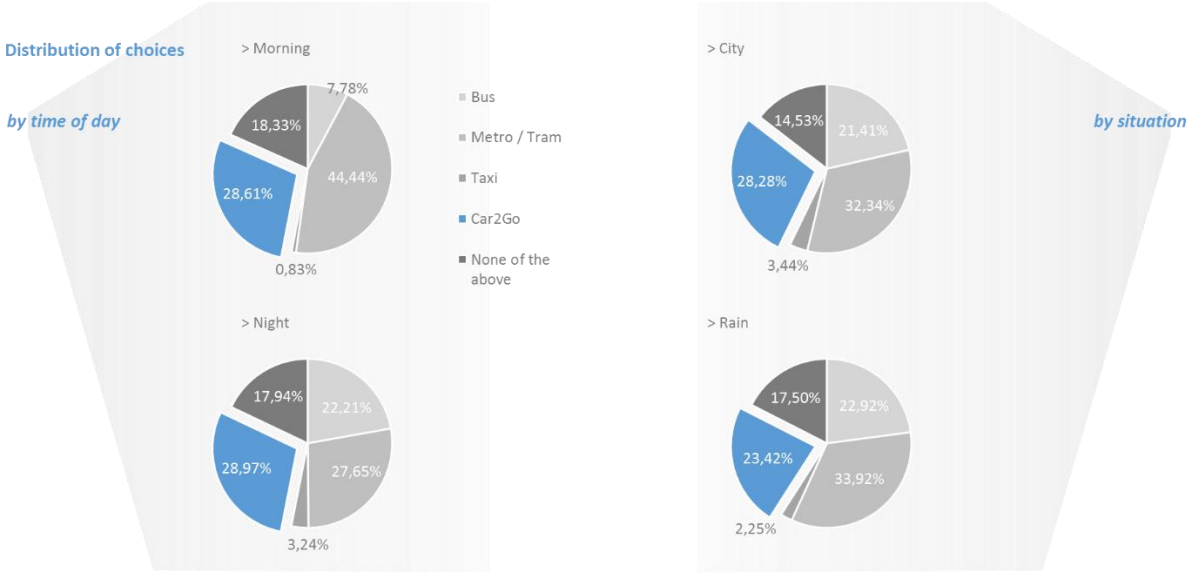


Figure 18: Distribution of choices by scenarios (other scenario variables set to average)

These effects become even more pronounced when comparing against the *BASE* scenario and keeping all other scenario variables at zero as seen in Figure 18. Interestingly in this case the substitution pattern for the *MORNING* scenario changes to benefit mostly *C2G* and also in the *NIGHT* scenario there is an increasing chance of observing *C2G* choices.

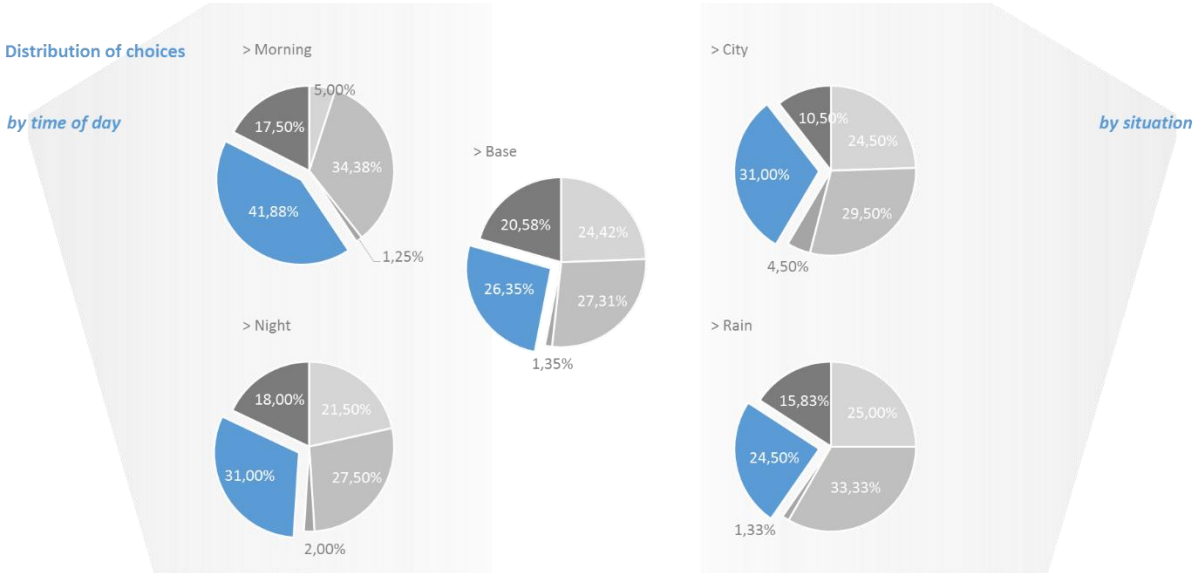


Figure 19: Distribution of choices by scenarios (other scenario variables set to zero)

Variable	Type	Mean	Std. Dev.	Min	Max
BUS	Dummy with "Bus" = 1	0.2000	0.4000	0.00	1.00
METRO	Dummy with "Metro" = 1	0.2000	0.4000	0.00	1.00
TAXI	Dummy with "Taxi" = 1	0.2000	0.4000	0.00	1.00
C2G	Dummy with "C2G" = 1	0.2000	0.4000	0.00	1.00
CITY	Dummy with "city" = 1	0.2667	0.4422	0.00	1.00
RAIN	Dummy with "rain" = 1	0.5000	0.5000	0.00	1.00
MORNING	Dummy with "night" = 1	0.1500	0.3571	0.00	1.00
NIGHT	Dummy with "morning" = 1	0.2833	0.4506	0.00	1.00
DISTANCE	in meters	197.8333	291.4802	0.00	1000.00
.BUS	in meters	390.0000	347.0593	0.00	1000.00
.METRO	in meters	376.6667	353.5642	0.00	1000.00
.TAXI	in meters	-	-	-	-
.C2G	in meters	222.5000	177.8537	0.00	500.00
TRAVELTIME	in minutes	19.7500	16.2137	0.00	66.00
.BUS	in minutes	28.5000	12.9738	5.00	66.00
.METRO	in minutes	32.3333	17.2631	5.00	66.00
.TAXI	in minutes	18.0833	10.2934	5.00	45.00
.C2G	in minutes	19.8333	10.7034	5.00	48.50
PRICE	in EUR	5.5470	7.7483	0.00	33.40
.BUS	in EUR	1.3333	0.9430	0.00	2.00
.METRO	in EUR	1.3333	0.9430	0.00	2.00
.TAXI	in EUR	19.0267	6.1908	8.00	33.40
.C2G	in EUR	6.0417	3.4419	1.00	13.40

Figure 20: Summary of parameters

Figure 20 finally sums up the parameters and scenario variables as presented in the DCE and discussed in the last two subchapters.

4. Estimation

4.1 Identification of relevant parameters

4.1.1 Simple conditional logit

Starting with a very simple definition of the model by only allowing the intercepts and main effects enter the equation the following utility function is derived for each alternative:

$$U_{nit}(X_{nit}) = \beta_{1i} + \beta_2 DISTANCE_{nit} + \beta_3 TRAVELTIME_{nit} + \beta_4 PRICE_{nit} + \varepsilon_{nit} \ .$$

Similar to the choice cards the access travel time is again stated in meters and therefore in the following labeled more fittingly as *DISTANCE*. The in-mode and egress travel time are for the rest of the paper combined in the variable *TRAVELTIME* and *PRICE* is the total price of the trip. In this model there is no discrimination for the scenario variables and as the parameter levels of in-mode travel time and *PRICE* are interlinked with the scenario *MORNING* there might be some bias in the coefficients.

	Estimate	Std. Error	t-value	Pr(>|t|)		Prior	Delta
BUS	1.4747	0.1009	14.6200	0.0000	***	1.4200	0.0547
METRO	1.9449	0.1019	19.0900	0.0000	***	1.4600	0.4849
TAXI	1.1794	0.2714	4.3500	0.0000	***	1.5300	-0.3506
C2G	2.0663	0.1076	19.2100	0.0000	***	1.3000	0.7663
DISTANCE	-0.0003	0.0001	-2.9600	0.0030	***	-0.0010	0.0007
TRAVELTIME	-0.0316	0.0025	-12.5600	0.0000	***	-0.1000	0.0684
PRICE	-0.1645	0.0156	-10.5300	0.0000	***	-0.2000	0.0355

Figure 21: Simple CLM

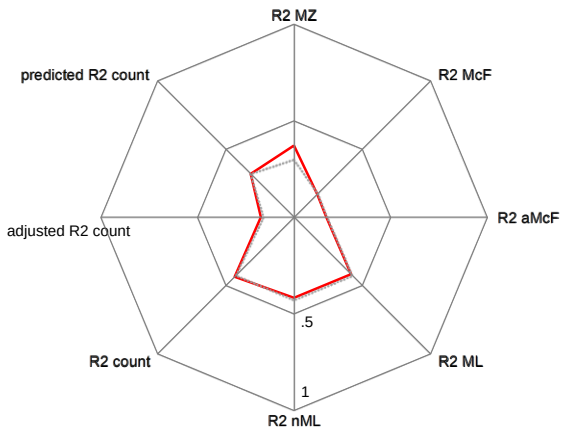


Figure 22: Goodness-of-fit indicators for simple CLM

All estimated coefficients are clearly significant at an one percent level and match expectations concerning sign and general ratio, but as seen in Figure 21 there are significant deviations from the before approximated priors. Especially the intercept for *TAXI* seems to have been over-approximated and similarly the intercept for *C2G* is not the lowest but the highest based in this model. Also the importance of *DISTANCE* and *TRAVELTIME* is three times lower than expected compared to the relevance of *PRICE*. Looking at the goodness-of-fit of the model as seen by the red line for various indicators in Figure 24, the fit is between 16.72% for the adjusted McFadden R^2 and 41.78% for the normalized maximum likelihood R^2 and the AIC is 6.435.

While adding a squared term to the main effects to check for potential quadratic functional form returns a significant coefficient for *PRICE* squared as seen in Figure 23, it also decrease the fit of the model slightly as seen by the grey line in Figure 24 and the AIC is slightly reduced to 6.398. As the squared term was not included in the design and also resulted in some nonsensical predictions for higher prices in the price analysis for C2G, it was dropped from further models in this paper. But for further future investigations a squared price term will be included already in the designing phase.

	Estimate	Std. Error	t-value	Pr(>|t|)		Prior	Delta
BUS	1.8549	0.1661	11.1700	0.0000	***	1.6065	0.2484
METRO	2.2854	0.1535	14.8900	0.0000	***	1.7534	0.5320
TAXI	1.4374	0.3066	4.6900	0.0000	***	1.3573	0.0801
C2G	2.6731	0.1533	17.4400	0.0000	***	2.0177	0.6554
DISTANCE	-0.0005	0.0003	-1.8800	0.0600	*	-0.0013	0.0008
TRAVELTIME	-0.0467	0.0099	-4.7000	0.0000	***	-0.0477	0.0010
PRICE	-0.2888	0.0282	-10.2200	0.0000	***	-0.1164	-0.1724
DISTANCE SQUARED	0.0001	0.0001	0.9900	0.3220			
TRAVELTIME SQUARED	0.0003	0.0001	1.7300	0.0830	*		
PRICE SQUARED	0.0068	0.0010	6.6800	0.0000	***		

Figure 23: Simple CLM with non-linear effects

4.1.2 Individual heterogeneity by attributes

To check for random parameters and possible distributions of those, a simulation exercise was carried out estimating the individual effect of each respondent for the simple model and comparing the returned coefficients. Figure 26 shows the variation in coefficients and while the main effects seem to be clearly negatively log-normal distributed, the intercepts for *BUS*, *METRO* and *C2G* could either be normal or lognormal distributed. But the variations of the intercepts are with a range of one to two percent and for the coefficients between two and seven percent relative small.

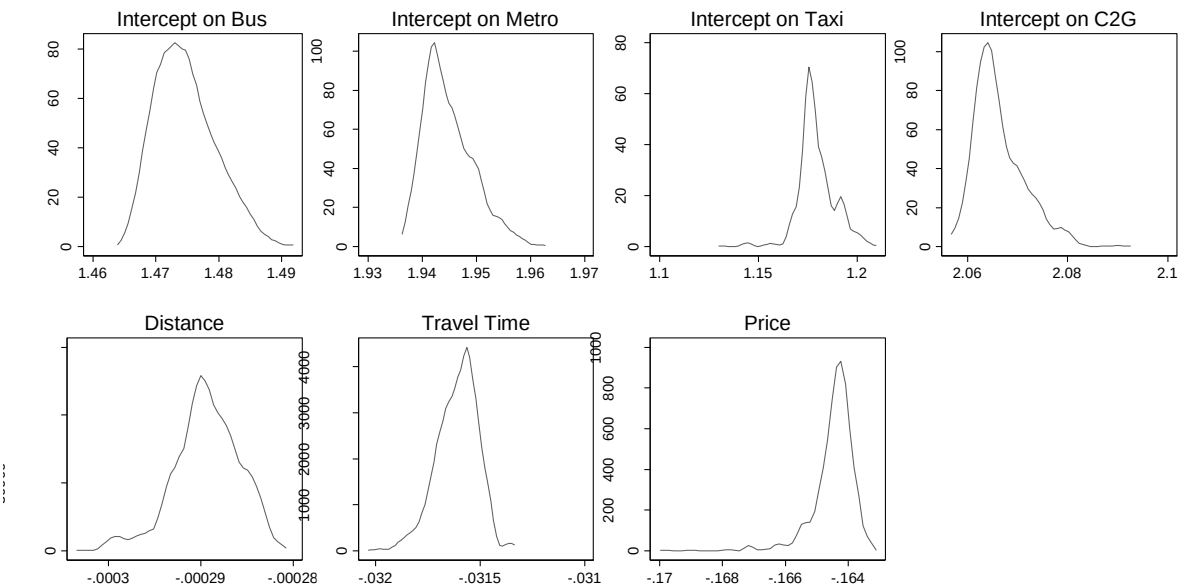


Figure 24: Individual impact in simple conditional logit

4.1.3 Simple mixed logit

Adding in a first step only a negatively log-normal distributed parameter for the main effects based on the individual impact, leaves the following utility function:

$$U_{nit}(X_{nit}) = \beta_{1nit} + \beta_{2nit}(-DISTANCE_{nit}) + \beta_{3nit}(-TRAVELTIME_{nit}) + \beta_{4nit}(-PRICE_{nit}) + \varepsilon_{nit}$$

with $\beta_{mnit} = \gamma_{mi} + \eta_{mnit}$ and $\eta_{mnit} \sim \ln N(0,1)$ for $m \in [2,4]$.

Compared to the CLM the log-likelihood of the estimation is slightly increased to -3 203.5 from before -3 213.5, but still lower than for the CLM with a squared term with -3 191.8. Based on the t-statistic, which can be seen in Figure 25, only the price seems to be impacted by random individual heterogeneity, while all other main effects return insignificant standard deviations.

	Estimate	Std. Error	t-value	Pr(>|t|)	
BUS	1.6099	0.1115	14.4404	0.0000	***
METRO	2.0757	0.1122	18.5070	0.0000	***
TAXI	1.5367	0.3011	5.1032	0.0000	***
C2G	2.4029	0.1441	16.6758	0.0000	***
DISTANCE	-8.2697	0.3869	-21.3769	0.0000	***
Standard Deviation	0.0019	1.3176	0.0015	0.9988	
TRAVELTIME	-3.4774	0.0821	-42.3721	0.0000	***
Standard Deviation	0.0005	0.2179	0.0025	0.9980	
PRICE	-1.5486	0.1080	-14.3383	0.0000	***
Standard Deviation	0.5727	0.0817	7.0083	0.0000	***

Figure 25: Simple MXL with random parameters for main effects

Extending the random parameter definition also to the intercepts, changes the utility function to

$$U_{nit}(X_{nit}) = \beta_{1nit} + \beta_{2nit}(-DISTANCE_{nit}) + \beta_{3nit}(-TRAVELTIME_{nit}) + \beta_{4nit}(-PRICE_{nit}) + \varepsilon_{nit}$$

with $\beta_{mnit} = \gamma_{mi} + \eta_{mnit}$

and $\eta_{mnit} \sim N(0,1)$ for $m \in [11,14]$ and $\eta_{mnit} \sim \ln N(0,1)$ for $m \in [2,4]$,

which returns the estimated coefficients in Figure 26. In this case none of the random parameters is significant at any common level and also the *TAXI* intercept seems not significantly differ from zero with a relatively high standard deviation. The log-likelihood is again slightly increased to -3 190.4. While the coefficients of *DISTANCE* and *TRAVELTIME* are robust to the change in model, the *PRICE* coefficient is slightly lower and the standard deviation now insignificant. The mean of the intercepts for *BUS*, *METRO* and *C2G* seem not to be affected by the change and the standard deviation is close to zero and insignificant for these variables.

	Estimate	Std. Error	t-value	Pr(>|t|)	
BUS	1.5487	0.1087	14.2516	0.0000	***
Standard Deviation	0.0008	0.2832	0.0027	0.9979	
METRO	2.0159	0.1096	18.3997	0.0000	***
Standard Deviation	0.0004	0.2685	0.0016	0.9987	
TAXI	-37.7040	39.9580	-0.9436	0.3454	
Standard Deviation	21.7420	19.7100	1.1031	0.2700	
C2G	2.3810	0.1391	17.1201	0.0000	***
Standard Deviation	0.0024	0.3124	0.0077	0.9939	
DISTANCE	-8.2298	0.3717	-22.1390	0.0000	***
Standard Deviation	0.0217	1.2074	0.0179	0.9857	
TRAVELTIME	-3.4858	0.0830	-41.9777	0.0000	***
Standard Deviation	0.0002	0.2096	0.0009	0.9993	
PRICE	-1.4622	0.0941	-15.5416	0.0000	***
Standard Deviation	0.1778	0.2490	0.7139	0.4753	

Figure 26: Simple MXL with random parameters for intercepts and main effects

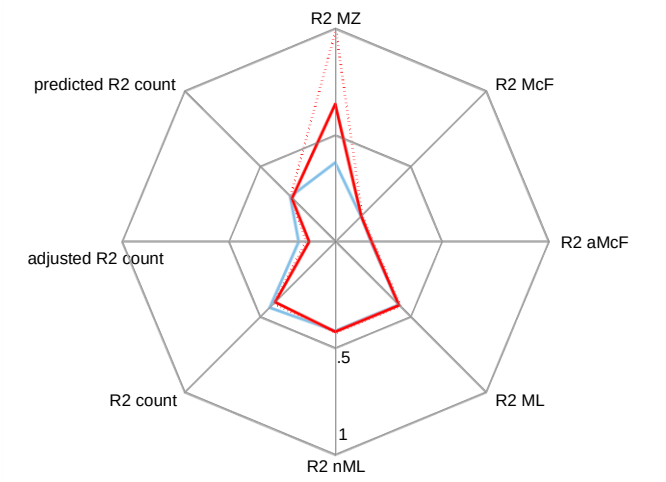


Figure 27: Goodness-of-fit indicators for simple MXL

Comparing the GoF for both models and the basic CLM (blue line), the introduction of random parameters leads to a small increase in fit with pseudo R^2's between 16.98% for the adjusted McFadden's measure and 42.26 for the MXL with randomness only in the main effects (solid red line) and between 17.32% for the adjusted McFadden and 42.89 for the ML R^2 for the MXL with randomness also in the intercepts (dotted red line). The Ratio of Variance R^2_{MZ} seems to be misleading when analysing a MXL as the variance introduced by the random parameters also enters the equation of the predicted utilities. The AIC is further reduced to 6 415.2 and 6 388.8 for both MXLs. As the introduction of random parameters for the intercepts returns only non-significant standard deviations and results in problems concerning the identification of the intercept for *TAXI* services, only the main effects were further analyzed for random effects in the following chapters.

4.2 Relevance of situational influences on valuation of travel time savings

4.2.1 Full conditional logit

Including the situational variables into the conditional logit extends the utility function to fit closely to the actual model as developed in the designing phase (see also Equation 2), but again in-mode and egress time are combined in one variable:

$$
\begin{aligned}
U_{nit}(X_{nit}) = \; & \beta_{11}BUS_{nit} + \beta_{12}METRO_{nit} + \beta_{13}TAXI_{nit} + \beta_{14}C2G_{nit} \\
& + (\beta_{111}CITY_{nt} + \beta_{112}RAIN_{nt} + \beta_{113}MORNING_{nt} + \beta_{114}NIGHT_{nt})\, BUS_{nit} \\
& + (\beta_{121}CITY_{nt} + \beta_{122}RAIN_{nt} + \beta_{123}MORNING_{nt} + \beta_{124}NIGHT_{nt})\, METRO_{nit} \\
& + (\beta_{131}CITY_{nt} + \beta_{132}RAIN_{nt} + \beta_{133}MORNING_{nt} + \beta_{134}NIGHT_{nt})\, TAXI_{nit} \\
& + (\beta_{141}CITY_{nt} + \beta_{142}RAIN_{nt} + \beta_{143}MORNING_{nt} + \beta_{144}NIGHT_{nt})\, C2G_{nit} \\
& + \beta_2 DISTANCE_{nit} \; + \beta_{21}DISTANCE_{nit}CITY_{nt} + \beta_{22}DISTANCE_{nit}RAIN_{nt} \\
& \qquad + \beta_{23}DISTANCE_{nit}MORNING_{nt} + \beta_{24}DISTANCE_{nit}Night_{nt} \\
& + \beta_3 TRAVEL_{nit} \; + \beta_{31}TRAVEL_{nit}CITY_{nt} + \beta_{32}TRAVEL_{nit}RAIN_{nt} \\
& \qquad + \beta_{33}TRAVEL_{nit}MORNING_{nt} + \beta_{34}TRAVEL_{nit}NIGHT_{nt} \\
& + \beta_4 PRICE_{nit} \; + \beta_{51}PRICE_{nit}CITY_{nt} + \beta_{52}PRICE_{nit}RAIN_{nt} \\
& \qquad + \beta_{53}PRICE_{nit}MORNING_{nt} + \beta_{54}PRICE_{nit}NIGHT_{nt} \\
& + \varepsilon_{nit}
\end{aligned}
$$

The log-likelihood of the model is increased to -3 159.1 and the AIC decreased to 6 326.3. The goodness-of-fit is increased for all measurements and ranges from 18.14% for the adjusted McFadden R^2 to 41.44 to 44.35% for MZ, ML and count R^2.

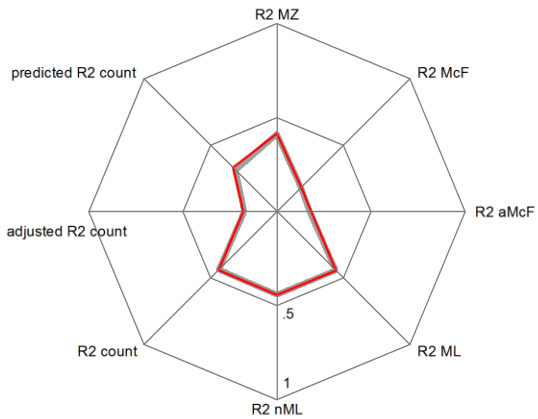


Figure 28: Goodness of fit full CLM

The estimated coefficients as seen in Figure 29 show especially for *DISTANCE* and *PRICE* an impact of the scenario. While the coefficient for *DISTANCE* increases as expected for *RAIN* and *NIGHT*, it decreases if the choice situation was situated in the *CITY*. Similar the price is less important, if in the city and with rain, but increases in importance if it is *MORNING* rush hour with its longer *TRAVELTIME* and therefore higher *PRICE* for *C2G* and *TAXI*. While the direction of the impacts of the scenarios seems logical over all main effects, there is a problem in the change of importance for *DISTANCE* as the coefficient turns positive for *CITY* and *MORNING*, which results in nonsensical predictions as the alternative which is farther away will be preferred, if all other attributes are equal. This effect might be due for the *MORNING* scenario to the interaction between *MORNING*, *TRAVELTIME* and *PRICE* for *C2G* and *TAXI*, which makes especially *METRO* more probable and due to the balancing of probabilities

in the designing phase is again interlinked to a slightly higher average distance for the *METRO* alternatives (381m versus 313m) while the opposite is true for *BUS* (159m versus 340m) and *C2G* (192m versus 237m). Similar distances for *BUS* (335m versus 329m) and *METRO* (489m versus 266m) are in the *CITY* scenario higher while the opposite is true for *C2G* (199m versus 242m). Therefore some of the higher *PRICE* effect for *C2G* might have interacted with the coefficient of *DISTANCE*. In conclusion there might be some collinearity between the main effects and scenarios, which biases the coefficients.

		Estimate	Std. Error	t-value	Pr(>|t|)	Prior	Delta
Intercept	BUS	1.6435	0.1877	8.7500	0.0000 ***	1.4200	0.2235
	METRO	2.0832	0.2018	10.3300	0.0000 ***	1.4600	0.6232
	TAXI	1.5670	0.5417	2.8900	0.0040 ***	1.5300	0.0370
	C2G	2.3079	0.2075	11.1200	0.0000 ***	1.3000	1.0079
City	BUS	0.3612	0.2623	1.3800	0.1680	0.5000	-0.1388
	METRO	0.3350	0.2779	1.2100	0.2280	0.5000	-0.1650
	TAXI	-0.0873	0.5809	-0.1500	0.8810	0.5000	-0.5873
	C2G	0.1790	0.2491	0.7200	0.4720	0.5000	-0.3210
Rain	BUS	-0.3318	0.2110	-1.5700	0.1160	0.5000	-0.8318
	METRO	-0.4092	0.2149	-1.9000	0.0570 *	0.8000	-1.2092
	TAXI	-1.1697	0.6270	-1.8700	0.0620 *	1.0000	-2.1697
	C2G	-0.6521	0.2444	-2.6700	0.0080 ***	0.8000	-1.4521
Morning	BUS	-0.9879	0.4403	-2.2400	0.0250 **	-0.5000	-0.4879
	METRO	-0.2076	0.2874	-0.7200	0.4700	-0.2000	-0.0076
	TAXI	3.8065	1.4190	2.6800	0.0070 ***	-0.5000	4.3065
	C2G	0.7329	0.3592	2.0400	0.0410 **	-1.0000	1.7329
Night	BUS	-0.0948	0.2450	-0.3900	0.6990	-0.2000	0.1052
	METRO	-0.0738	0.2652	-0.2800	0.7810	-0.1200	0.0462
	TAXI	0.6692	0.6379	1.0500	0.2940	0.3000	0.3692
	C2G	0.1143	0.2620	0.4400	0.6630	-0.3000	0.4143
Main Effects	DISTANCE	-0.0003	0.0002	-2.0700	0.0380 **	-0.0010	0.0007
	.CITY	0.0007	0.0003	2.6200	0.0090 ***	0.0005	0.0002
	.RAIN	-0.0006	0.0002	-2.6100	0.0090 ***	-0.0020	0.0014
	.MORNING	0.0005	0.0003	1.5700	0.1160	-0.0005	0.0010
	.NIGHT	-0.0006	0.0003	-1.9600	0.0510 *	-0.0010	0.0004
	TRAVELTIME	-0.0352	0.0047	-7.4800	0.0000 ***	-0.1000	0.0648
	.CITY	-0.0061	0.0064	-0.9600	0.3380	0.0100	-0.0161
	.RAIN	0.0205	0.0055	3.7100	0.0000 ***	-0.0050	0.0255
	.MORNING	0.0013	0.0088	0.1500	0.8830	-0.0500	0.0513
	.NIGHT	-0.0060	0.0075	-0.8000	0.4250	-0.0500	0.0440
	PRICE	-0.2103	0.0319	-6.5900	0.0000 ***	-0.2000	-0.0103
	.CITY	0.0654	0.0337	1.9400	0.0520 *	0.0100	0.0554
	.RAIN	0.0850	0.0367	2.3200	0.0210 **	0.0500	0.0350
	.MORNING	-0.2361	0.0706	-3.3400	0.0010 ***	-0.0100	-0.2261
	.NIGHT	-0.0585	0.0380	-1.5400	0.1240	0.0500	-0.1085

Figure 29: Full CLM

For the intercepts only the rain and morning interactions return any significant relevant coefficients, where rain reduces the intercept for all alternatives and increases the likelihood of the *NONE* option if all other parameters are equal, while *MORNING* reduces the likelihood of *BUS* and *METRO* and positively effects *TAXI* and *C2G*. Again these last impacts need to be seen in context with the higher *TRAVELTIME* and therefore *PRICE*, which counteract the increase in intercept.

		Estimate	Std. Error	t-value	Pr(>|t|)	Full CLM	Delta	in %
Intercept	BUS	1.5395	0.1823	8.4500	0.0000 ***	1.6435	-0.1040	-6.33%
	METRO	2.0207	0.1998	10.1100	0.0000 ***	2.0832	-0.0625	-3.00%
	TAXI	1.6130	0.5431	2.9700	0.0030 ***	1.5670	0.0460	2.93%
	C2G	2.2671	0.2074	10.9300	0.0000 ***	2.3079	-0.0407	-1.76%
City	BUS	0.6999	0.2356	2.9700	0.0030 ***	0.3612	0.3387	93.77%
	METRO	0.7566	0.2410	3.1400	0.0020 ***	0.3350	0.4215	125.82%
	TAXI	-0.3628	0.5709	-0.6400	0.5250	-0.0873	-0.2755	315.69%
	C2G	0.2550	0.2460	1.0400	0.3000	0.1790	0.0759	42.41%
Rain	BUS	-0.2581	0.2083	-1.2400	0.2150	-0.3318	0.0737	-22.21%
	METRO	-0.3818	0.2151	-1.7800	0.0760 *	-0.4092	0.0274	-6.69%
	TAXI	-1.2055	0.6264	-1.9200	0.0540 *	-1.1697	-0.0358	3.06%
	C2G	-0.6235	0.2433	-2.5600	0.0100 **	-0.6521	0.0286	-4.38%
Morning	BUS	-0.6306	0.3970	-1.5900	0.1120	-0.9879	0.3573	-36.17%
	METRO	0.0149	0.2576	0.0600	0.9540	-0.2076	0.2225	-107.17%
	TAXI	3.2341	1.3801	2.3400	0.0190 **	3.8065	-0.5725	-15.04%
	C2G	0.7680	0.3553	2.1600	0.0310 **	0.7329	0.0352	4.80%
Night	BUS	-0.1745	0.2414	-0.7200	0.4700	-0.0948	-0.0798	84.15%
	METRO	-0.2045	0.2600	-0.7900	0.4310	-0.0738	-0.1307	177.18%
	TAXI	0.7128	0.6347	1.1200	0.2610	0.6692	0.0436	6.51%
	C2G	0.0727	0.2600	0.2800	0.7800	0.1143	-0.0415	-36.34%
Main Effects	DISTANCE	-0.0001	0.0001	-0.5700	0.5710	-0.0003	0.0003	-77.47%
	.CITY	-	-	-	-	0.0007		
	.RAIN	-0.0008	0.0002	-3.4900	0.0000 ***	-0.0006	-0.0002	28.15%
	.MORNING	-	-	-	-	0.0005		
	.NIGHT	-0.0006	0.0003	-1.8700	0.0620 **	-0.0006	0.0000	-5.45%
	TRAVELTIME	-0.0345	0.0046	-7.4700	0.0000 ***	-0.0352	0.0007	-2.10%
	.CITY	-0.0109	0.0061	-1.7900	0.0740 *	-0.0061	-0.0048	79.16%
	.RAIN	0.0192	0.0055	3.5000	0.0000 ***	0.0205	-0.0013	-6.53%
	.MORNING	-0.0049	0.0085	-0.5800	0.5650	0.0013	-0.0062	-477.60%
	.NIGHT	-0.0026	0.0074	-0.3500	0.7240	-0.0060	0.0034	-56.32%
	PRICE	-0.2145	0.0318	-6.7400	0.0000 ***	-0.2103	-0.0042	1.99%
	.CITY	0.0898	0.0326	2.7600	0.0060 ***	0.0654	0.0243	37.19%
	.RAIN	0.0881	0.0366	2.4100	0.0160 **	0.0850	0.0031	3.66%
	.MORNING	-0.2023	0.0680	-2.9800	0.0030 ***	-0.2361	0.0338	-14.31%
	.NIGHT	-0.0638	0.0378	-1.6900	0.0910 *	-0.0585	-0.0054	9.22%

Figure 30: Full CLM without Interaction DISTANCE with MORNING / CITY

Dropping the interaction between *DISTANCE* and *MORNING* / *CITY* decreases the GoF of the model very slightly to log-likelihood of -3 163.9 and a pseudo R^2 between 17.34% for the adjusted McFadden and 44.14% for R^2_{ML} as can be seen by the almost parallel grey line in Figure 28. The estimates of the coefficients differ in some cases widely as Figure 30. Obviously this change in coefficients is greatest in the interactions of *CITY* and *MORNING*, but also some of the other interactions are changed. The base

coefficient of *DISTANCE* drops almost to zero and is no longer relevant. Therefore in this model *DISTANCE* only has an impact in the *RAIN* and *NIGHT* scenario. While risking a bias of coefficients due to the exclusion of significant variables, which are correlated with other explanatories as explained above and therefore introducing a correlation between explanatories and the error term, this model would have the benefit of more rational predictions of the WTP as now all ratios will have the expected sign.

4.2.2 Individual heterogeneity by attributes

Repeating the simulation exercise from 4.1.2 for the full conditional logit with all interactions, returns a similar result for main effects and while the intercepts are now clearly normal distributed. The interactions seem to be normal distributed as well, but again the magnitude of the individual impact is very low.

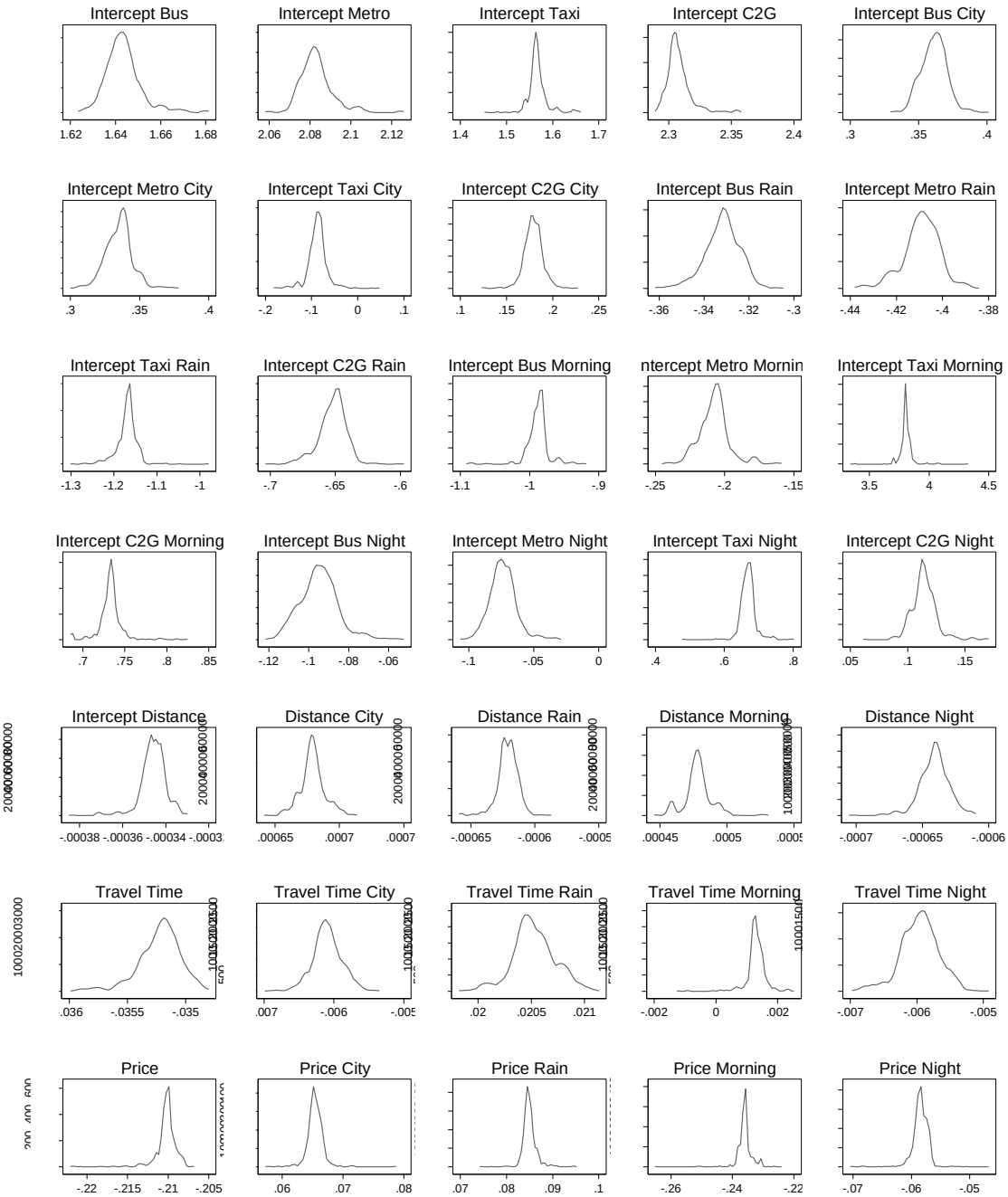


Figure 31: Individual impact in full CLM

4.2.3 Full Mixed logit with situational parameters

Extending the full model again to an MXL including random parameters for the main effects, yields the following utility function:

$$U_{nit}(X_{nit}) = \beta_{1nit} + \beta_{2nit}(-DISTANCE_{nit}) + \beta_{3nit}(-TRAVELTIME_{nit}) + \beta_{4nit}(-PRICE_{nit}) + \varepsilon_{nit}$$

with $\beta_{mnit} = \gamma_{mi} + \gamma_{m1i}CITY_{nt} + \gamma_{m2i}RAIN_{nt} + \gamma_{m3i}MORNING_{nt} + \gamma_{m4i}NIGHT_{nt} + \eta_{mnit}$

and $\eta_{mnit} = 0$ for $m \in [11, 14]$ and $\eta_{mnit} \sim \ln N(0,1)$ for $m \in [2, 4]$.

The log-likelihood of this model is further increased to -3 153.8 and the GoF are between 17.54 for the adjusted McFadden and 44.61 for the Maximum Likelihood R^2 as seen in Figure 32. The AIC is reduced to 6 377.6.

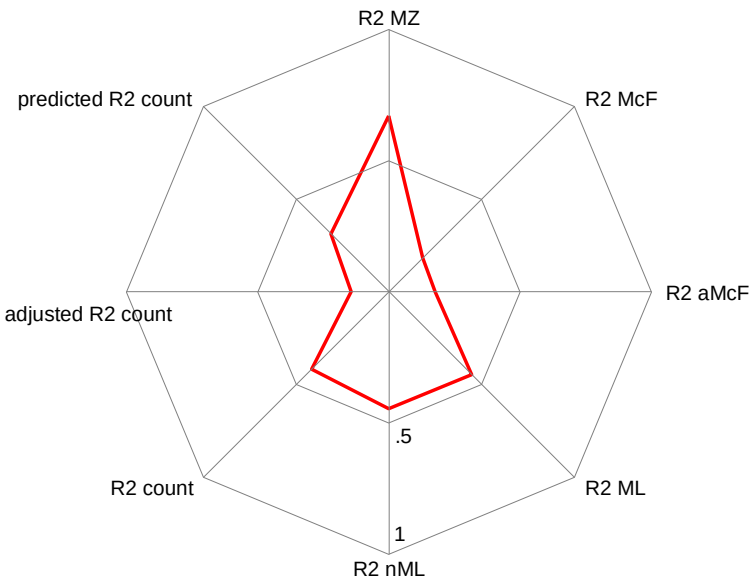


Figure 32: Goodness of fit full MXL

Again all alternative specific constants and main effects are significant at below five percent. The intercepts are only effected by the situational variable *RAIN*, which decreases the base utility of all alternatives compared to the *NONE* option. The importance of *TRAVELTIME* is increased by *CITY*, *MORNING* and *NIGHT* and decreased by *RAIN*. The *PRICE* is less important in the *CITY* and there is a significant random heterogeneity to it. *DISTANCE* and *TRAVELTIME* are not affected by the random parameters as the *t*-statistic is very low and the coefficient of the standard deviation goes to zero.

Again there is some problem in identifying the effect of the *MORNING* and *CITY* scenario on *DISTANCE* as in both scenarios the total effect of the access time goes to zero as the maximization of the log-likelihood function seems to hit a plateau for both coefficients. Dropping both interactions therefore from the model increases the predictive value of model without losing any efficiency in form of decreases in the GoF. Actually the adjusted measures gain slightly as the number of parameters drops and the *t*-values of some coefficients can be reduced thereby increasing the statistical relevance of those parameters as can be seen by comparing Figure 33 and Figure 34.

		Estimate	Std. Error	t-value	Pr(>|t|)	
Intercept	BUS	1.8234	0.2050	8.8943	0.0000	***
	METRO	2.2263	0.2172	10.2516	0.0000	***
	TAXI	1.9605	0.7649	2.5631	0.0104	**
	C2G	2.7564	0.3443	8.0049	0.0000	***
City	BUS	0.3860	0.2465	1.5661	0.1173	
	METRO	0.4368	0.2506	1.7429	0.0814	*
	TAXI	-0.3260	0.7971	-0.4090	0.6825	
	C2G	-0.0490	0.3556	-0.1377	0.8905	
Rain	BUS	-0.5303	0.2127	-2.4934	0.0127	**
	METRO	-0.6154	0.2114	-2.9115	0.0036	***
	TAXI	-0.8527	0.7035	-1.2120	0.2255	
	C2G	-0.7537	0.3337	-2.2586	0.0239	**
Morning	BUS	-0.5028	0.3841	-1.3089	0.1906	
	METRO	0.1632	0.2862	0.5702	0.5685	
	TAXI	1.4062	1.4911	0.9431	0.3457	
	C2G	0.6843	0.5847	1.1705	0.2418	
Night	BUS	0.0665	0.2409	0.2760	0.7826	
	METRO	0.1306	0.2594	0.5034	0.6146	
	TAXI	0.3757	0.6436	0.5838	0.5594	
	C2G	0.0959	0.3085	0.3108	0.7559	
Main Effects	DISTANCE	-7.7524	0.5103	-15.1904	0.0000	***
	.CITY	-1.4162	1.8916	-0.7487	0.4541	
	.RAIN	0.6713	0.5038	1.3325	0.1827	
	.MORNING	-10.1036	256.2387	-0.0394	0.9685	
	.NIGHT	0.5302	0.4770	1.1116	0.2663	
	TRAVELTIME	-3.5750	0.1828	-19.5527	0.0000	***
	.CITY	0.3409	0.1850	1.8424	0.0654	*
	.RAIN	-0.8536	0.1979	-4.3124	0.0000	***
	.MORNING	0.5118	0.2705	1.8922	0.0585	*
	.NIGHT	0.6494	0.2098	3.0955	0.0020	***
	PRICE	-1.2724	0.2533	-5.0237	0.0000	***
	.CITY	-0.5483	0.3139	-1.7471	0.0806	*
	.RAIN	-0.3640	0.2428	-1.4993	0.1338	
	.MORNING	0.5762	0.3808	1.5131	0.1303	
	.NIGHT	0.1035	0.2276	0.4548	0.6492	
Std. Dev.	DISTANCE	0.0012	0.4610	0.0025	0.9980	
	TRAVELTIME	0.0001	0.2080	0.0005	0.9996	
	PRICE	0.4982	0.0754	6.6101	0.0000	***

Figure 33: Full MXL with scenario variables

		Estimate	Std. Error	t-value	Pr(>|t|)		Full MXL	Delta	in %
Intercept	BUS	1.6756	0.2015	8.3148	0.0000	***	1.8234	-0.1478	-8.11%
	METRO	2.1017	0.2181	9.6372	0.0000	***	2.2263	-0.1245	-5.59%
	TAXI	2.2810	0.6368	3.5819	0.0003	***	1.9605	0.3205	16.35%
	C2G	2.7851	0.2969	9.3806	0.0000	***	2.7564	0.0286	1.04%
City	BUS	0.5415	0.2365	2.2892	0.0221	**	0.3860	0.1555	40.29%
	METRO	0.6012	0.2392	2.5135	0.0120	**	0.4368	0.1644	37.63%
	TAXI	-0.6693	0.6769	-0.9887	0.3228		-0.3260	-0.3432	105.28%
	C2G	-0.0535	0.3069	-0.1745	0.8615		-0.0490	-0.0046	9.38%
Rain	BUS	-0.3769	0.2160	-1.7449	0.0810	*	-0.5303	0.1534	-28.93%
	METRO	-0.4829	0.2171	-2.2246	0.0261	**	-0.6154	0.1325	-21.53%
	TAXI	-1.1286	0.6343	-1.7792	0.0752	*	-0.8527	-0.2759	32.36%
	C2G	-0.7649	0.3049	-2.5083	0.0121	**	-0.7537	-0.0112	1.48%
Morning	BUS	-0.2996	0.3767	-0.7954	0.4264		-0.5028	0.2032	-40.41%
	METRO	0.2696	0.2680	1.0058	0.3145		0.1632	0.1064	65.19%
	TAXI	0.8507	1.1537	0.7374	0.4609		1.4062	-0.5555	-39.50%
	C2G	0.5539	0.4524	1.2243	0.2208		0.6843	-0.1305	-19.06%
Night	BUS	0.0361	0.2358	0.1533	0.8782		0.0665	-0.0303	-45.64%
	METRO	0.0845	0.2532	0.3336	0.7387		0.1306	-0.0461	-35.32%
	TAXI	0.3053	0.6196	0.4928	0.6222		0.3757	-0.0704	-18.74%
	C2G	0.0063	0.3010	0.0208	0.9834		0.0959	-0.0896	-93.48%
Main Effects	DISTANCE	-9.3152	1.4393	-6.4719	0.0000	***	-7.7524	-1.5628	20.16%
	.CITY	-	-	-	-		-1.4162		
	.RAIN	2.2701	1.4766	1.5373	0.1242		0.6713	1.5988	238.15%
	.MORNING	-	-	-	-		-10.1036		
	.NIGHT	0.4344	0.3861	1.1252	0.2605		0.5302	-0.0958	-18.07%
	TRAVELTIME	-3.6078	0.1769	-20.3950	0.0000	***	-3.5750	-0.0328	0.92%
	.CITY	0.4006	0.1754	2.2838	0.0224	**	0.3409	0.0597	17.52%
	.RAIN	-0.8265	0.1924	-4.2968	0.0000	***	-0.8536	0.0271	-3.17%
	.MORNING	0.6507	0.2405	2.7056	0.0068	***	0.5118	0.1389	27.14%
	.NIGHT	0.5972	0.2069	2.8866	0.0039	***	0.6494	-0.0522	-8.04%
	PRICE	-1.1624	0.1941	-5.9895	0.0000	***	-1.2724	0.1101	-8.65%
	.CITY	-0.7164	0.2875	-2.4919	0.0127	**	-0.5483	-0.1681	30.65%
	.RAIN	-0.4669	0.2151	-2.1703	0.0300	**	-0.3640	-0.1029	28.27%
	.MORNING	0.3756	0.3090	1.2158	0.2240		0.5762	-0.2006	-34.81%
	.NIGHT	0.0861	0.2111	0.4078	0.6834		0.1035	-0.0174	-16.84%
Std. Dev.	DISTANCE	0.0007	0.4694	0.0014	0.9989		0.0012	-0.0005	-42.94%
	TRAVELTIME	0.0003	0.2116	0.0016	0.9987		0.0001	0.0002	236.59%
	PRICE	0.4891	0.0739	6.6161	0.0000	***	0.4982	-0.0091	-1.82%

Figure 34: Full MXL with scenario variables without interaction distance and morning / city

4.3 Relevance of individual characteristics on valuation of travel time savings

4.3.1 Binary logit on influence of individual characteristics on DCE

To gain a further inside into the relevance of the socio-economic variables in the DCE a binary logit was estimated on the probability of observing a choice of *C2G*:

$$\Pr(y_{nt} = C2G|X_{nt}) = \alpha_1(PRICE_{C2G_{nt}}) + \alpha_2(DISTANCE_{C2G_{nt}} - DISTANCE_{Min_{nt}}) + \alpha_3(TRAVELTIME_{C2G_{nt}} - TRAVELTIME_{Min_{nt}}) + X_n'\delta + \mu_{nt},$$

where X_n is a vector of socio-demographic variables and individual stated preferences. While α should control for the major influences of the scenario, the interesting part is which of the variables in X are relevant and should be included into the MXL.

	Estimate	Std. Err.	z	P>|z|	[95% Conf. Interval]	
PRICE C2G	-0.1351	0.0181	-7.45	0.00	-0.171	-0.100
DELTA TRAVELTIME C2G TO MIN	0.0341	0.0053	6.43	0.00	0.024	0.044
DELTA DISTANCE C2G TO MIN	-0.0009	0.0005	-2.07	0.04	-0.002	0.000
GENDER	-0.3154	0.0958	-3.29	0.00	-0.503	-0.128
AGE	-0.0087	0.0046	-1.88	0.06	-0.018	0.000
ETHNIC GROUP						
MINORITY ONE	0.5621	0.2913	1.93	0.05	-0.009	1.133
MINORITY TWO	-0.7604	0.3103	-2.45	0.01	-1.369	-0.152
INCOME						
1.000€-2.000€	0.0340	0.1633	0.21	0.84	-0.286	0.354
2.000€-3.000€	0.7113	0.1902	3.74	0.00	0.339	1.084
3.000€-4.000€	0.9658	0.3499	2.76	0.01	0.280	1.652
4.000€-5.000€	1.6913	0.4640	3.65	0.00	0.782	2.601
>5.000€	1.4295	0.7248	1.97	0.05	0.009	2.850
NA	0.1214	0.1652	0.73	0.46	-0.203	0.445
HOUSEHOLD SIZE	0.1237	0.0422	2.93	0.00	0.041	0.206
CARS	0.0463	0.0387	1.20	0.23	-0.029	0.122
SMARTPHONE	0.4963	0.1638	3.03	0.00	0.175	0.817
AGREEMENT C2G (Continous in 1 to 5)	0.2350	0.0750	3.13	0.00	0.088	0.382
LIKELINESS C2G (Continous in 1 to 5)	-0.3407	0.0700	-4.87	0.00	-0.478	-0.203
KNOW CS	0.0972	0.0878	1.11	0.27	-0.075	0.269
PRIOR USE CS	-0.2551	0.4084	-0.62	0.53	-1.056	0.545
PREFERENCE PUBLIC TRANSPORT	-0.0549	0.2006	-0.27	0.79	-0.448	0.338
PREFERENCE TAXI	1.2364	0.5599	2.21	0.03	0.139	2.334
PREFERENCE OTHER	-0.3531	0.1424	-2.48	0.01	-0.632	-0.074
NO PREFERENCE CS	0.1995	0.1289	1.55	0.12	-0.053	0.452
OPINION CARS						
HIGHLY POSITIVE	0.0310	0.0379	0.82	0.41	-0.043	0.105
POSITIVE	0.0272	0.0350	0.78	0.44	-0.041	0.096
MEDIUM	-0.0190	0.0271	-0.70	0.48	-0.072	0.034
LOW	-0.1456	0.0296	-4.92	0.00	-0.204	-0.088
OPINION CS						
POSITIVE	0.1410	0.2182	0.65	0.52	-0.287	0.569
MEDIUM	0.0412	0.1978	0.21	0.84	-0.346	0.429
LOW	-0.3870	0.2069	-1.87	0.06	-0.793	0.019
INTERCEPT	1.2714	0.6471	1.96	0.05	0.003	2.540

Figure 35: Binary logit on probability of observing C2G in DCE

As Figure 35 shows, individuals choosing *C2G* are slightly younger and female. They have an above average income and while membership in minority one increases the probability the opposite is true for minority two. As already seen in 3.3.2 the ownership of a smartphone and a higher income increases the probability of choosing *C2G* and similar the bigger the household the more likely a choice of *C2G*. Again as expected and already seen a higher agreement to the concept and stated likeliness to use C2G will increase the chances of observing the same in the DCE, while preferences for bus, taxi or no preferences for CS decrease that probability. Interestingly preferences for other private modes of transportation have the opposite effect in the DCE and neither prior use nor knowledge of CS services are relevant for the choice. Having a very low opinion on cars or CS in general also reduces the likeliness to choose *C2G* in the DCE.

4.3.2 Full Mixed logit with demographic parameters

Using the relevant variables from the binary logit as only explanatories for the random parameters in the MXL leaves the following utility function:

$$U_{nit}(X_{nit}) = \beta_{1nit} + \beta_{2nit}(-DISTANCE_{nit}) + \beta_{3nit}(-TRAVELTIME_{nit}) + \beta_{4nit}(-PRICE_{nit}) + \varepsilon_{nit}$$

$$\text{with } \beta_{mnit} = \delta_{mi} + \delta_{m1i}AGE_{nt} + \delta_{m2i}GENDER_{nt} + \eta_{mnit} \text{ and } \eta_{mnit} \sim \ln N(0,1) \text{ for } m \in [2,3].$$

$$\text{and } \beta_{mnit} = \delta_{mi} + \delta_{m1i}AGE_{nt} + \delta_{m2i}GENDER_{nt} + INCOMECAT'_{nt}\delta_{m3ci} + \delta_{m4i}SMARTPHONE_{nt} + \eta_{mnit} \text{ and } \eta_{mnit} \sim \ln N(0,1) \text{ for } m = 4.$$

$INCOMECAT'_{nt}$ is in this case the vector of income categories while AGE_{nt}, $GENDER_{nt}$ and $SMARTPHONE_{nt}$ are the in 3.3.2 and 4.3.1 discussed socio-economic parameters per individual. The stated preferences and opinions on cars, CS and C2G in general as used in the binary logit in 4.3.1 were not included in the MXL as these would be endogenous and the error term correlated to the error term of the DCE. Including these variables however in a hybrid model approach should be considered for further research.

Even in this reduced setup the MXL already performs equally to the CLM and MXL including the situational variables, explaining a similar amount of deviation and obtaining a fit of between 17.82% for R^2_{aMcF} and 44.33% for R^2_{ML} as seen in Figure 36. The AIC is with 6 359.6 close to the 6 326.3 of the full CLM and better than the 6 377.6 of full MXL. The log-likelihood is on the same level to the full MXL with -3 159.8.

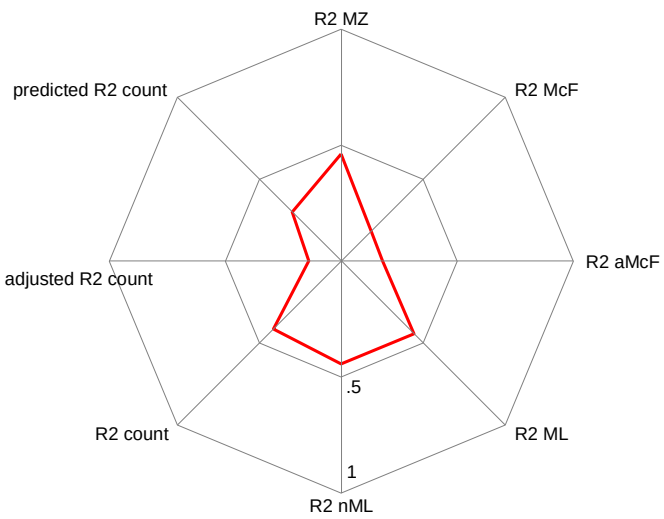


Figure 36: Goodness-of-fit simple MXL with demographics

As Figure 37 shows again all ASC and main effects are statistically relevant at any level, but in contrast to the earlier MXL also the distance seems to include an at ten percent relevant random parameter. Interesting the coefficients of all main effects and therefore their importance seem to increase with a switch from female to male. Similar a higher age indicates a higher importance of *DISTANCE* and *TRAVELTIME*, but reduces the importance of *PRICE.* As to be expected a higher income seemingly decreases the relevance of the pricing coefficient and for individuals owning a smartphone the price has less meaning. But of all these effects only the effect of *AGE* and one *INCOME* category on *PRICE* are actually relevant.

		Estimate	Std. Error	t-value	Pr(>|t|)	
Intercept	BUS	1.6688	0.1163	14.3486	0.0000	***
	METRO	2.1325	0.1175	18.1465	0.0000	***
	TAXI	1.6312	0.3040	5.3652	0.0000	***
	C2G	2.5258	0.1479	17.0808	0.0000	***
Main Effects	DISTANCE	-14.7263	4.9229	-2.9914	0.0028	**
	.GENDER	1.6124	1.1202	1.4394	0.1500	
	.AGE	0.4122	0.7771	0.5304	0.5958	
	TRAVELTIME	-3.6766	0.1607	-22.8807	0.0000	***
	.GENDER	0.0836	0.0532	1.5736	0.1156	
	.AGE	0.0038	0.1094	0.0347	0.9723	
	PRICE	-1.4707	0.2357	-6.2401	0.0000	***
	.GENDER	0.0755	0.0466	1.6183	0.1056	
	.AGE	-0.1872	0.0912	-2.0523	0.0401	**
	.SMARTPHONE	-0.0935	0.1437	-0.6509	0.5151	
	.INCOME2	0.0314	0.1473	0.2133	0.8311	
	.INCOME3	-0.1584	0.1718	-0.9220	0.3565	
	.INCOME4	-1.1170	0.4427	-2.5230	0.0116	**
	.INCOME5	-2.8471	1.8941	-1.5032	0.1328	
	.INCOME6	-2.0384	1.6921	-1.2047	0.2283	
	.INCOME_NA	0.1524	0.1473	1.0347	0.3008	
Std. Dev.	DISTANCE	1.9008	1.0291	1.8470	0.0648	*
	TRAVELTIME	0.0091	0.2274	0.0399	0.9682	
	PRICE	0.5198	0.0762	6.8227	0.0000	***

Figure 37: Simple MXL with demographics

Adapting the utility function to include both the demographic and the situational variables, increases the fit of the model even further to a log-likelihood of -3 110.6 and leaves the following model:

$$U_{nit}(X_{nit}) = \beta_{1nit} + \beta_{2nit}(-DISTANCE_{nit}) + \beta_{3nit}(-TRAVELTIME_{nit}) + \beta_{4nit}(-PRICE_{nit}) + \varepsilon_{nit}$$

with $\beta_{mnit} = \gamma_{mi} + \gamma_{m1i}CITY_{nt} + \gamma_{m2i}RAIN_{nt} + \gamma_{m3i}MORNING_{nt} + \gamma_{m4i}NIGHT_{nt} + \eta_{mnit}$
and $\eta_{mnit} = 0$ for $m \in [11, 14]$

with $\beta_{mnit} = \gamma_{mi} + \gamma_{m2i}RAIN_{nt} + \gamma_{m4i}NIGHT_{nt}$
$+\delta_{m1i}AGE_{nt} + \delta_{m2i}GENDER_{nt} + \eta_{mnit}$
and $\eta_{mnit} \sim \ln N(0,1)$ for $m = 2$

with $\beta_{mnit} = \gamma_{mi} + \gamma_{m1i}CITY_{nt} + \gamma_{m2i}RAIN_{nt} + \gamma_{m3i}MORNING_{nt} + \gamma_{m4i}NIGHT_{nt}$
$+\delta_{m1i}AGE_{nt} + \delta_{m2i}GENDER_{nt} + \eta_{mnit}$
and $\eta_{mnit} \sim \ln N(0,1)$ for $m = 3$

with $\beta_{mnit} = \gamma_{mi} + \gamma_{m1i}CITY_{nt} + \gamma_{m2i}RAIN_{nt} + \gamma_{m3i}MORNING_{nt} + \gamma_{m4i}NIGHT_{nt}$
$+\delta_{m1i}AGE_{nt} + \delta_{m2i}GENDER_{nt} + INCOMECAT'_{nt}\delta_{m3ci}$
$+\delta_{m4i}SMARTPHONE_{nt} + \eta_{mnit}$
and $\eta_{mnit} \sim \ln N(0,1)$ for $m = 4$.

Similar to the log-likelihood also the AIC is decreased to 6 313.2 and the GoF are increased to between 18.38% for R^2_{aMcF} and 46.56% for R^2_{ML} as seen in Figure 40 (Solid red line compared to dotted grey line). Again all ASCs and main effects are statistically relevant at any level and for the interactions and explanatories of the main effect coefficients similar results are found as in the independent MXL for situational parameters and demographic effects. Interestingly in the combined MXL the effect of *RAIN*, *NIGHT* and *GENDER* on the coefficient of *DISTANCE* are also relevant at a ten / five percent level. For the effect of *INCOME* on *PRICE* the results are still challenging as only one categories coefficient is statistically significant and for the category of *INCOME* between 4 000 and 4 999€ the maximum likelihood function seems to be flat therefore allowing no interpretation of that coefficient. Similar to before dropping the interactions between *DISTANCE* and *MORNING* / *PRICE* only very slightly decreases the fit and returns more robust estimations for most coefficients as can be seen in the comparison of Figure 39 and Figure 40.

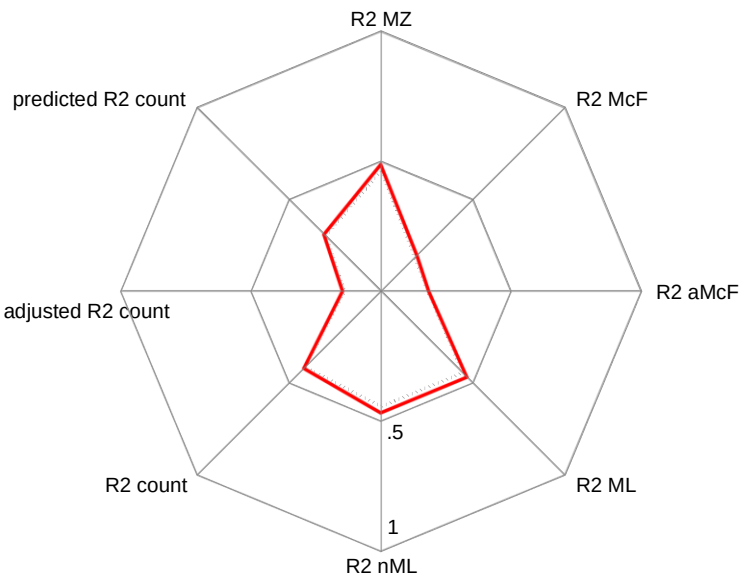


Figure 38: Goodness-of-fit full MXL with demographics

		Estimate	Std. Error	t-value	Pr(>|t|)	
Intercept	BUS	1.7640	0.1959	9.0030	0.0000	***
	METRO	2.1640	0.2116	10.2261	0.0000	***
	TAXI	1.5479	0.5625	2.7520	0.0059	**
	C2G	2.6245	0.2725	9.6316	0.0000	***
City	BUS	0.4248	0.2427	1.7500	0.0801	*
	METRO	0.4910	0.2467	1.9905	0.0465	**
	TAXI	0.0217	0.5950	0.0365	0.9709	
	C2G	0.0669	0.3020	0.2214	0.8248	
Rain	BUS	-0.4989	0.2068	-2.4122	0.0159	**
	METRO	-0.5839	0.2070	-2.8209	0.0048	***
	TAXI	-0.6567	0.5528	-1.1881	0.2348	
	C2G	-0.6861	0.2764	-2.4820	0.0131	**
Morning	BUS	-0.4330	0.3366	-1.2865	0.1983	
	METRO	0.2871	0.2708	1.0600	0.2892	
	TAXI	2.1307	1.1920	1.7875	0.0739	*
	C2G	1.0597	0.5002	2.1188	0.0341	**
Night	BUS	0.1963	0.2387	0.8222	0.4110	
	METRO	0.2572	0.2563	1.0036	0.3156	
	TAXI	0.6258	0.5721	1.0937	0.2741	
	C2G	0.2624	0.2901	0.9044	0.3658	
Main Effects	DISTANCE	-9.3909	0.9388	-10.0032	0.0000	***
	.CITY	-1.6803	1.6940	-0.9919	0.3212	
	.RAIN	0.7026	0.3840	1.8295	0.0673	*
	.MORNING	-3.7938	151.0329	-0.0417	0.9932	
	.NIGHT	0.7469	0.3651	2.0459	0.0408	**
	.GENDER	0.5760	0.2146	2.6842	0.0073	***
	.AGE	-0.0035	0.3051	-0.0115	0.9908	
	TRAVELTIME	-3.8628	0.2581	-14.9690	0.0000	***
	.CITY	0.3516	0.1813	1.9397	0.0524	*
	.RAIN	-0.8123	0.1959	-4.1453	0.0000	***
	.MORNING	0.4263	0.2677	1.5926	0.1112	
	.NIGHT	0.6905	0.2068	3.3399	0.0008	***
	.GENDER	0.0886	0.0575	1.5403	0.1235	
	.AGE	0.1151	0.1147	1.0027	0.3160	
	PRICE	-1.2770	0.2684	-4.7585	0.0000	***
	.CITY	-0.4078	0.2137	-1.9079	0.0564	*
	.RAIN	-0.3101	0.1833	-1.6921	0.0906	*
	.MORNING	0.7910	0.2590	3.0540	0.0023	***
	.NIGHT	0.1902	0.1978	0.9615	0.3363	
	.GENDER	0.0863	0.0437	1.9754	0.0482	**
	.AGE	-0.1513	0.0848	-1.7831	0.0746	*
	.SMARTPHONE	-0.1597	0.1274	-1.2535	0.2100	
	.INCOME2	-0.0127	0.1370	-0.0925	0.9263	
	.INCOME3	-0.2556	0.1659	-1.5406	0.1234	
	.INCOME4	-1.0734	0.4143	-2.5911	0.0096	***
	.INCOME5	-6.3038	27.2131	-0.2316	0.8168	
	.INCOME6	-1.2707	0.9157	-1.3877	0.1652	
	.INCOME_NA	0.0205	0.1398	0.1466	0.8835	
Std. Dev.	DISTANCE	0.0870	2.4100	0.0361	0.9712	
	TRAVELTIME	0.0093	0.2743	0.0338	0.9730	
	PRICE	0.4358	0.0750	5.8084	0.0000	***

Figure 39: Full MXL with demographics

		Estimate	Std. Error	t-value	Pr(>|t|)	Full MXL	Delta	in %
Intercept	BUS	1.6622	0.1934	8.5963	0.0000 ***	1.7640	-0.1018	-5.77%
	METRO	2.0878	0.2121	9.8411	0.0000 ***	2.1640	-0.0763	-3.52%
	TAXI	1.7014	0.3598	4.7282	0.0000 ***	1.5479	0.1535	9.92%
	C2G	2.6115	0.2461	10.6128	0.0000 ***	2.6245	-0.0130	-0.50%
City	BUS	0.5638	0.2356	2.3932	0.0167 **	0.4248	0.1390	32.73%
	METRO	0.6473	0.2395	2.7029	0.0069 ***	0.4910	0.1563	31.83%
	TAXI	-0.1888	0.5333	-0.3540	0.7233	0.0217	-0.2105	-968.90%
	C2G	0.0981	0.2935	0.3343	0.7381	0.0669	0.0312	46.72%
Rain	BUS	-0.4062	0.2156	-1.8842	0.0595 *	-0.4989	0.0927	-18.58%
	METRO	-0.5161	0.2180	-2.3671	0.0179 **	-0.5839	0.0678	-11.62%
	TAXI	-0.7834	0.5447	-1.4383	0.1503	-0.6567	-0.1267	19.29%
	C2G	-0.6676	0.2813	-2.3736	0.0176 **	-0.6861	0.0185	-2.70%
Morning	BUS	-0.2658	0.3769	-0.7053	0.4806	-0.4330	0.1671	-38.60%
	METRO	0.3785	0.2755	1.3739	0.1695	0.2871	0.0914	31.85%
	TAXI	1.7995	1.1928	1.5087	0.1314	2.1307	-0.3311	-15.54%
	C2G	0.9927	0.4709	2.1079	0.0350 **	1.0597	-0.0670	-6.33%
Night	BUS	0.1578	0.2359	0.6692	0.5034	0.1963	-0.0384	-19.59%
	METRO	0.2011	0.2525	0.7962	0.4259	0.2572	-0.0561	-21.83%
	TAXI	0.5740	0.5689	1.0090	0.3130	0.6258	-0.0518	-8.28%
	C2G	0.1909	0.2897	0.6589	0.5100	0.2624	-0.0715	-27.25%
Main Effects	DISTANCE	-9.8961	0.7659	-12.9206	0.0000 ***	-9.3909	-0.5052	5.38%
	.CITY	-	-	-	-	-1.6803		
	.RAIN	1.6104	0.8018	2.0085	0.0446 **	0.7026	0.9078	129.21%
	.MORNING	-	-	-	-	-3.7938		
	.NIGHT	0.6639	0.3798	1.7481	0.0805 *	0.7469	-0.0830	-11.11%
	.GENDER	0.4856	0.1779	2.7292	0.0063 ***	0.5760	-0.0904	-15.69%
	.AGE	-0.3162	0.3164	-0.9993	0.3177	-0.0035	-0.3127	8926.08%
	TRAVELTIME	-3.8748	0.2468	-15.6995	0.0000 ***	-3.8628	-0.0119	0.31%
	.CITY	0.3501	0.1748	2.0026	0.0452 **	0.3516	-0.0015	-0.42%
	.RAIN	-0.8438	0.1999	-4.2213	0.0000 ***	-0.8123	-0.0315	3.88%
	.MORNING	0.5482	0.2526	2.1705	0.0300 **	0.4263	0.1219	28.59%
	.NIGHT	0.6351	0.2044	3.1081	0.0019 ***	0.6905	-0.0554	-8.02%
	.GENDER	0.0900	0.0578	1.5580	0.1192	0.0886	0.0014	1.59%
	.AGE	0.1551	0.1150	1.3485	0.1775	0.1151	0.0401	34.83%
	PRICE	-1.2460	0.2650	-4.7024	0.0000 ***	-1.2770	0.0310	-2.43%
	.CITY	-0.4939	0.2319	-2.1295	0.0332 **	-0.4078	-0.0860	21.10%
	.RAIN	-0.3482	0.1905	-1.8281	0.0675 *	-0.3101	-0.0381	12.28%
	.MORNING	0.6716	0.2657	2.5279	0.0115 **	0.7910	-0.1194	-15.09%
	.NIGHT	0.1863	0.1946	0.9575	0.3383	0.1902	-0.0039	-2.03%
	.GENDER	0.0898	0.0442	2.0329	0.0421 **	0.0863	0.0035	4.05%
	.AGE	-0.1612	0.0861	-1.8722	0.0612 *	-0.1513	-0.0099	6.56%
	.SMARTPHONE	-0.1545	0.1216	-1.2712	0.2037	-0.1597	0.0052	-3.24%
	.INCOME2	-0.0026	0.1379	-0.0192	0.9847	-0.0127	0.0100	-79.09%
	.INCOME3	-0.2378	0.1662	-1.4315	0.1523	-0.2556	0.0178	-6.96%
	.INCOME4	-1.0771	0.4109	-2.6214	0.0088 ***	-1.0734	-0.0037	0.35%
	.INCOME5	-6.3020	27.2131	-0.2316	0.8168	-6.3038	0.0018	-0.03%
	.INCOME6	-1.2961	0.9603	-1.3497	0.1771	-1.2707	-0.0254	2.00%
	.INCOME_NA	0.0278	0.1409	0.1972	0.8437	0.0205	0.0073	35.58%
Std. Dev.	DISTANCE	0.0072	0.5571	0.0130	0.9896	0.0870	-0.0797	-91.67%
	TRAVELTIME	0.0226	0.3028	0.0746	0.9405	0.0093	0.0133	143.54%
	PRICE	0.4290	0.0734	5.8442	0.0000 ***	0.4358	-0.0067	-1.54%

Figure 40: Full MXL with demographics without interaction DISTANCE with MORNING / CITY

5. Results

5.1 Effects of situational and individual attributes on valuation of travel time

Figure 43 wraps up the results of the in 4.1 to 4.3 developed models and following Train (2009) converts the log-normal distributed coefficients of *DISTANCE*, *TRAVELTIME* and *PRICE* within the MXLs back for better comparison to the CLMs by applying

$$\tilde{\beta}_m = -\exp(\overline{\beta_m}) \quad \text{and} \quad \tilde{\sigma}_m = \left(\exp(2\overline{\beta_m} + \sigma_m)(\exp(\sigma_m) - 1)\right)^{\frac{1}{2}}.$$

The best fit is reached by including situational and socio-economic parameters in the MXL. While adding only demographics will increase the overall fit to a similar level as adding only the situational factors, the predictive quality of the model in terms of R^2_{count} is slightly lower as there is a bigger part of randomness in the coefficients of the MXL.

		Simple CLM		Full CLM		Simple MXL		Full MXL		Simple MXL incl. demographics		Full MXL incl. demographics	
Intercept	BUS	1.47470	***	1.53952	***	1.60986	***	1.67562	***	1.66876	***	1.66216	***
	METRO	1.94487	***	2.02073	***	2.07570	***	2.10174	***	2.13246	***	2.08775	***
	TAXI	1.17937	***	1.61299	***	1.53669	***	2.28096	***	1.63122	***	1.70142	***
	C2G	2.06629	***	2.26714	***	2.40295	***	2.78506	***	2.52576	***	2.61147	***
City	BUS	-		0.69990	***	-		0.54147	**	-		0.56375	**
	METRO	-		0.75659	***	-		0.60118	**	-		0.64733	***
	TAXI	-		-0.36278		-		-0.66926		-		-0.18879	
	C2G	-		0.25498		-		-0.05354		-		0.09812	
Rain	BUS	-		-0.25807		-		-0.37685	*	-		-0.40624	*
	METRO	-		-0.38179	*	-		-0.48289	**	-		-0.51608	**
	TAXI	-		-1.20551	*	-		-1.12856	*	-		-0.78342	
	C2G	-		-0.62354	**	-		-0.76487	**	-		-0.66761	**
Morning	BUS	-		-0.63058		-		-0.29961		-		-0.26584	
	METRO	-		0.01489		-		0.26959		-		0.37853	
	TAXI	-		3.23406	**	-		0.85074		-		1.79953	
	C2G	-		0.76805	**	-		0.55387		-		0.99269	**
Night	BUS	-		-0.17454		-		0.03614		-		0.15783	
	METRO	-		-0.20453		-		0.08447		-		0.20106	
	TAXI	-		0.71276		-		0.30530		-		0.57396	
	C2G	-		0.07275		-		0.00625		-		0.19090	
Main Effects	DISTANCE (in 100m)	-0.02894	***	-0.00779		-0.02562	***	-0.00900	***	-0.00004	**	-0.00504	***
	.CITY	-		-		-		-		-		-	
	.RAIN	-		-0.07957	***	-		-0.07816		-		-0.02017	**
	.MORNING	-		-		-		-		-		-	
	.NIGHT	-		-0.06060	**	-		-0.00490		-		-0.00475	*
	.GENDER	-		-		-		-		-0.00016		-0.00315	***
	.AGE	-		-		-		-		-0.00002		0.00137	
	STD. DEV. (LN-NORMAL)	-		-		0.00113		0.00023		0.00025	*	0.00043	
	TRAVELTIME (in Min)	-0.03161	***	-0.03448	***	-0.03089	***	-0.02711	***	-0.02531	***	-0.02076	***
	.CITY	-		-0.01092	*	-		-0.01336	**	-		-0.00870	**
	.RAIN	-		0.01919	***	-		0.01525	***	-		0.01183	***
	.MORNING	-		-0.00491		-		-0.02486	***	-		-0.01516	**
	.NIGHT	-		-0.00260		-		-0.02215	***	-		-0.01842	***
	.GENDER	-		-		-		-		-0.00221		-0.00196	
	.AGE	-		-		-		-		-0.00010		-0.00348	
	STD. DEV. (LN-NORMAL)	-		-		0.00072		0.00050		0.00243		0.00317	
	PRICE (in €)	-0.16446	***	-0.21449	***	-0.21254	***	-0.31275	***	-0.22977	***	-0.28766	***
	.CITY	-		0.08979	***	-		0.15997	**	-		0.11211	**
	.RAIN	-		0.08812	**	-		0.11666	**	-		0.08458	*
	.MORNING	-		-0.20231	***	-		-0.14259		-		-0.27542	**
	.NIGHT	-		-0.06384	*	-		-0.02812		-		-0.05892	
	.GENDER	-		-		-		-		-0.01801		-0.02704	**
	.AGE	-		-		-		-		0.03923	**	0.04282	*
	.SMARTPHONE	-		-		-		-		0.02052		0.04119	
	.INCOME2	-		-		-		-		-0.00733		0.00076	
	.INCOME3	-		-		-		-		0.03367		0.06089	
	.INCOME4	-		-		-		-		0.15458	**	0.18969	***
	.INCOME5	-		-		-		-		0.21644		0.28713	
	.INCOME6	-		-		-		-		0.19985		0.20896	
	.INCOME7	-		-		-		-		-0.03784		-0.00810	
	STD. DEV. (LN-NORMAL)	-		-		0.24882	***	0.31724	***	0.24603	***	0.26093	***
	Log-likelihood	-3213.54		-3159.15		-3203.60		-3153.80		-3159.80		-3110.60	
	AIC	6435.08		6326.30		6415.20		6377.60		6359.60		6313.20	
	Adjusted McFadden R2	0.16727		0.18135		0.16984		0.17549		0.17782		0.18382	
	Adjusted ML R2	0.41779		0.44359		0.42259		0.44607		0.44329		0.46565	
	Count R2	0.43500		0.44042		0.43208		0.42542		0.41292		0.43417	

Figure 41: Summary of estimates

Almost all interactions of the situational variables with the main effects and of *GENDER* and *AGE* with *PRICE* are statistically relevant in the full MXL with demographics. This and the significant random part of the coefficient of *PRICE* will influence the individual WTP in a certain situation greatly and is a sign of the heterogeneity in assessment of VoT. Applying the definition of the WTP as the ratio between the coefficients of the variable in question over *PRICE* to estimate the VoT in the simple CLM returns

$$\widehat{WTP}_{DISTANCE} = \frac{\hat{\beta}_2}{\hat{\beta}_4} = 0.1760\frac{EUR}{100m} \approx 0.1478\frac{EUR}{Min} \quad \text{and}$$

$$\widehat{WTP}_{TRAVELTIME} = \frac{\hat{\beta}_3}{\hat{\beta}_4} = 0.1922\frac{EUR}{Min}.$$

While in this case the $\widehat{WTP}_{TRAVELTIME}$ is above Wardmans (2011) estimates for the UK, the VoT is in the range of predictions by Shires and de Jong (2009). The VoT for walking distance is however not higher than the VoT for in-mode travel time but slightly lower, which does not fit expectations as seen in 2.1.

Differently using the full CLM the $\widehat{WTP}_{DISTANCE}$ ranges from $0.0363\frac{EUR}{100m} \approx 0.0305\frac{EUR}{Min}$ for *BASE* up to $0.2854\frac{EUR}{100m} \approx 0.2678\frac{EUR}{Min}$ for *NIGHT* and $0.3710\frac{EUR}{100m} \approx 0.3421\frac{EUR}{Min}$ for *RAIN* compared to a $\widehat{WTP}_{TRAVELTIME}$ between $0.1608\frac{EUR}{Min}$ for *BASE* / $0.1729\frac{EUR}{Min}$ for *NIGHT* and $0.0713\frac{EUR}{Min}$ for *RAIN*. Therefore in this case the VoT for walking distances is way higher compared to in-mode travel time if the situational factors are negative, while dropping massively if not so.

Allowing the coefficients to include a random parameter, the WTP cannot be stated as a single value anymore, but follows a distribution

$$Dist(\widehat{WTP}_{TravT}) = \frac{-\exp(N(\hat{\beta}_3, \hat{\sigma}_3)}{-\exp(N(\hat{\beta}_4, \hat{\sigma}_4)}.$$

The distribution of the WTP was generated by simulating for each individual its WTP within a certain scenario using the estimated coefficients and standard errors. For the BASE Scenario in the full MXL all demographic variables were kept at their original value and the situational variables were set to zero.

Again the simple model reveals a smaller VoT for walking versus in-mode travel times as seen in Figure 42 while the full MXL including the situational parameters reveals a higher VoT for walking distances for *RAIN* but not for *NIGHT* as seen in Figure 43. Also the spread of both distributions is increased by adding the situational variables, where *DISTANCE* shows more heterogeneity than *TRAVELTIME*. The WTP of *DISTANCE* is mostly positively affected by *RAIN* and *NIGHT*. For *TRAVELTIME* the origin of travel in the *CITY* and the time of day *NIGHT* increase the WTP almost by factor two. While *MORNING* has a less prominent but also positive impact, *RAIN* reduces the WTP slightly.

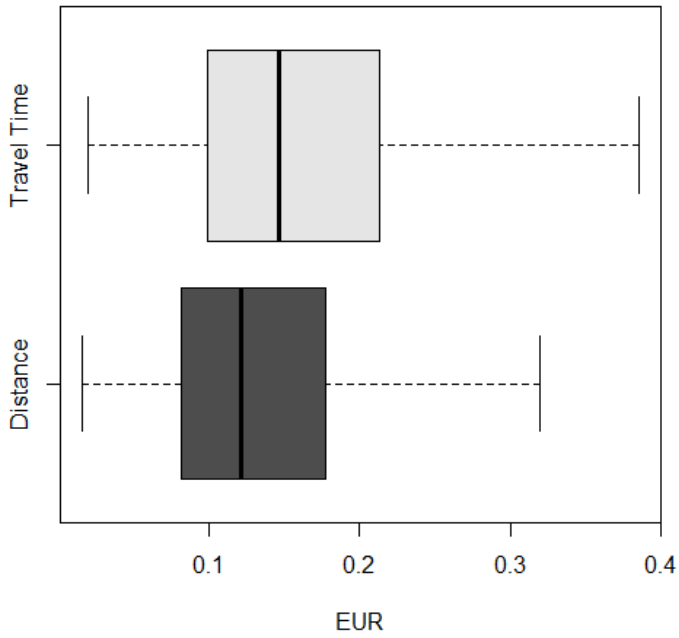


Figure 42: VoT for simple MXL

Adding the individual heterogeneity reduces the impact of *DISTANCE* even further and under most combinations of demographic and situational parameters as seen in Figure 44 to Figure 46 the *DISTANCE* seems not to be identifiable within the given data set. Only with *RAIN* in the *BASE* and *NIGHT* scenario the expected behavior of a higher value for walking than in-mode time is observed. While the direction and relative impact of effects are very similar between the MXL without and with the demographics, the importance of *DISTANCE* approaches zero without *RAIN* in the later model while *TRAVELTIME* explains almost all of the observed behavior.

Similar the general relevance of *PRICE* is reduced by adding the *INCOME* categories and goes for incomes above 3 000€ down by two third. While this is the extreme case, the importance of *PRICE* is relatively to *TRAVELTIME* and the intercepts in general reduced. Therefore comparing both MXL the VoT *TRAVELTIME* benefits in the model with demographics both from the reduction of importance of *DISTANCE* and *PRICE*.

The direction of effects between the variables on the WTP are other than that as expected: A higher income increases the WTP for savings in both walking and in-mode time. Interestingly older individuals are estimated to have a higher VoT for in-mode time compared to younger while at the same time valuing walking time less. Males or females appear to value in-mode time similarly, but men seem to be willing to pay slightly more to reduce walking times than women. In total there is a wide heterogeneity of VoT for both the situational, socio-economic and unobserved random parameters, but with the number of observations and the complexity of the model, the standard errors of the estimates are too high to return robust results for the combined analysis. Future research should therefore include more choice tasks per individual and more observations in general to disentangle the observed effects. Especially for the value of savings in walking distance a closer and more detailed understanding seems to be necessary. To this moment only the existence of heterogeneity can be inferred.

WTP DISTANCE
[in EUR / 100m]

WTP TRAVEL TIME
[in EUR / Min.]

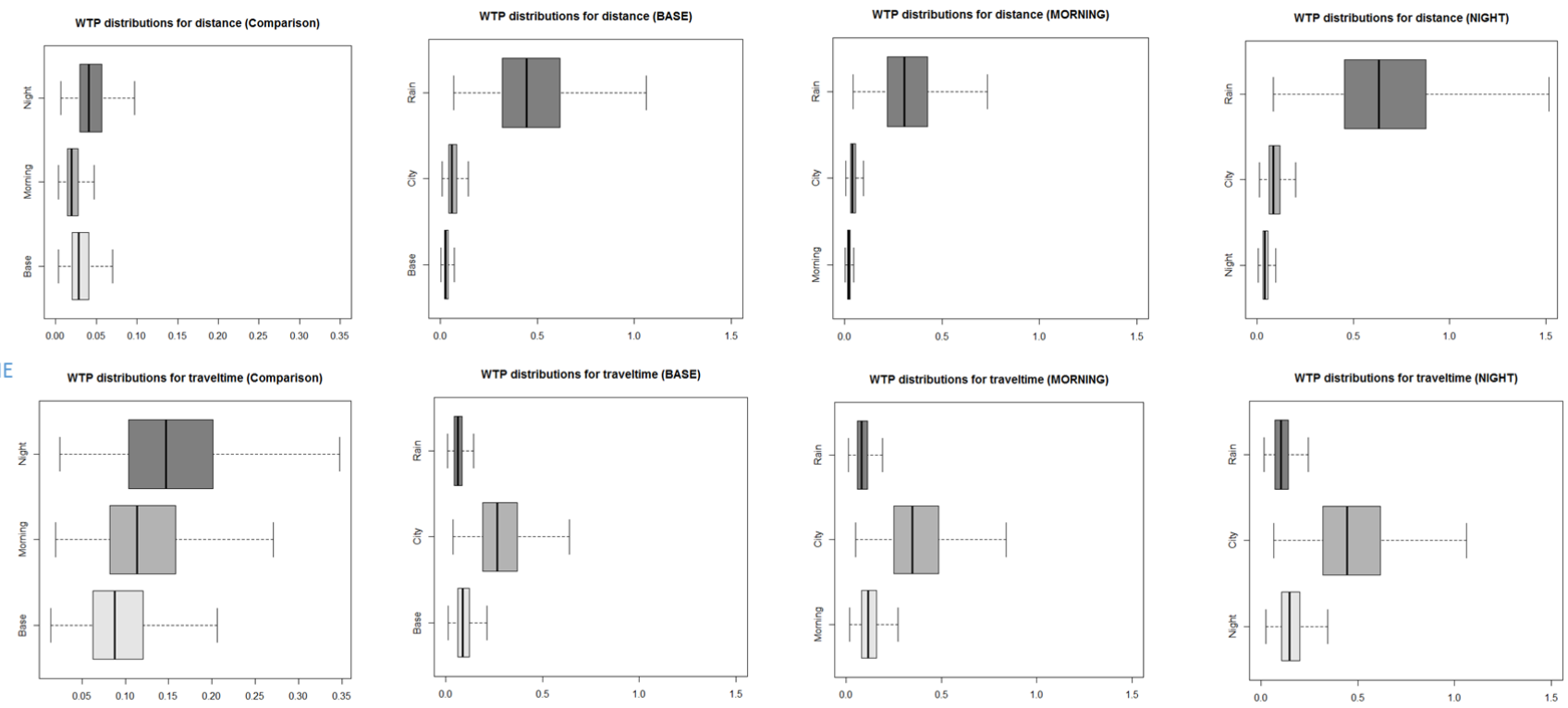


Figure 43: VoT for full MXL

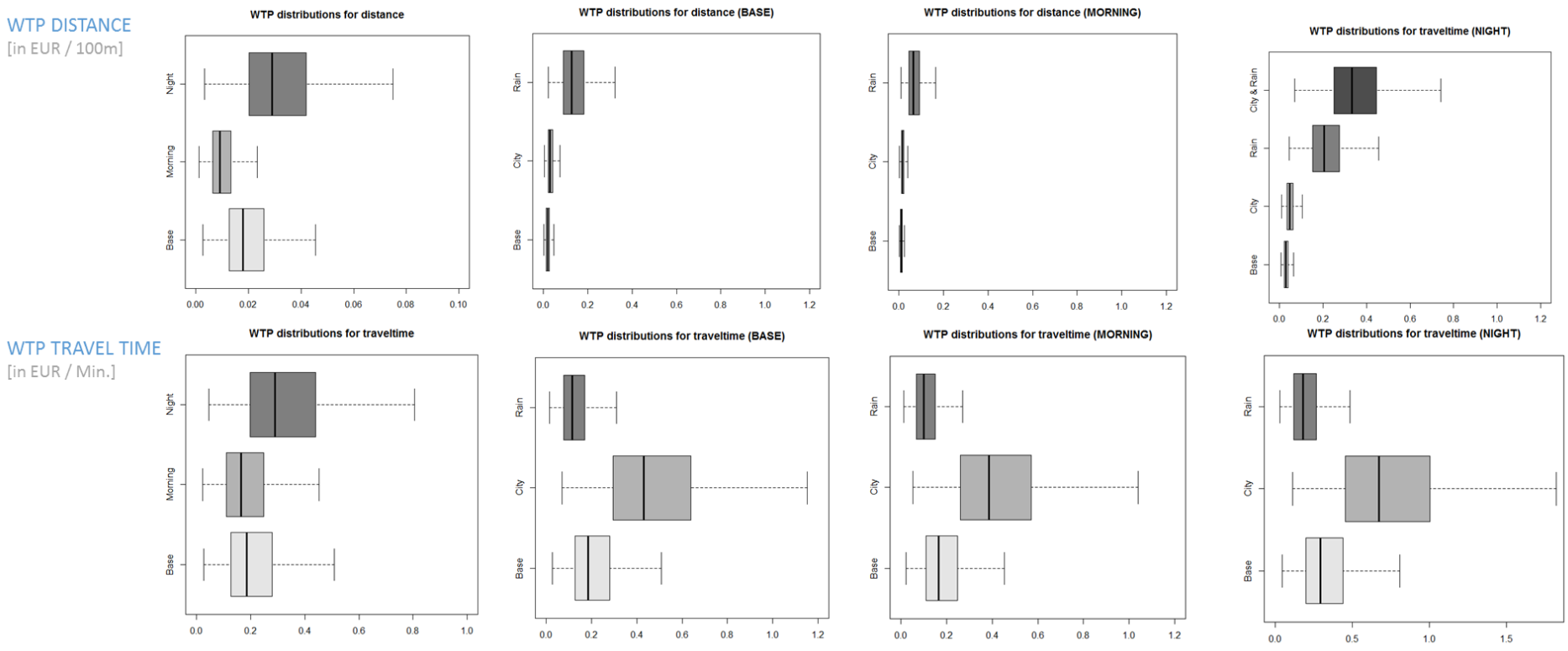


Figure 44: VoT for full MXL – Situation

WTP DISTANCE
[in EUR / 100m]

WTP TRAVEL TIME
[in EUR / Min.]

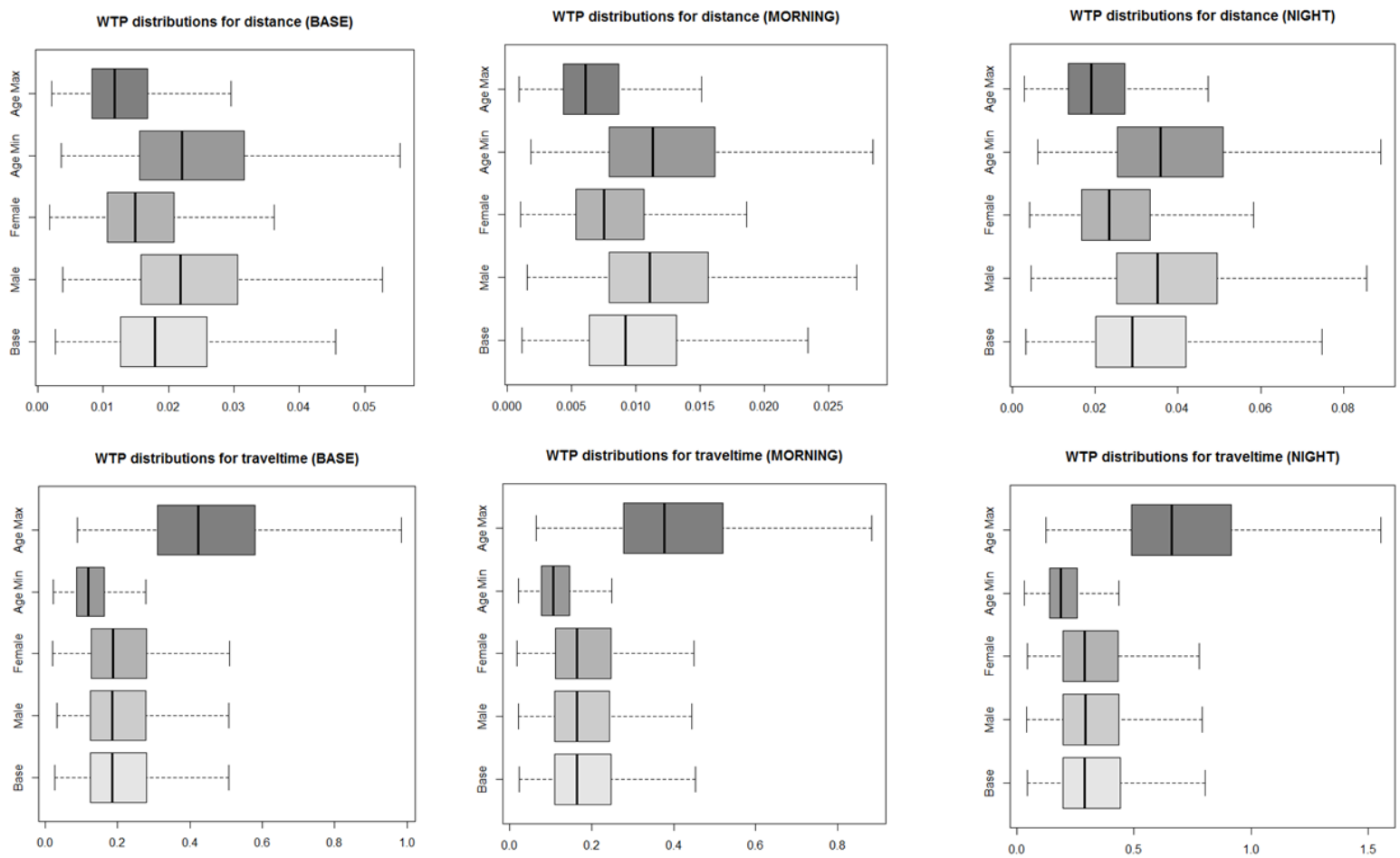


Figure 45: VoT for full MXL – GENDER and AGE

WTP DISTANCE
[in EUR / 100m]

WTP TRAVEL TIME
[in EUR / Min.]

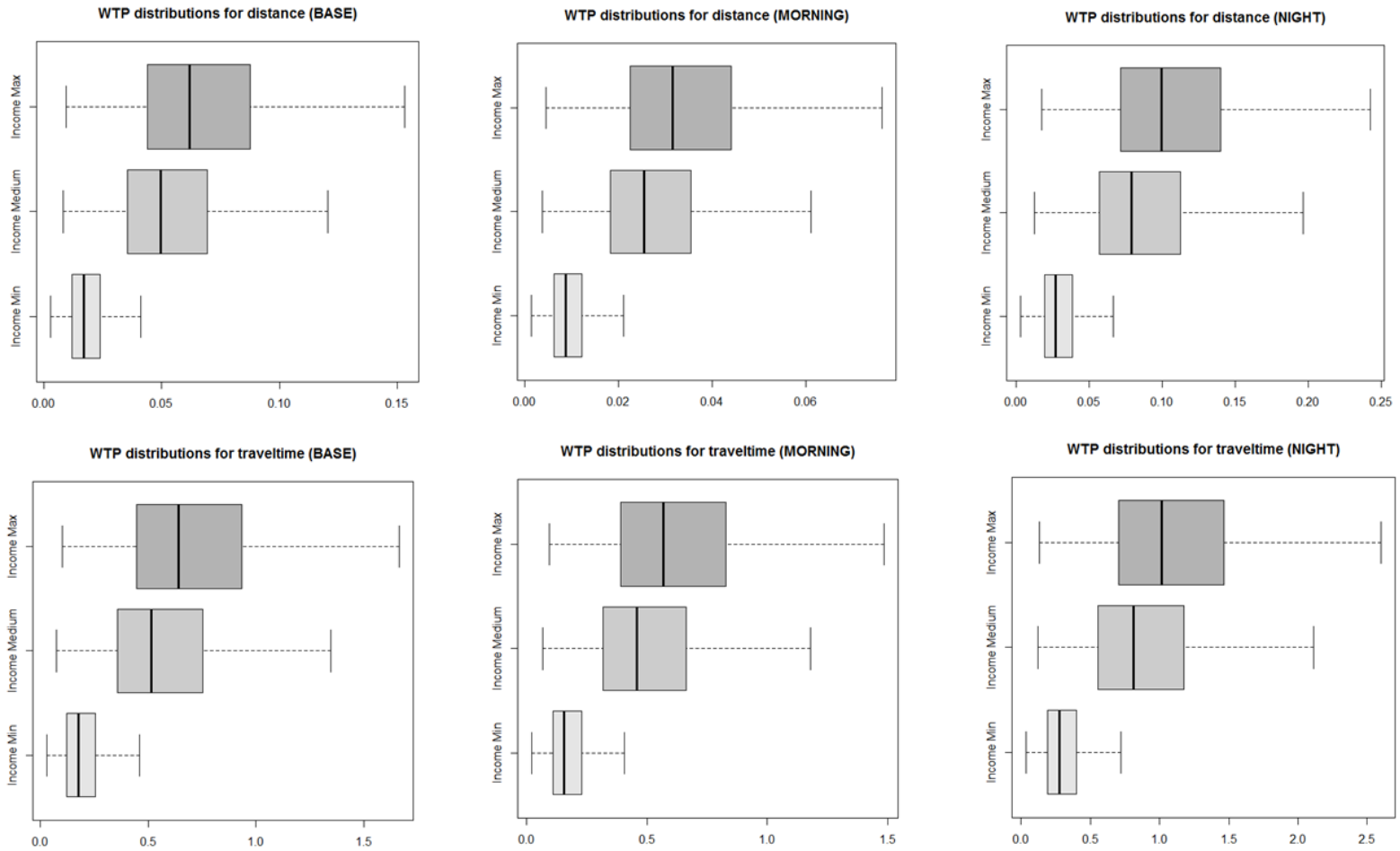


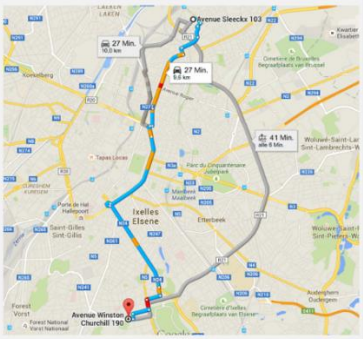
Figure 46: VoT for full MXL – INCOME

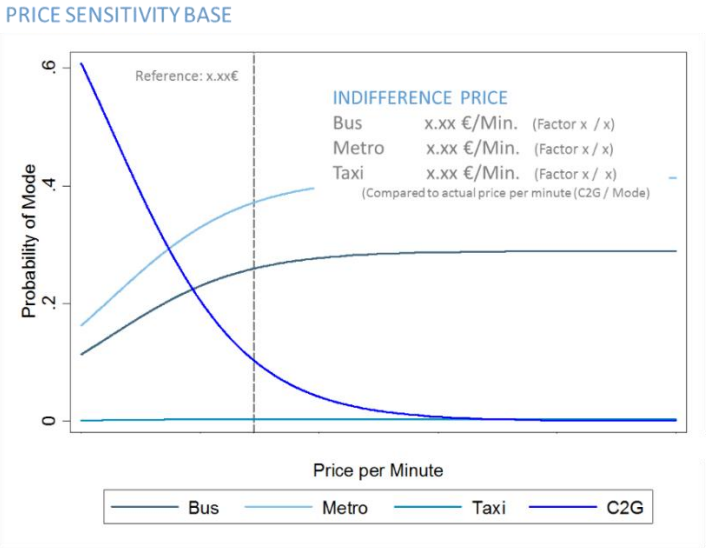
5.2 Pricing effects and optimization problem

Using the full MXL without demographics to estimate indifference prices and market shares under certain scenarios allows for optimal pricing within specified target groups. The indifference price was calculated by simulating different utilities over all individuals and estimating at which price the mean of the population would be indifferent between two alternative transport modes. All simulations were done by varying one variable and keeping all other variables either at a specified scenario level (parameters), the original value (demographics) or at zero (situation). As these information are private property of moovel the figures cannot be published in this paper, but a short synapse of findings will be presented.

As with the general WTP for *DISTANCE* and *TRAVELTIME* the utilities of the different alternative modes are highly impacted by the situational variables. As to be seen for example in Figure 47 the probability of choosing C2G drops faster for the *MORNING* versus the *BASE* scenario. While in this case again the correlation between *TRAVELTIME*, *PRICE* and *MORNING* needs to be taken into account, similar and more product related effects occur also in the comparison of other scenarios and when changing the demographic variables.

Scenario 2	Walking Distance*	Travel Time	Price
Bus	300m	41 Min.	
Metro	400m	41 Min.	
Taxi	-	27 Min.	
Car2Go	450m	27 Min. + 5 Min. for Parking	





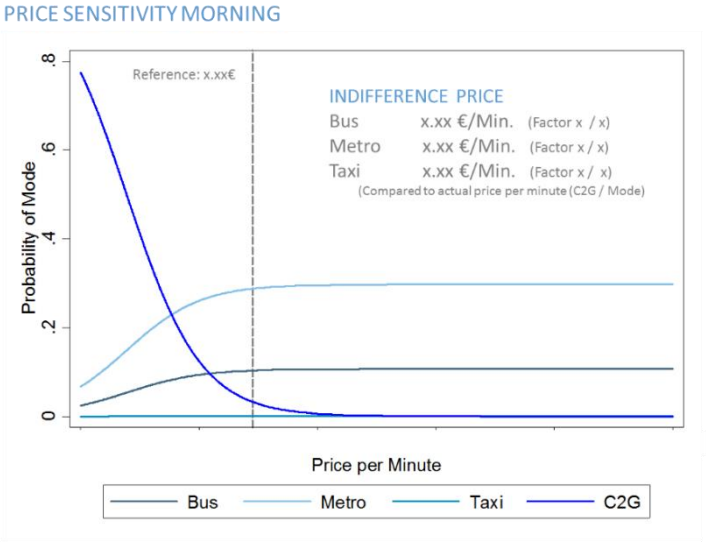


Figure 47: Comparison Indifference price under different situational parameters

While the simulations were used to compare different pricing effects on the market share of each alternative another analysis looked at the impact of travel distances on the choice of transport mode revealing that with increasing *TRAVELTIME* the probability of *BUS* and *METRO* will increase due to the comparison of fixed prices (*METRO*, *BUS*) versus relative prices (*C2G*, *TAXI*).

To obtain an optimal pricing the definition of the target group should therefore be taken into account and as seen above group, situation or even individual focused prices could be used to increase market shares and revenues as a significant heterogeneity in prices occurs. Also estimating separated MXL for different levels of agreement to C2G from the stated preferences part of the survey showed significant differences in market shares and possible indifference prices between these groups. As mentioned before a further step of analysis should therefore be to use hybrid models to deepen the understanding of the impact of opinions on the pricing willingness.

All estimations of market shares and indifference prices were based on the assumption that the comparable target group was in a similar situation to the respondents in the DCE. This means all alternatives and parameters are known and the user is registered at C2G with a direct possibility to access the car

5.3 Caveats

While both research questions could be answered using the above model, there are some caveats to the survey design and models used. The correlation between the situational variable *MORNING* and the higher *TRAVELTIMES* and *PRICES* due to the underlying assumption of rush hour created a situation in which the interaction between *DISTANCE* and *MORNING* was not perfectly identifiable. While the Leeds group approach of using value boundaries and restrictions as discussed in 2.3.3 creates realistic choice situations, it also impacts strongly the statistical performance of the model and the problem of inducing correlations should be taken into account.

Also the access time should have been presented more prominent in the survey design. It seems as if it was only taken into account by respondents, if the situational variable *RAIN* or *NIGHT* were active. For further iterations the access time should therefore be only represented as distance and separately from the in-mode time instead of as “thereof”. In general, there were too many changes between designed model for the survey and the actual model used, which resulted in non-retrievable coefficients for some parameter interactions. While the original model was designed to differentiate between access, in-mode and egress travel time, in the end only total travel time and thereof access time were included. Also the original model was missing a non-linear term, which was later seen to be statistically relevant, and was only designed for CLM. While the later was at least tested upfront using simulation, the model was not designed at all to include demographics.

A further problem of using MXL models seems to be the usage of GoF indicators. While in this study several measures were applied the results were mixed for the MXL and certain measures as i.e. the variance ratio seemed to not deal well with the random part of the coefficients. Also, throughout the investigation the typical adjusted McFadden measure reported lower fits than all other measures. As the results were mainly used to predict behavior and estimate optimal prices the reached fit for the count R^2 seems slightly disappointing. A R^2_{count} of 45% seems in this case not very high and looking at the adjusted R^2_{count} with a maximum of 17% and the predicted R^2_{count} with a maximum of 32%, there is still a large unexplained part within the estimations. While it would for example be interesting to also see the interaction of the socio-demographic parameters with the different situational factors, the model design and obtained number of observations seem to not allow for deeper analysis and identification of more interaction effects.

6. Conclusion and Outlook

The aim of this paper was to identify those factors, that influence the probability of choosing a certain mode of transportation under different circumstances and socio-economic conditions, and to assess the individual heterogeneity in creating an optimal pricing for urban transportation services.

Using surveyed data of 800 individuals in a DCE with three choice task per individual and five labeled alternatives, it could be shown, that the choice probability and therefore utility of alternatives was clearly different for the time of day, the weather and the point of origin. Also it could be seen that the parameters of the alternatives were valued differently in each scenario. If it rains or it is night the access distance was valued higher compared to within rush hour or at midday with sun where the in-mode travel time was valued above.

Using a conditional logit model VoT of 0.15 €/Min. to 0.18 €/Min. were retrieved, which is similar to comparable studies within Europe. But by allowing the VoT to differ for situations, the range was increased to 0.03 €/Min. to 0.34 €/Min. for access distance and 0.07 €/Min. to 0.17 €/Min. for in-mode travel time. Switching to MXL and adding individual socio-economic parameters and a random parameter to the main effects increased this range even further, showing that the VoT and substitution effect between walking and in-mode travel time are very dependent on the situation, the socio-economic background like age, gender and income of the individual and unobserved random parameters.

Simulating based on these estimates the different market shares under flexible prices showed a high impact of situational and individual parameters on the indifference prices between alternatives. While for example public transport was strong for certain demographics, in other groups the C2G indifference prices were much lower. Separated MXL for different levels of the stated agreement to the concept of C2G showed also a significant difference between groups and optimal prices. Therefore decision makers from urban transport modes should consider to adopt their prices to situation and individual to increase market shares and revenues.

While all used models were tested using different GoF indicators there is not enough literature on the performance and interpretation of certain measures within MXLs. While the adjusted McFadden R^2 stated a fit between 16.7% and 18.3% the adjusted maximum likelihood R^2 ranged from 41.8% to 46.6%. Even more troubling was the high increase of the Ratio of Variance R^2 when switching to MXL. Further research should therefore be done to analysis the efficiency and robustness of these GoF measures within MXL applications.

Also additional research on the influence of situational variables on the walking distance and difference in the assessment of respondents to walking, in-mode and transit activities should be taken into consideration. While this study revealed, that there is an impact of the situation and that this impact differs for different activities, due to the design not all effects could be sufficiently identified. Further more detailed designs and a more prominent representation of the different activities should help to improve the retrieval of these valuations. Similar a hybrid model approach could reveal more details on the choice process, as the separated MXL and a binary logit on the probability of observing a C2G choice indicate a high influence of opinions and psychological preferences on the consumer decision.

IV. Literature

Akaike, H. (1973). Information Theory and and extension of the maximum likelihood principle. In B. C. Petrov, *Second international symposium on information theory* (pp. 267-281). Budapest: Akademiai Kiado.

Algers, S. B. (1998). Mixed logit estimation of the value of travel time. *Tech. Rep., Uppsala working paper series*(1998: 15). Retrieved from <http://ideas.repec.org/p/fth/uppaal/199815.html>

Amador, F. G. (2005). Preference heterogeneity and willingness to pay for travel time savings. *Transportation*(32), pp. 627–647.

Assessing the Impacts of Alternative 'Opt-out' Formats in Choice Experiment Studies: Consumer Preferences for Genetically Modified Content and Production Information in Food. (2003). *Agricultural Policy Res.*(5), 1-43.

Baltas, G. D. (1998). A flexible model for consumer choice in packaged goods categories. *Journal of the Market Research Society*(40), pp. 141-153.

Baltas, G. D. (2001). Random utility models in marketing research: a survey. *Journal of Business Research*(51), pp. 115-125.

Beesley, M. (1973). *Urban Transport: Studies in Economic Policy.* London: Buttersworths.

Bekhor, S. B.-A. (2002). Adaptation of logit kernel to route choice situations. . *Transportation Research Record*(1805), pp. 78–85.

Ben-Akiva, M. B. (2001). Specification, identification, and estimation of the logit kernel (or continuous mixed logit) model. *Tech. rep., Department of Civil Engineering, MIT*. Cambridge, MA.

Ben-Akiva, M. L. (1985). *Discrete choice analysis: Theory and application to travel demand.* Cambridge, MA: MIT Press.

Bhat, C. (2001). Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research, Part B*, pp. 677–693.

Bhat, C. C. (2002). A unified mixed logit framework for modeling revealed and stated preferences: Formulation and application to congestion pricing analysis in the San Francisco Bay area. *Transportation Research*(36), pp. 593–616.

Bishop, Y. F. (1975). *Discrete multivariate analysis: Theory and Practice.* Cambridge, MA: MIT Press.

Bliemer, M. R. (2005). Efficiency and sample size requirements for stated choice studies. *working paper, Institute of Transportation Studies, University of Sydney*.

Bliemer, M. R. (2009). Designing stated choice experiments: the state of art. In R. Y. Kitamura, *The Expending Sphere of Travel Behaviour Research, Selected Papers from the 11th International Conference on Travel Behaviour Research* (pp. 495-508). Bingley: Emerald.

Bliemer, M. R. (2010). Construction of experimental designs for mixed logit models allowing for correlation acrosse choice observations. *Transportation Resarch, Part B*(46), pp. 720-734.

Bliemer, M. R. (2011). Experimental design influences on stated choice outputs: an empirical study in air travel choice. *Transportation Research*(45), pp. 63-79.

Browning, R. B. (2005). Effects of obesity and sex on the energetic cost and preferred speed of walking. *Journal of Applied Physiology*(100), 390–398.

Brownstone, D. B. (2000). Joint mixed logit models of stated and revealed preferences for alternative-fuel vehicles. *Transportation Research*(34), pp. 315–338.

Brownstone, D. S. (2005). Valuing time and reliability: Assessing the evidence from road pricing demonstrations. . *Transportation Research*(39), pp. 279–293.

Bunch, D. L. (1996). A comparison of experimental design strategies for multinomial logit models: the case of generic attributes. *working paper, Graduate School of Management, University of California at Davis*.

Burge, P. R. (2004). Review of international experience in VOT study design. *Proceedings of the EuropeanTransport Conference.*

Burgess, L. S. (2005). Optimal designs for choice experiments with asymmetric attributes. *Journal of Statistical Planning and Inference*(134), pp. 288-301.

Car2Go. (2015). *Unternehmensprofil*. Retrieved 06 24, 2015, from Car2Go: https://www.car2go.com/de/stuttgart/unternehmen/

Carrier, E. (2003). *Modeling airline passenger choice: Passenger preference for schedule in the passenger origin–destination simulator (PODS). .* Massachusetts, MA: Ph.D. thesis. MIT.

ChoiceMetric Pty Ltd. (2014). *Ngene 1.1.2 User Manual & Reference Guide - The Cutting Edge in Experimental Design.*

Cragg, J. G. (1970). The demand for automobiles. *Canadian Journal of Economics*(3), pp. 386-406.

De Donnea, F. (1972). Consumer behaviour, transport mode choice and value of time: some micro-economic models. *Regional and Urban Economics*(1), pp. 355-382.

Devarasetty, P. B. (2012). The value of travel time and reliability-evidence from a stated preference survey and actual usage. *Transportation Research, Part A*(46), pp. 1227-1240.

Dhar, R. S. (2001). The Effect of Forced Choice on Choice. *Working Paper, Yale University*.

Efron, B. (1978). Regression and ANOVA with zero-one data: Measures of residual variation. *Journal of the American Statistical Association*(73), pp. 113-121.

European Environment Agency. (2012, 01 12). *Share of household expenditure on transport services across countries (% of total spending, 2008)*. Retrieved 06 24, 2015, from European Environment Agency: http://www.eea.europa.eu/data-and-maps/figures/term24-share-of-household-expenditure-on-transport-services-across-countries-of-total-spending

Ferrini, S. S. (2007). Designs with a priori information for nonmarket valuation with choice-experiments: a Monte Carlo study. *Journal of Environmental Economics and Management*(53), pp. 342-363.

finanzen.net. (2015, 06 17). *Exchange Rates*. Retrieved from http://www.finanzen.net/: http://www.finanzen.net/waehrungsrechner/euro_us-dollar

Flynn, T. L. (2008). Estimating preferences for a dermatology consultation using best-worst scaling: comparison of various methods of analysis. *BMC Medical Research Methodology*(76), pp. 1-12.

Fosgerau, M. B. (2007). A practical test for the choice of mixing distribution in discrete choice models. *Transportation Research*(41), pp. 784-794.

Fowkes, A. W. (1988). The design of stated preference travel choice experiments with particular regard to inter-personal taste variations. *Journal of Transport Economics and Policy*(22), pp. 27-44.

Fowkes, A. W. (1993). Non-orthogonal stated preference design. *Proceedings of the PTRC Summer Annual Meeting* (pp. 91-97). London: PTRC Education Research Services Limited.

Frejinger, E. B. (2007). Capturing correlation with subnetworks in route choice models. *Transportation Research*(41), pp. 363–378.

Google Maps. (2015, 03 12). *Maps*. Retrieved 03 12, 2015, from Google: https://www.google.de/maps

Greene, W. H. (2006). Accounting for heterogeneity in the variance of unobserved effects in mixed logit models. *Transportation Research*(40), pp. 75–92.

Group, A. M. (1994). *The value of travel time on UK roads.* London: Report prepared for UK Department of Transport.

Gunn, H. (2007). An introduction to the valuation of travel-time savings and losses. In D. B. Hensher, *Handbook of Transport Modelling (2nd Edition)* (pp. 433-448). Amsterdam: Elsevier Science Ltd.

Hagle, T. M. (1992). Goodness-of-fit measures for probit and logit. *American Journal of Political Science*(36), pp. 762-784.

Han, B. A. (2001). Accommodating drivers' taste variation and repeated choice correlation in route choice modeling by using the mixed logit model. Eightieth Annual Meeting of the Transportation Research Board.

Hensher, D. (2001a). The sensitivity of the valuation of travel time savings to the specification of unobserved effects. *Transportation Research, Part E - Special Issue on Travel-Time Savings*(37), pp. 129-142.

Hensher, D. (2001b). The valuation of commuter travel time savings for car drivers: evaluating alternative model specifications. *Transportation*(28), pp. 101-118.

Hensher, D. (2011). Valuation of travel time savings. In A. L. De Palma, *A Handbook of Transport Economics.* (pp. 135-160). Cheltenham: Edward Elgar Publishing.

Hensher, D. G. (2003). The mixed logit model: The state of practice. *Transportation*(30), pp. 133–176.

Hess, S. B. (2005). Estimation of value of travel-time savings using mixed logit models. *Transportation Research*(39), pp. 221–236.

Hess, S. P. (2005). Mixed logit modelling of airport choice in multi-airport regions. *Journal of Air Transport Management*(11), pp. 59–68.

Hess, S. T. (2006). On the use of modified latin hypercube sampling (MLHS) method in the estimation of mixed logit model for vehicle choice. *Transportation Research*(40), pp. 147–163.

Hoyos, D. M.-M. (2009). The influence of cultural identity on the WTP to protect natural resources: Some empirical evidence. *Ecological Economics*(68), pp. 2372-2381.

Huber, J. Z. (1996). The importance of utility balance and efficient choice designs. *Journal of Marketing Research*(33), pp. 307-317.

Kahneman D, T. A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*(47), pp. 263-91.

Kamga, C. Y. (2014). Temporal and weather related variation patterns of urban travel time: Considerations and caveats for value of travel time, value of variability, and mode choice studies. *Transportation Research, Part C*(45), pp. 4-16.

Kanninen, B. (2002). Optimal design for multinomial choice experiments. *Journal of Marketing Research*(39), pp. 214-217.

Levine, R. N. (1999). The Pace of Life in 31 Countries. *Journal of Cross-Cultural Psychology*(30), 178–205.

Long, S. (1997). *Regression Models for Categorical and Limited Dependent Variables.* Thousand Oaks (CA): Sage Publications Inc.

Louviere, J. (1992). Experimental choice analysis: Introduction and overview. *Journal of Business Research*(24), pp. 89–95.

Louviere, J. H. (1983). Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event. *Journal of Consumer Research*(10), pp. 348-361.

Louviere, J. H. (2000). *Stated Choice Methods: Analysis and Applications.* Cambridge: Cambridge University Press.

Louviere, J. I. (2008). Designing discrete choice experiments: do optimal designs come at a price? *Journal of Consumer Research*(35), pp. 360-375.

Louviere, J. P. (2010). Design of Discrete Choice Experiments: A Discussion of Issues That Matter in Future Applied Research. *Journal of Choice Modelling*(4), pp. 1-8.

Louviere, J. W. (1983). Design and analysis of simulated consumer choice or allocation experiments: an approach based on aggregated data. *Journal of Marketing Research*(20), pp. 350-367.

Maddala, G. (1983). *Limited-dependent and qualitative variables in econometrics.* Cambridge: Cambridge University Press.

Maddala, G. S. (1992). *Introduction to econometrics (2nd. edition).* New York: Macmillan.

Main Regional Public Transportation Provider (2014). *Yearly Activity Report.*

Manrai, A. (1995). Mathematical models of brand choice behavior. *European Journal of Operational Research*(82), pp. 1-17.

Manski, C. (1977). The structure of random utility models. *Theory and decision, 8*(3), pp. 229-254.

Mariel, P. d. (2013). Selecting random parameters in discrete choice experiment for environmental valuation: A simulation experiment. *The Journal of Choice Modelling*(7), pp. 33-57.

Marschak, J. (1960). Binary choice constraints on random utility indications. In K. Arrow, *Stanford Symposium on Mathematical Methods in the Social Sciences* (pp. 312–329). Stanford, CA: Stanford University Press.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. *Zarembka, P. (Ed.), Frontiers in Econometrics, Academic Press*.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behaviour. *Zarembka, P. (Ed.), Frontiers in Econometrics, Academic Press*.

McFadden, D. (1981). Econometric models of probabilistic choice. In C. M. Manski, *Structural Analysis of Discrete Data with Econometric Applications.* (pp. 198-272). Cambridge, MA: MIT Press.

McFadden, D. T. (2000). Mixed MNL models for discrete response. *Journal of Applied Econometrics*, pp. 447–470.

McKelvey, R. D. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*(4), pp. 103-120.

Mohler, B. T.-R. (2007). Visual flow influences gait transition speed and preferred walking speed. *Experimental Brain Research*(181), 221–228.

Raftery, A. (1996). Bayesian model selection in social research. In P. Marsden, *Sociological methodology (Vol. 25)* (pp. 111-163). Oxford: Basil Blackwell.

Regional Institute for Statistics & Analysis. (2013). *City Region in figures.* Regional Public Service.

Revelt, D. T. (1998). Mixed logit with repeated choices: Households' choices of applicance efficiency level. *Review of Economics*.

Rose, J. B. (2005). Sample optimality in the design of stated choice experiments. *working paper, Institute of Transportation Studies, University of Sydney*.

Rose, J. B. (2006). Designing efficient data for stated choice experiments. *11th International Conference on Travel Behaviour Research.* Kyoto.

Rose, J. B. (2014). Stated Choice experimental design theory: the who, the what and the why. In S. D. Hess, *Handbook of Choice Modelling.* Cheltenham, UK: Edward Elgar Publishing.

Rose, J. S. (2009). Incorporating model uncertainty into the generation of efficient stated choice experiments: a model averaging approach. *International Choice Modelling Conference.* Harrowgate.

Sándor, Z. W. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*(38), pp. 430-444.

Sándor, Z. W. (2002). Profile construction in experimental choice designs for mixed logit models. *Marketing Science*(21), pp. 455-475.

Sándor, Z. W. (2005). Heterogenous conjoint choice designs. *Journal of Marketing Research*(42), pp. 210-218.

Shires, J. d. (2009). An international meta-analysis of values of travel time savings. *Evaluation and Program Planning*(32), pp. 315-325.

Small, K. W. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*(73), pp. 1367–1382.

Thurstone, L. (1927). A law of comparative judgement. *Psychological Review*(34), pp. 273-286.

Toner, J. C.-M. (1998). Anything you can do, we can do better: a provocative introduction to a new approach to stated preference design. (pp. 107-120). Antwerp: WCTR Proceedings .

Toner, J. W. (1999). Testing recent advances in stated preference design. *Proceedings of the European Transport Conference.* Cambridge.

Train, K. (1998). Recreation demand models with taste differences over people. *Land Economics*(74), pp. 230–239.

Train, K. (1999). Halton Sequences for Mixed Logit. *Econometrics*.

Train, K. (2009). *Discrete Choice Methods with Simulation (2nd. Edition).* Cambridge: Cambridge University Press.

Tversky, A. S. (1992). Choice Under Conflict: The Dynamics of Deferred Decision. *Psychological Science*(6), pp. 358-361.

UK, O. f. (2015). *Retail Price Index*. Retrieved 06 17, 2015, from http://ons.gov.uk/: http://ons.gov.uk/ons/taxonomy/index.html?nscl=Retail+Prices+Index

USA, B. o. (2015). *Consumer Price Index*. Retrieved 06 17, 2015, from http://www.bls.gov/: http://www.bls.gov/cpi/

Viton, P. (2004). Will mixed logit change urban transport policies? *Journal of Transport Economics and Policy*(38), pp. 403–423.

Wardman, M. (2004). Public transport values of time. *Transport Policy*(11), pp. 363–377.

Wardman, M. A. (2011). Meta-analysis of UK values of travel time: An update. *Transportation Research, Part A*(45), pp. 1–17.

Watson, S. T. (2000). Efficiency properties of orthogonal stated preference designs. In J. d. Ortuzar, *Stated Preference Modelling Techniques* (pp. 91-101). London: PTRC Education and Research Services Limited.

Wikipedia. (2015a, 05 16). *Carsharing*. Retrieved 06 24, 2015, from Wikipedia: https://de.wikipedia.org/wiki/Carsharing

Wikipedia. (2015b, 06 17). *Car2Go*. Retrieved 06 24, 2015, from Wikipedia: https://de.wikipedia.org/wiki/Car2go

Wikipedia. (2015c, 06 24). *Carsharing (English)*. Retrieved 06 24, 2015, from Wikipedia: https://en.wikipedia.org/wiki/Carsharing

Windmeijer, F. (1995). Goodness-of-fit measures in binary choice models. *Econometric Reviews*(55), pp. 876-890.

V. Appendix

Appendix 1: NGENE 1.1.2 Design Code

```
Design
;alts = Bus, Metro, Taxi, C2G, None

? MNL model optimizing d-error
;rows   = 60
;block  = 20
? 700 individuals * 3 choices = 2100 choicesets --> 35 Obs per row
;eff    = (MNL, d)

? Conditions:
;cond:
? Traveltime adder fitting taveltime
  if(Bus.btraveltime   =  5, Bus.addedtraveltime   = 10),
  if(Bus.btraveltime   = 15, Bus.addedtraveltime   = 30),
  if(Bus.btraveltime   = 30, Bus.addedtraveltime   = 30),
  if(Taxi.ctraveltime  =  5, Taxi.addedtraveltime  = 10),
  if(Taxi.ctraveltime  = 15, Taxi.addedtraveltime  = 30),
  if(Taxi.ctraveltime  = 30, Taxi.addedtraveltime  = 30),
  if(C2G.ctraveltime   =  5, C2G.addedtraveltime   = 10),
  if(C2G.ctraveltime   = 15, C2G.addedtraveltime   = 30),
  if(C2G.ctraveltime   = 30, C2G.addedtraveltime   = 30),

? Pricemarkup fitting price
  if(Taxi.tprice  =  6.40, Taxi.mtpricemarkup = 10.00),
  if(Taxi.tprice  = 14.40, Taxi.mtpricemarkup = 15.00),
  if(Taxi.tprice  = 26.40, Taxi.mtpricemarkup = 15.00),
  if(C2G.cprice   =  1.00, C2G.mcpricemarkup  =  2.00),
  if(C2G.cprice   =  1.25, C2G.mcpricemarkup  =  1.50),
  if(C2G.cprice   =  1.50, C2G.mcpricemarkup  =  3.00),
  if(C2G.cprice   =  1.75, C2G.mcpricemarkup  =  3.50),
  if(C2G.cprice   =  2.00, C2G.mcpricemarkup  =  4.00),
  if(C2G.cprice   =  3.00, C2G.mcpricemarkup  =  6.00),
  if(C2G.cprice   =  3.75, C2G.mcpricemarkup  =  7.50),
  if(C2G.cprice   =  4.50, C2G.mcpricemarkup  =  9.00),
  if(C2G.cprice   =  5.25, C2G.mcpricemarkup  = 10.50),
  if(C2G.cprice   =  6.00 and C2G.ctraveltime = 30, C2G.mcpricemarkup  = 12.00),
  if(C2G.cprice   =  6.00 and C2G.ctraveltime = 45, C2G.mcpricemarkup  =  4.25),
  if(C2G.cprice   =  7.50, C2G.mcpricemarkup  =  3.75),
  if(C2G.cprice   =  9.00, C2G.mcpricemarkup  =  6.00),
  if(C2G.cprice   = 10.50, C2G.mcpricemarkup  =  7.00),
  if(C2G.cprice   = 12.00, C2G.mcpricemarkup  =  8.00),

? Canceling out dominant szenarios
  if(C2G.cdistance   =  500, Bus.bdistance  >= 150 and Metro.mdistance >= 150)   ,

? Night vs Morning
  if(Bus.morning = 1, Bus.night = 0),

? Traveltime & Price:
?taxi
  if(Taxi.ctraveltime =  5, Taxi.tprice = [ 6.4])                                ,
  if(Taxi.ctraveltime = 15, Taxi.tprice = [14.4])                                ,
  if(Taxi.ctraveltime = 30, Taxi.tprice = [26.4])                                ,
?C2G
  if(C2G.ctraveltime =  5, C2G.cprice = [ 1.0,  1.25,  1.50,  1.75,  2.0])       ,
  if(C2G.ctraveltime = 15, C2G.cprice = [ 3.0,  3.75,  4.50,  5.25,  6.0])       ,
  if(C2G.ctraveltime = 30, C2G.cprice = [ 6.0,  7.50,  9.00, 10.50, 12.0])       ,

? Egresstime & Price
 if(C2G.cegresstime = 0, C2G.egresspricemarkup = 0)                             ,
 if(C2G.cprice = [1.00, 3.00,  6.00] and C2G.cegresstime = 3.5, C2G.egresspricemarkup = 0.70),
 if(C2G.cprice = [1.25, 3.75,  7.50] and C2G.cegresstime = 3.5, C2G.egresspricemarkup = 0.90),
 if(C2G.cprice = [1.50, 4.50,  9.00] and C2G.cegresstime = 3.5, C2G.egresspricemarkup = 1.05),
 if(C2G.cprice = [1.75, 5.25, 10.50] and C2G.cegresstime = 3.5, C2G.egresspricemarkup = 1.20),
 if(C2G.cprice = [2.00, 6.00, 12.00] and C2G.cegresstime = 3.5, C2G.egresspricemarkup = 1.40),

? Traveltime over alternatives
  if(C2G.ctraveltime = 15 and C2G.morning = 0, metro.mtraveltime   >= 15),
  if(C2G.ctraveltime =  5      , taxi.ctraveltime   =  5 and metro.mtraveltime   <= 45),
  if(C2G.ctraveltime = 15      , bus.btraveltime   >= 15 and metro.mtraveltime   <= 45 and
       taxi.ctraveltime = 15),
  if(C2G.ctraveltime = 30      , bus.btraveltime   >= 30 and metro.mtraveltime   >= 15 and
       taxi.ctraveltime = 30),
  if(C2G.ctraveltime = 45      , bus.btraveltime   >= 45 and metro.mtraveltime   >= 15 and
       taxi.ctraveltime = 45) ,
  if(C2G.ctraveltime = 60      , bus.btraveltime    = 60 and metro.mtraveltime   >= 15 and
```

```
        taxi.ctraveltime = 60)
? Assumption: No additional time for Parking (Dif goes to base Utility Taxi)

? Model:
? Szenarios
? 1.morning:  Midday [0], Morning[1]
? 2 night:    Midday [0], Night  [1]
? 3. rain:    Sun    [0], Rain   [1]
? 4. city:    Home   [0], City   [1]

;model:
U(Bus)   = b11[1.45]
            + start[.5]       *city[0,1]
            + time11[-.20]    *night[0, 0.00000001 ,1]
            + traffic11[-0.5] *morning[0, 0.00000001 ,1]
            + weather11[0.5]  *rain[0,1]
            + b2[-.001]       *bdistance[0, 150, 300, 500, 1000]
            + cd[.0005]       *cbdistance[bdistance]*city[city]
            + td[-.001]       *nbdistance[bdistance]*night[night]
            + md[-.0005]      *mbdistance[bdistance]*morning[morning]
            + rd[-.002]       *rbdistance[bdistance]*rain[rain]
            + b3[-.1]         *btraveltime[5, 15, 30        ]
            + c3[.01]         *cbtraveltime[btraveltime]*city[city]
            + n3[-.05]        *nbtraveltime[btraveltime]*night[night]
            + m3[-.05]        *mbtraveltime[btraveltime]*morning[morning]
            + b3              *addedtraveltime[10, 30]*morning[morning]
            + r3[-.005]       *rbtraveltime[btraveltime]*rain[rain]
            + b4[-.1]         *begresstime[0, 3.5, 6]
            + c4[.01]         *cbegresstime[begresstime]*city[city]
            + n4[-.1]         *nbegresstime[begresstime]*night[night]
            + m4[-.05]        *mbegresstime[begresstime]*morning[morning]
            + r4[-.2]         *rbegresstime[begresstime]*rain[rain]
            + b5[-.2]         *bprice[0, 2.0, 2.000001]
            + c5[.01]         *cbprice[bprice]*city[city]
            + n5[.05]         *nbprice[bprice]*night[night]
            + m5[-.01]        *mbprice[bprice]*morning[morning]
            + r5[.01]         *rbprice[bprice]*rain[rain]  /
U(Metro) = b12[1.45]
            + start           *city[city]
            + time12[-.12]    *night[night]
            + traffic12[-0.2] *morning[morning]
            + weather12[0.8]  *rain[rain]
            + b2              *mdistance[0, 150, 300, 500, 1000]
            + cd              *cmdistance[mdistance]*city[city]
            + td              *nbdistance[mdistance]*night[night]
            + md              *mmdistance[mdistance]*morning[morning]
            + rd              *rmdistance[mdistance]*rain[rain]
            + b3              *mtraveltime[5, 15, 30, 45, 60]
            + c3              *cmtraveltime[mtraveltime]*city[city]
            + n3              *nmtraveltime[mtraveltime]*night[night]
            + m3              *mmtraveltime[mtraveltime]*morning[morning]
            + r3              *rmtraveltime[mtraveltime]*rain[rain]
            + b4              *megresstime[0, 3.5, 6]
            + c4              *cmegresstime[megresstime]*city[city]
            + n4              *nmegresstime[megresstime]*night[night]
            + m4              *mmegresstime[megresstime]*morning[morning]
            + r4              *rmegresstime[megresstime]*rain[rain]
            + b5              *mprice[bprice]
            + c5              *cmprice[bprice]*city[city]
            + n5              *nmprice[bprice]*night[night]
            + m5              *mmprice[bprice]*morning[morning]
            + r5              *rmprice[bprice]*rain[rain]  /
U(Taxi)  = b13[1.40]
            + start           *city[city]
            + time13[ .30]    *night[night]
            + traffic13[-0.5] *morning[morning]
            + weather13[1.0]  *rain[rain]
            + b3              *ctraveltime[5, 15, 30        ]
            + c3              *cctraveltime[ctraveltime]*city[city]
            + n3              *nctraveltime[ctraveltime]*night[night]
            + m3              *mctraveltime[ctraveltime]*morning[morning]
            + b3              *addedtraveltime*morning[morning]
            + r3              *rctraveltime[ctraveltime]*rain[rain]
            + b5              *tprice[6.4, 14.4, 26.4]
            + c5              *ctprice[tprice]*city[city]
            + b5              *ctpricemarkup[-2.40]*city[city]
            + n5              *ntprice[tprice]*night[night]
            + m5              *mtprice[tprice]*morning[morning]
```

```
            + b5              *mtpricemarkup[10, 15]*morning[morning]
            + r5              *rtprice[tprice]*rain[rain]  /
U(C2G)    = b14[1.30]
            + start           *city[city]
            + time14[-.30]    *night[night]
            + traffic14[-1.0] *morning[morning]
            + weather14[0.8]  *rain[rain]
            + b2              *cdistance[0, 150, 300, 500]
            + cd              *ccdistance[bdistance]*city[city]
            + td              *ncdistance[cdistance]*night[night]
            + md              *mcdistance[cdistance]*morning[morning]
            + rd              *rcdistance[cdistance]*rain[rain]
            + b3              *ctraveltime[ctraveltime      ]
            + c3              *cctraveltime[ctraveltime]*city[city]
            + n3              *nctraveltime[ctraveltime]*night[night]
            + m3              *mctraveltime[ctraveltime]*morning[morning]
            + b3              *addedtraveltime*morning[morning]
            + r3              *rctraveltime[ctraveltime]*rain[rain]
            + b4              *cegresstime[0, 3.5]
            + c4              *ccegresstime[cegresstime]*city[city]
            + n4              *ncegresstime[cegresstime]*night[night]
            + m4              *mcegresstime[cegresstime]*morning[morning]
            + r4              *rcegresstime[cegresstime]*rain[rain]
            + b5              *cprice[1.0, 1.25, 1.50, 1.75, 2.0, 3.0, 3.75, 4.50, 5.25, 6.0,
                                      7.50, 9.00, 10.50, 12.0]
            + c5              *ccprice[cprice]*city[city]
            + n5              *ncprice[cprice]*night[night]
            + m5              *mcprice[cprice]*morning[morning]
            + b5              *mcpricemarkup[2.00, 1.50, 3.00, 3.50, 4.00,  7.50,  9.00,
                                             10.50, 12.00,  4.25,  3.75,  6.00,  7.00,
                                             8.00]*morning[morning]
            + r5              *rcprice[cprice]*rain[rain]
            + b5              *egresspricemarkup[0, 0.7, 0.9, 1.05, 1.20, 1.4]

$
```

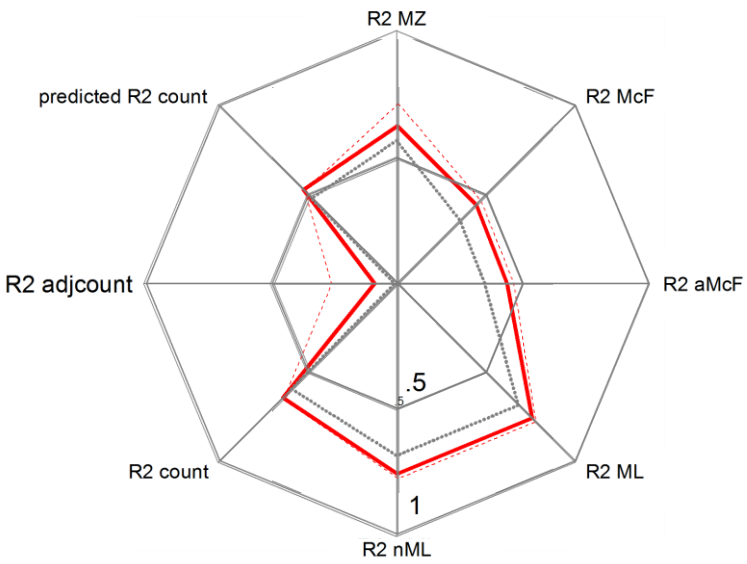
Appendix 2: Stata Code Simulation (CLM as MXL is very similar)

Please see separate file simulation.do

Appendix 3: Results Simulation Experiment

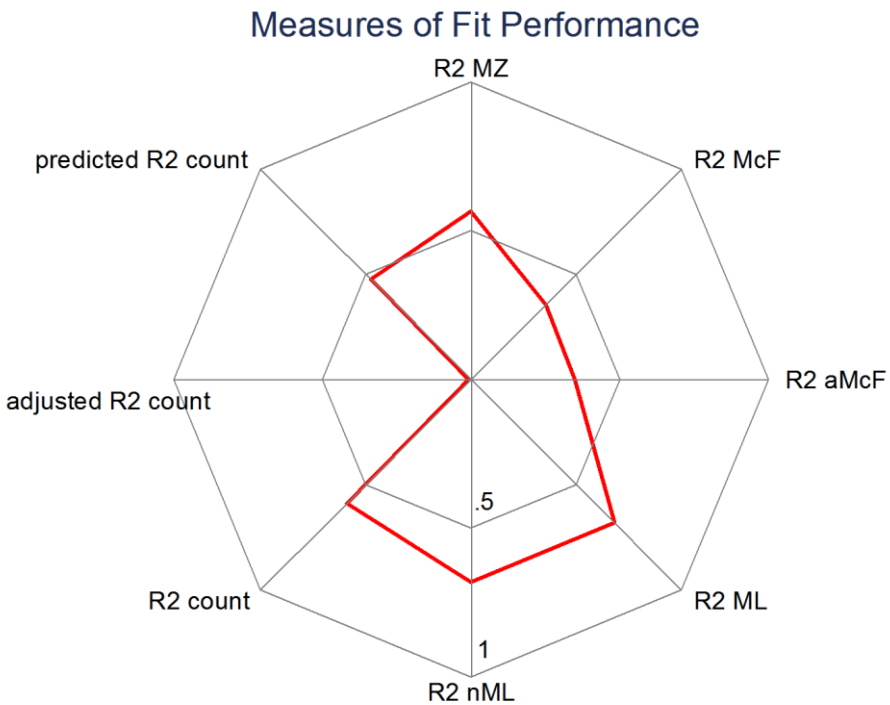
Results CLM

			Coef	Std.Def.	Real Coef
Main Effects	access	(in 100m)	-0.1303	(0.0254)	-0.1288
	in-mode	(in Min.)	-0.0913	(0.0087)	-0.0908
	egress	(in Min.)	-0.0477	(0.3263)	-0.0485
	price	(in EUR)	-0.1164	(0.0286)	-0.1149
CITY	access	(in 100m)	0.0378	(0.0371)	0.0372
	in-mode	(in Min.)	0.0084	(0.0125)	-0.0485
	egress	(in Min.)	0.0039	(0.0491)	0.0068
	price	(in EUR)	0.0125	(0.0318)	0.0111
MORNING	access	(in 100m)	-0.0378	(0.0368)	-0.0345
	in-mode	(in Min.)	-0.0814	(0.0150)	0.0807
	egress	(in Min.)	-0.0771	(0.0511)	-0.0761
	price	(in EUR)	-0.0319	(0.0189)	-0.0325
NIGHT	access	(in 100m)	-0.1088	(0.0582)	-0.1033
	in-mode	(in Min.)	-0.0509	(0.0173)	-0.0509
	egress	(in Min.)	-0.0916	(0.0448)	-0.0761
	price	(in EUR)	-0.0058	(0.0391)	-0.0062
RAIN	access	(in 100m)	-0.3124	(0.0433)	-0.3123
	in-mode	(in Min.)	-0.0046	(0.0117)	-0.0043
	egress	(in Min.)	-0.0641	(0.0394)	-0.0907
	price	(in EUR)	0.0194	(0.0324)	0.0199



Results MXL

	Estimate	Std. Error	t-value	Pr(>|t|)	Prior	Delta
Bus	2.1337	0.4831	4.4200	0.0000 ***	1.4200	0.7137
Standard Deviation	0.9320	0.7044	1.3200	0.1860		
Metro	2.0139	0.3853	5.2300	0.0000 ***	1.4600	0.5539
Standard Deviation	-0.0148	0.7615	-0.0200	0.9840		
Taxi	1.8538	0.4260	4.3500	0.0000 ***	1.5300	0.3238
Standard Deviation	-0.2944	0.9309	-0.3200	0.7520		
C2G	1.2858	0.4908	2.6200	0.0090 ***	1.3000	-0.0142
Standard Deviation	-1.0343	0.7554	-1.3700	0.1710		
Distance	-6.4758	0.1820	-35.5900	0.0000 ***	-6.9078	0.4320
Standard Deviation	1.1877	0.3609	3.2900	0.0010 ***		
In-Mode Time	-2.0037	0.1541	-13.0000	0.0000 ***	-2.3026	0.2989
Standard Deviation	0.3546	0.1553	2.2800	0.0220 **		
Egress Time	-2.2479	0.3806	-5.9100	0.0000 ***	-2.3026	0.0547
Standard Deviation	1.0274	0.7372	1.3900	0.1630		
Price	-2.3661	0.5168	-4.5800	0.0000 ***	-1.6094	-0.7567
Standard Deviation	1.4109	0.5120	2.7600	0.0060 ***		



Appendix 4: Overview Choice Tasks

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
1	Bus	Home	Normal	22:00	Rain	40	800		2.00 €
	Metro / Tram					62	150		2.00 €
	Taxi					30	0		24.00 €
	C2G					37	300	3.5	13.40 €
	None					0	0		- €
1	Bus	Home	Normal	13:00	Sun	36	500		2.00 €
	Metro / Tram					24	800		2.00 €
	Taxi					30	0		24.00 €
	C2G					37	300	3.5	13.40 €
	None					0	0		- €
1	Bus	City	Normal	13:00	Rain	27	1000		2.00 €
	Metro / Tram					10	450		2.00 €
	Taxi					15	0		18.40 €
	C2G					15	0	0	5.25 €
	None					0	0		- €
2	Bus	City	Normal	22:00	No Rain	14	800		- €
	Metro / Tram					5	0		- €
	Taxi					5	0		10.40 €
	C2G					9	300	0	2.00 €
	None					0	0		- €
2	Bus	City	Normal	13:00	Rain	34	300		- €
	Metro / Tram					19	300		- €
	Taxi					15	0		18.40 €
	C2G					20	150	3.5	6.45 €
	None					0	0		- €
2	Bus	Home	Normal	13:00	Rain	30	0		- €
	Metro / Tram					35	450		- €
	Taxi					30	0		24.00 €
	C2G					32	150	0	10.50 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
3	Bus	City	Normal	22:00	Rain	32	150		2.00 €
	Metro / Tram					55	800		2.00 €
	Taxi					30	0		26.40 €
	C2G					37	300	3.5	11.70 €
	None					0	0		- €
3	Bus	City	Rushhour	08:00	Sun	23	650		2.00 €
	Metro / Tram					20	1300		2.00 €
	Taxi					15	0		20.40 €
	C2G					22	300	3.5	7.40 €
	None					0	0		- €
3	Bus	Home	Normal	22:00	Rain	39	800		2.00 €
	Metro / Tram					51	500		2.00 €
	Taxi					5	0		8.00 €
	C2G					7	150	0	1.50 €
	None					0	0		- €
4	Bus	City	Normal	22:00	Rain	15	0		2.00 €
	Metro / Tram					60	1300		2.00 €
	Taxi					5	0		10.40 €
	C2G					10	150	3.5	1.70 €
	None					0	0		- €
4	Bus	City	Normal	13:00	Sun	39	800		2.00 €
	Metro / Tram					78	1500		2.00 €
	Taxi					30	0		26.40 €
	C2G					37	300	3.5	8.40 €
	None					0	0		- €
4	Bus	City	Normal	13:00	Rain	32	150		2.00 €
	Metro / Tram					57	1000		2.00 €
	Taxi					15	0		18.40 €
	C2G					24	500	3.5	3.70 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
	Bus					36	500		- €
	Metro / Tram					52	600		- €
5	Taxi	Home	Normal	22:00	Rain	30	0		24.00 €
	C2G					34	0	3.5	10.05 €
	None					0	0		- €
	Bus					40	800		2.00 €
	Metro / Tram					42	1000		2.00 €
5	Taxi	City	Normal	13:00	Sun	30	0		26.40 €
	C2G					36	500	0	7.50 €
	None					0	0		- €
	Bus					23	650		- €
	Metro / Tram					45	1300		- €
5	Taxi	City	Normal	13:00	Rain	5	0		10.40 €
	C2G					9	0	3.5	3.40 €
	None					0	0		- €
	Bus					63	1500		2.00 €
	Metro / Tram					21	500		2.00 €
6	Taxi	City	Rushhour	08:00	Sun	15	0		20.40 €
	C2G					15	0	0	2.75 €
	None					0	0		- €
	Bus					55	800		- €
	Metro / Tram					21	500		- €
6	Taxi	City	Rushhour	08:00	Rain	45	0		33.40 €
	C2G					49	0	3.5	12.15 €
	None					0	0		- €
	Bus					72	1000		2.00 €
	Metro / Tram					55	800		2.00 €
6	Taxi	City	Rushhour	08:00	Sun	15	0		20.40 €
	C2G					24	500	3.5	3.65 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
	Bus					45	1300		2.00 €
	Metro / Tram					57	1000		2.00 €
7	Taxi	Home	Normal	13:00	Sun	5	0		8.00 €
	C2G					11	500	0	1.50 €
	None					0	0		- €
	Bus					37	600		- €
	Metro / Tram					49	300		- €
7	Taxi	City	Normal	13:00	Sun	15	0		18.40 €
	C2G					21	500	0	3.75 €
	None					0	0		- €
	Bus					27	1000		2.00 €
	Metro / Tram					32	150		2.00 €
7	Taxi	City	Normal	13:00	Sun	15	0		18.40 €
	C2G					17	150	0	6.00 €
	None					0	0		- €
	Bus					39	800		2.00 €
	Metro / Tram					54	800		2.00 €
8	Taxi	City	Normal	13:00	Rain	30	0		26.40 €
	C2G					36	500	0	10.50 €
	None					0	0		- €
	Bus					38	650		2.00 €
	Metro / Tram					34	300		2.00 €
8	Taxi	City	Normal	13:00	Rain	15	0		18.40 €
	C2G					22	300	3.5	6.45 €
	None					0	0		- €
	Bus					27	1000		- €
	Metro / Tram					49	300		- €
8	Taxi	City	Normal	13:00	Rain	15	0		18.40 €
	C2G					19	300	0	3.00 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
	Bus					60	1300		2.00 €
	Metro / Tram					27	1000		2.00 €
9	Taxi	Home	Rushhour	08:00	Sun	15	0		18.00 €
	C2G					20	150	3.5	3.65 €
	None					0	0		- €
	Bus					20	1300		2.00 €
	Metro / Tram					53	650		2.00 €
9	Taxi	Home	Normal	13:00	Sun	5	0		8.00 €
	C2G					10	150	3.5	1.70 €
	None					0	0		- €
	Bus					21	500		2.00 €
	Metro / Tram					32	150		2.00 €
9	Taxi	City	Normal	22:00	No Rain	15	0		18.40 €
	C2G					19	300	0	4.50 €
	None					0	0		- €
	Bus					30	1300		2.00 €
	Metro / Tram					47	150		2.00 €
10	Taxi	City	Normal	13:00	Sun	15	0		18.40 €
	C2G					19	300	0	4.50 €
	None					0	0		- €
	Bus					30	0		2.00 €
	Metro / Tram					15	0		2.00 €
10	Taxi	City	Normal	22:00	Rain	30	0		26.40 €
	C2G					34	300	0	12.00 €
	None					0	0		- €
	Bus					38	650		2.00 €
	Metro / Tram					60	1300		2.00 €
10	Taxi	Home	Normal	13:00	Sun	30	0		24.00 €
	C2G					39	500	3.5	8.40 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
	Bus					24	800		2.00 €
	Metro / Tram					9	300		2.00 €
11	Taxi	Home	Normal	22:00	Rain	15	0		16.00 €
	C2G					15	0	0	4.50 €
	None					0	0		- €
	Bus					11	500		- €
	Metro / Tram					23	650		- €
11	Taxi	City	Normal	22:00	No Rain	5	0		10.40 €
	C2G					9	300	0	1.75 €
	None					0	0		- €
	Bus					19	300		- €
	Metro / Tram					9	300		- €
11	Taxi	City	Rushhour	08:00	Sun	15	0		20.40 €
	C2G					19	300	0	3.00 €
	None					0	0		- €
	Bus					11	500		- €
	Metro / Tram					15	0		- €
12	Taxi	City	Normal	13:00	Rain	5	0		10.40 €
	C2G					9	0	3.5	2.15 €
	None					0	0		- €
	Bus					34	300		- €
	Metro / Tram					23	650		- €
12	Taxi	City	Normal	13:00	Sun	15	0		18.40 €
	C2G					21	500	0	6.00 €
	None					0	0		- €
	Bus					25	800		2.00 €
	Metro / Tram					39	800		2.00 €
12	Taxi	City	Normal	22:00	Rain	15	0		18.40 €
	C2G					15	0	0	6.00 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
	Bus					20	1300		- €
	Metro / Tram					15	800		- €
13	Taxi	City	Normal	13:00	Sun	5	0		10.40 €
	C2G					12	300	3.5	2.95 €
	None					0	0		- €
	Bus					30	1300		2.00 €
	Metro / Tram					27	1000		2.00 €
13	Taxi	City	Normal	13:00	Sun	15	0		18.40 €
	C2G					24	500	3.5	5.55 €
	None					0	0		- €
	Bus					22	600		- €
	Metro / Tram					57	1000		- €
13	Taxi	City	Normal	13:00	Sun	5	0		10.40 €
	C2G					12	300	3.5	2.55 €
	None					0	0		- €
	Bus					36	500		2.00 €
	Metro / Tram					27	1000		2.00 €
14	Taxi	City	Normal	13:00	Sun	15	0		18.40 €
	C2G					15	0	0	3.75 €
	None					0	0		- €
	Bus					34	300		2.00 €
	Metro / Tram					72	1000		2.00 €
14	Taxi	City	Normal	13:00	Rain	30	0		26.40 €
	C2G					36	500	0	9.00 €
	None					0	0		- €
	Bus					15	0		- €
	Metro / Tram					40	800		- €
14	Taxi	Home	Normal	22:00	No Rain	15	0		16.00 €
	C2G					20	150	3.5	3.70 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
	Bus					42	1000		- €
	Metro / Tram					15	0		- €
15	Taxi	City	Normal	13:00	Rain	30	0		26.40 €
	C2G					32	150	0	7.50 €
	None					0	0		- €
	Bus					17	1000		- €
	Metro / Tram					17	150		- €
15	Taxi	City	Normal	13:00	Sun	5	0		10.40 €
	C2G					11	500	0	1.00 €
	None					0	0		- €
	Bus					20	450		2.00 €
	Metro / Tram					53	650		2.00 €
15	Taxi	City	Normal	13:00	Rain	15	0		18.40 €
	C2G					19	0	3.5	6.45 €
	None					0	0		- €
	Bus					32	150		2.00 €
	Metro / Tram					22	600		2.00 €
16	Taxi	City	Normal	13:00	Rain	30	0		26.40 €
	C2G					34	0	3.5	10.05 €
	None					0	0		- €
	Bus					49	300		2.00 €
	Metro / Tram					49	300		2.00 €
16	Taxi	Home	Rushhour	08:00	Rain	15	0		18.00 €
	C2G					20	150	3.5	5.55 €
	None					0	0		- €
	Bus					62	150		- €
	Metro / Tram					15	0		- €
16	Taxi	City	Rushhour	08:00	Rain	15	0		20.40 €
	C2G					15	0	0	6.00 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
	Bus					39	800		2.00 €
	Metro / Tram					57	1000		2.00 €
17	Taxi	City	Normal	13:00	Rain	15	0		18.40 €
	C2G					19	300	0	6.00 €
	None					0	0		- €
	Bus					34	300		2.00 €
	Metro / Tram					42	1000		2.00 €
17	Taxi	City	Normal	13:00	Rain	30	0		26.40 €
	C2G					35	150	3.5	10.05 €
	None					0	0		- €
	Bus					30	0		2.00 €
	Metro / Tram					78	1500		2.00 €
17	Taxi	City	Normal	13:00	Sun	30	0		26.40 €
	C2G					30	0	0	6.00 €
	None					0	0		- €
	Bus					39	800		2.00 €
	Metro / Tram					62	150		2.00 €
18	Taxi	City	Normal	13:00	Sun	30	0		26.40 €
	C2G					34	0	3.5	11.70 €
	None					0	0		- €
	Bus					15	800		2.00 €
	Metro / Tram					12	600		2.00 €
18	Taxi	City	Normal	22:00	Rain	5	0		10.40 €
	C2G					9	0	3.5	2.95 €
	None					0	0		- €
	Bus					37	600		2.00 €
	Metro / Tram					27	1000		2.00 €
18	Taxi	City	Normal	22:00	No Rain	15	0		18.40 €
	C2G					17	150	0	6.00 €
	None					0	0		- €

Choiceset		Szenario				Values			
Series	Alternative	Starting Point	Traffic	Time of Day	Weather	Travel Time incl. Walking, Transfers & Parking	Thereof Walking Distance (in m)	Thereof Time for Parking (in Min)	Price
19	Bus	City	Normal	22:00	No Rain	38	650		2.00 €
	Metro / Tram					34	300		2.00 €
	Taxi					15	0		18.40 €
	C2G					21	500	0	6.00 €
	None					0	0		- €
19	Bus	Home	Normal	22:00	No Rain	42	1000		- €
	Metro / Tram					15	800		- €
	Taxi					5	0		8.00 €
	C2G					7	150	0	1.75 €
	None					0	0		- €
19	Bus	Home	Rushhour	08:00	Rain	49	300		2.00 €
	Metro / Tram					21	500		2.00 €
	Taxi					45	0		31.00 €
	C2G					49	0	3.5	9.70 €
	None					0	0		- €
20	Bus	Home	Normal	13:00	Sun	48	1500		2.00 €
	Metro / Tram					63	1500		2.00 €
	Taxi					30	0		24.00 €
	C2G					37	300	3.5	13.40 €
	None					0	0		- €
20	Bus	Home	Normal	22:00	Rain	38	650		- €
	Metro / Tram					37	600		- €
	Taxi					15	0		16.00 €
	C2G					20	150	3.5	4.65 €
	None					0	0		- €
20	Bus	City	Normal	13:00	Rain	17	150		2.00 €
	Metro / Tram					19	300		2.00 €
	Taxi					15	0		18.40 €
	C2G					17	150	0	6.00 €
	None					0	0		- €

Appendix 5: Estimation Code in R

Please see separated file estimation.R

Appendix 6: Estimation Code in Stata

Please see separated file estimation.do