



TI 2010-014/3

Tinbergen Institute Discussion Paper

# Biases in Willingness-To-Pay Measures from Multinomial Logit Estimates due to Unobserved Heterogeneity

*Vincent van den Berg<sup>1,3</sup>*

*Eric Kroes<sup>1,2</sup>*

*Erik Verhoef<sup>1,3</sup>*

<sup>1</sup> VU University Amsterdam;

<sup>2</sup> Significance BV;

<sup>3</sup> Tinbergen Institute.

### **Tinbergen Institute**

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

### **Tinbergen Institute Amsterdam**

Roetersstraat 31  
1018 WB Amsterdam  
The Netherlands  
Tel.: +31(0)20 551 3500  
Fax: +31(0)20 551 3555

### **Tinbergen Institute Rotterdam**

Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900  
Fax: +31(0)10 408 9031

Most TI discussion papers can be downloaded at  
<http://www.tinbergen.nl>.

# Biases in Willingness-To-Pay Measures from Multinomial Logit Estimates due to Unobserved Heterogeneity

Vincent van den Berg<sup>1#</sup>, Eric Kroes<sup>1,2</sup>, Erik Verhoef<sup>1</sup>

1 Department of Spatial Economics, VU University, De Boelelaan 1105, 1081HV, Amsterdam, The Netherlands

2 Significance BV, Koninginnegracht 23, 2514 AB, The Hague, The Netherlands

# Corresponding author: Email: vberg@feweb.vu.nl;

Tel: +31 20 598 6049;

Fax: +31 20 598 6005

---

## Abstract

It is a common finding in empirical discrete choice studies that the estimated mean relative values of the coefficients (i.e. WTP's) from multinomial logit (MNL) estimations differ from those calculated using mixed logit estimations, where the mixed logit has the better statistical fit. However, it is less clear under exactly what circumstances such differences arise, whether they are important, and if they can be seen as biases in the WTP estimates from MNL. We use datasets created by Monte Carlo simulation to test, in a controlled environment, the effects of the different possible sources of bias on the accuracy of WTP's estimated by MNL. Consistent with earlier research we find that random unobserved heterogeneity in the marginal utilities does not in itself bias the MNL estimates. Furthermore, whether or not the unobserved heterogeneity is symmetrically shaped also does not affect the accuracy of the WTP estimates of MNL. However, we find that if two heterogeneous marginal utilities are correlated then the WTP's from MNL may be biased. If the correlation between the marginal utilities is negative, then the bias in the MNL estimate is negative, whereas if the correlation is positive the bias is positive.

*Keywords:* Discrete Choice, Biases in WTP's, Multinomial Logit, Correlated Heterogeneous Marginal Utilities

---

## 1 Introduction

Multinomial Logit (MNL) models are often used in empirical discrete choice studies. However, as Bhat (1998a) notes, if there is heterogeneity in the marginal utilities and/or in the alternative specific constants, then ignoring this, by estimating a MNL, could lead to biased parameter estimates. With MNL it is only possible to control for observed heterogeneity. For example, the marginal utility of a price attribute depends on income. To control for unobserved heterogeneity, a mixed logit can be used. Still, ignoring heterogeneity by using MNL does not necessarily result in biased estimates. Rizzi and Ortúzar (2006) find that the Values of Risk Reduction (i.e. the WTP's for Risk) in the three surveys they analyse are somewhat lower with MNL. Still, the average WTP's from their mixed logits are within the 95% confidence intervals of the WTP's of the respective MNL's. Train (1998) concludes that there is probably no general answer to whether or not MNL gives correct estimates when heterogeneity is present. Hence, it remains unclear under what circumstances MNL gives biased mean estimates of a Willingness-To-Pay (WTP) (i.e. the ratios of marginal utilities).

This paper examines some circumstances under which heterogeneous marginal utilities might cause biased WTP estimates from MNL. We analyse three cases of heterogeneity, using discrete choice datasets created by Monte Carlo simulation. The first is symmetric random heterogeneity

in the marginal utility. The second is asymmetric heterogeneity. Thirdly, we consider what happens when two heterogeneous marginal utilities are correlated.

Because we create the datasets by means of simulation, we can test the effects of heterogeneous marginal utilities in a clean laboratory-type setting. Conversely, with real world empirical datasets, there are a large number of issues at play, making it difficult to analyse separate issues.

On the simulated datasets we perform MNL estimates. The WTP from a MNL is given by the coefficient of some attribute divided by the coefficient of price. The resulting WTP's are then compared to their design values. The design level for a dataset is calculated by dividing the mean of the simulated marginal utility of an attribute by the mean of marginal utility of price. This is also what Horowitz (1980) compares when he studied how well MNL could approximate a probit with random heterogeneity in the parameters. Given that the true model has random heterogeneity, the coefficients of the MNL estimation should equal the means of the random marginal utilities, ignoring the arbitrary scaling.

Hence, we first simulate the choices with random heterogeneity in the marginal utilities. Then, we do MNL estimates on these simulated datasets. We have not performed mixed logit estimates.

Section 2 begins by discussing some of the literature. Section 3 studies the effect of (non-symmetric) heterogeneity in the marginal utilities and discusses the basic set up of the Monte Carlo simulation, which is also used in the following section. Section 4 studies what happens when two heterogeneous marginal utilities are correlated. Finally, Section 5 concludes.

## 2 Literature discussion

Bhat (1998a) finds for his dataset that the WTP's from mixed logit are on average slightly larger than with MNL. Further, the absolute value of the elasticity of the choice probability to costs is larger with mixed logit than with MNL. Similarly, Bhat (2000) finds that MNL underestimates the WTP's for *out-of-* and *in-vehicle travel time*. Bhat (1998b) notes that the WTP's for *out-of-* and *in-vehicle travel time* are somewhat smaller with his MNL than with his mixed logit.

Train (1998) finds that for his data the WTP's are larger with mixed logit than with MNL. However, the WTP's from his mixed logit with correlated marginal utilities are smaller than those from MNL or the mixed logit without free correlation. He concludes that there is probably no general conclusion as to whether MNL gives good estimates for the WTP's in a given dataset: that the performance of MNL will be different for each dataset.

Algers et al. (1999) note that for their dataset the mean WTP for travel time differs significantly between the MNL and mixed logit. Hensher et al. (2008) find that, for their dataset, the mean price elasticities with nested logit are for some alternatives higher and for others lower than with mixed logit. Van den Berg, Kroes and Verhoef (2009) find for their dataset that MNL underestimates the WTP for travel time compared with mixed logit, but overestimates the WTP's for the other attributes.

The empirical papers often find that MNL estimations give different estimates of the WTP's to mixed logits. However, there is no clear pattern in whether MNL over- or underestimates WTP's. One study finds an overestimation, a second finds an underestimation and a third finds no real difference. These empirical results do, however, contradict the results of the theoretical analysis of Horowitz (1980) using simulated datasets. He analyses the performance of MNL when there is random heterogeneity in the marginal utilities and finds that the ratio of the two coefficients he estimates on his simulated datasets (i.e. the WTP) is for all amounts of heterogeneity in the marginal utilities almost identical to the design value. He concludes that the ratio of the

coefficients is unbiased when one does not control for heterogeneity in the marginal utilities. He finds, however, that the MNL choice probabilities are biased by heterogeneity. The question is: why would unobserved heterogeneity cause biased WTP estimates from MNL? Further, if there is a bias, in what direction will it be?

This paper focuses on the effect of heterogeneity on WTP estimates from MNL. An interesting question, though beyond the scope of this paper, is whether a mixed logit estimation will result in correct WTP estimates. Hess et al. (2005) note that the mean WTP for travel time for their dataset is sensitive to the assumed distribution of the random unobserved components of the cost and time parameters: the mean WTP differs substantially across the different distributions they use. Sillano and Ortúzar (2005) find that the mean WTP can also depend on the used estimation procedure (i.e. Maximum simulated likelihood or Bayesian). Daly, Hess and Train (2009) show that if the price coefficient has a normally distributed random element, and the model is estimated in utility space, then the mean WTP is undefined. This is because the price coefficient then has non-zero probability of being zero and one cannot divide by zero. This issue arises for any distribution of the price coefficient that has a non-zero probability of being zero.

### 3 Non-symmetric marginal utilities

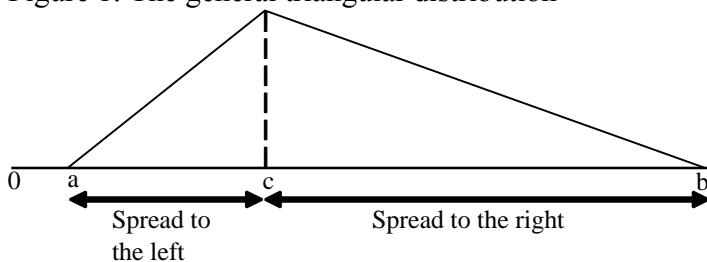
This section tests whether (non-symmetric) heterogeneity in the marginal utilities could bias the WTP's from MNL. The idea is that asymmetry in the heterogeneity might not middle out to the mean of the distribution; for example, because the logit probability function is non-linear.

#### 3.1 Set-up of the dataset simulations

We investigate this through dataset simulations in Gauss 6.0, in which we use several different design forms of the random elements of two marginal utilities. Hereby, we test whether the differences in the WTP's from MNL increase as the random elements of the marginal utilities become more skewed. We simulate 2000 datasets per version of the marginal utilities and then perform MNL estimates on these simulated datasets. Each dataset in a version has the same design. The datasets of another version has, for example, a more skewed distribution of a marginal utility. Each simulated dataset has 500 individuals, who all face one choice situation.

Symmetric triangular distributions for random elements were first applied in a discrete choice study by Train (2001). The triangular distribution is increasingly used in empirical studies using mixed logit (Hensher et al., 2008). In a more general form, the triangular distribution can also allow for non-symmetric distributions. Figure 1 gives an example of a general triangular. The  $a$  is the minimum of the distribution,  $b$  the maximum and  $c$  the mode. With a non-symmetric distribution, the mode is not the same as the mean. The mean of the distribution equals  $(a+b+c)/3$ . If  $(c-a) < (b-c)$ , then the mean is larger than the mode. If  $(c-a) > (b-c)$ , it is smaller.

Figure 1: The general triangular distribution



With a symmetric triangular (i.e. with (c-a)=(b-c)), the c-a is called the *spread*, as it gives the spread of the distribution. We define  $d_k=c_k-a_k$  (i.e. the (negative) spread to the left) and  $e_k=b_k-c_k$  (i.e. the spread to the right). For the non-symmetric version, the marginal utilities ( $\beta_{kq}$ ) are given by (1). In the formula  $c_k$  is the mode of  $k$ th marginal utility.  $S_{kq}$  is a standard uniform draw for the  $k$ th marginal utility of individual  $q$ .

$$\begin{aligned}\beta_{kq} &= c_k + \sqrt{S_{kq} * (e_k + d_k) * (d_k)} - d_k & \text{if } S_{kq} \leq d_k / (e_k - d_k) \\ \beta_{kq} &= c_k + e_k - \sqrt{(1-S_{kq}) * (e_k + d_k) * e_k} & \text{if } S_{kq} > d_k / (e_k - d_k)\end{aligned}\tag{1}$$

We now introduce the layout of the datasets we create by simulation. We use general triangular distributions for the marginal utilities (i.e. the marginal utilities can be non-symmetric). There are three alternatives. The utility of alternative  $i$  for individual  $q$  depends on three explanatory variables (vector  $\mathbf{X}_{iq}$ ), an unobserved element  $\varepsilon_{iq}$  and the individual  $q$  specific parameter vector ( $\beta_q$ ). The parameters are the same for all alternatives. The parameters are also referred to as marginal utilities. The utility functions are given in (2a-c).

$$U_{1q} = \beta_{1q} * X_{11q} + \beta_{2q} * X_{21q} + \varepsilon_{1q}\tag{2a}$$

$$U_{2q} = \beta_{1q} * X_{12q} + \beta_{2q} * X_{22q} + \beta_{3q} * X_{32q} + \varepsilon_{2q} + \text{ASC}_2\tag{2b}$$

$$U_{3q} = \varepsilon_{3q} + \text{ASC}_3\tag{2c}$$

The first number in subscript behind the X variables ( $k=1,2,3$ ) indicates which variable is used (e.g. whether it is the price or travel time). The second number indicates the alternative ( $i=1,2,3$ ). Finally,  $q$  ( $q=1,2,\dots,500$ ) indicates the individual. For ease of interpretation, we view  $X_{1iq}$  as the price. The  $\varepsilon_{iq}$ 's are randomly generated to be Independently and Identically Distributed, with an Extreme Type I distribution form. Discrete choice models use the assumption of utility maximization. Thus, a simulated individual chooses the alternative which gives the highest utility. The utility of the third alternative is not influenced by any attribute. The alternative can be seen as an ‘‘opting out’’ alternative.

The  $X_{11q}$  and  $X_{12q}$  (i.e. price) variables are created from the variable  $Z_{1q}$  and the  $X_{21q}$  and  $X_{22q}$  variables from the variable  $Z_{2q}$ . The  $X_{1iq}$  ( $i=1$  or  $2$ ) variables are created from  $Z_{1q}$  by a randomly generated difference variable, which can take the values of 10%, 20%, 30% and 40%. Hence,  $Z_{1q}$  times the difference variable gives the difference between the two  $X_{1iq}$  variables. This difference is allocated to an increase relative to  $Z_{1q}$  for  $X_{11q}$  and a decrease for  $X_{12q}$ . This allocation is determined by the standard uniform random variable  $r_{1q}$ . For  $X_{21q}$  and  $X_{22q}$  these variables are difference relative to  $Z_2$  (i.e.  $\text{diff}_{2q}$ ) and  $r_{2q}$ . Here  $\text{diff}_{2q}$  can take four values, namely 0%, -10%, -20% and -30%. Hence, the  $X_{21q}$  for alternative one is always lower or equal to the  $X_{22q}$  for alternative two. The variables are determined by the formulas (3a-d). The design levels and forms for  $Z_{1q}$  and  $Z_{2q}$  and  $X_{32q}$  are shown in Table 1. Note that  $X_{32q}$  has a uniform distribution, which starts at mean – spread/2 = 5 and ends at mean + spread/2 = 25.

$$X_{11q} = Z_{1q} + Z_{1q} * \text{diff}_{1q} * r_{1q}\tag{3a}$$

$$X_{12q} = Z_{1q} - Z_{1q} * \text{diff}_{1q} * (1 - r_{1q})\tag{3b}$$

$$X_{21q} = Z_{2q} + Z_{2q} * \text{diff}_{2q} * r_{2q}\tag{3c}$$

$$X_{22q} = Z_{2q} - Z_{2q} * \text{diff}_{2q} * (1 - r_{2q})\tag{3d}$$

Table 1: The design of the variables of the simulated datasets

	Mean	Standard Deviation	Spread	Distribution shape
$Z_{1q}$	10	5	.	Lognormal
$Z_{2q}$	70	40	.	Lognormal
$X_{32q}$	15	.	20	Uniform

The marginal utilities and Alternative Specific Constants (ASC's) are determined by the design in Table 2. The spread to the left follows from the design variable  $Min\_k$ , the spread to the right follows from  $Max\_k$ . The spread to the left is given by  $Min\_k * mode\_k$  and to the right by  $Max\_k * mode\_k$ . The mode of a distribution is set so that, given the spread, the mean of the marginal utility equals the design mean. Thus, the expected value of a marginal utility is the same for each design version of the dataset, which helps with interpreting the results.

Table 2: The designs of the marginal utilities and ASC's

	Mean	Spread to the left	Spread to the right	Distribution shape
$\beta_{1q}$	-0.15	Varies	Varies	Triangular
$\beta_{2q}$	-0.025	Varies	Varies	Triangular
$\beta_3$	-0.02	0	0	Fixed
$ASC_2$	-0.4	0	0	Fixed
$ASC_2$	-4.5	0	0	Fixed

### 3.2 Calculation of the design levels of the WTP's

Table 3 shows that we use 18 different combinations of the  $Min\_k$  and  $Max\_k$  design variables. We generate 2000 different datasets per version, where each dataset has new values for the explanatory variables, marginal utilities and unobserved elements. Hence, for Table 3 we performed 36000 MNL estimations. Each dataset contains 500 individuals, who face one choice situation. Each version has a different design. In this section the difference between versions is in the skewedness of the heterogeneous marginal utilities. We simulated the datasets in Gauss 6.0 using the pseudo-random number generator and performed MNL estimates with the maxlik module. The starting values for the coefficients and ASC's were, for each dataset, set randomly within certain bounds. The bounds were different for each coefficient and ASC.

We calculate for each dataset the WTP's for  $X_{2i}$  and  $X_{32}$  (i.e.  $\beta_2/\beta_1$  and  $\beta_3/\beta_1$ ), and then calculate the averages per version. These averages are reported in Table 3. We also give the average design levels of the WTP's and the average relative sizes of the two. This last figure equals the mean of the estimated WTP divided by the design WTP.

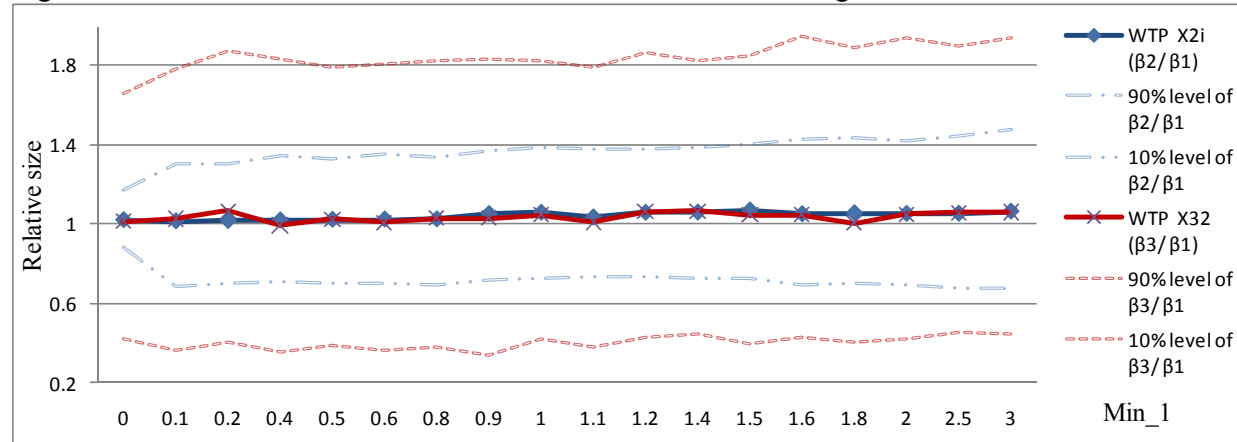
The design level of a WTP for a dataset is calculated by dividing the dataset mean of  $\beta_{kq}$  by the mean  $\beta_{1q}$ . Given the assumption that the true model has heterogeneous marginal utilities, the coefficients of the MNL estimation should be equal to the means of the heterogeneous marginal utilities. Hence, a WTP for an attribute from a MNL should equal the mean of the simulated marginal utility of that attribute divided by the mean marginal utility of price (i.e. it should equal  $E[\beta_{kq}]/E[\beta_{1q}]$ ). Accordingly, the WTP from MNL is not necessarily the same as the mean WTP from mixed logit. The mean WTP of an attribute with mixed logit is given by the average of the division of its marginal utility by the marginal utility of price (i.e. by  $E[\beta_{kq}/\beta_{1q}]$ ). The WTP from a MNL estimate equals the coefficient of the attribute divided by the price coefficient.

Table 3: Results of estimations on the datasets with non-symmetric random marginal utilities.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
<b>Design variables</b>																		
Min_1	0	0.1	0.2	0.3	0.5	0.6	0.8	0.9	0.999	1	1.1	1.2	1.4	1.5	1.6	2	2.5	3
Min_2	0	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
Max_1	0	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
Max_2	0	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
<b>WTP variable <math>X_{2iq}</math> (<math>\beta_2/\beta_1</math>)</b>																		
Estimated	0.170	0.169	0.169	0.170	0.170	0.170	0.170	0.175	0.176	0.172	0.176	0.176	0.178	0.175	0.176	0.175	0.175	0.177
Design value	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167
Relative value	1.018	1.012	1.015	1.017	1.020	1.022	1.023	1.049	1.055	1.032	1.056	1.056	1.067	1.048	1.053	1.050	1.050	1.061
<b>WTP variable <math>X_{32q}</math> (<math>\beta_3/\beta_1</math>)</b>																		
Estimated	0.135	0.137	0.142	0.132	0.136	0.134	0.137	0.137	0.140	0.135	0.141	0.142	0.139	0.139	0.134	0.140	0.141	0.141
Design value	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133
Relative value	1.013	1.027	1.065	0.992	1.023	1.008	1.030	1.028	1.047	1.010	1.061	1.064	1.045	1.045	1.006	1.050	1.058	1.060

Note: Min\_1 multiplied by the mode gives the spread to the left for  $\beta_{1q}$ . Hence, the minimum is at  $(1-\text{Min}_1) \cdot \text{mode}_1$ . The spread to the right for the second random parameter is equal to Max\_2 multiplied by the mode of the random parameter.

Figure 2: Relative values estimated WTP's in Table 3 to their design values over the values of  $\text{Min}_1$



Note: Min\_1 multiplied by the mode gives the spread to the left for  $\beta_{1q}$ . Hence, the minimum is at  $(1-\text{Min}_1) \cdot \text{mode}_1$ . In the most left column there is not heterogeneity in the marginal utilities. In the other columns, if Min\_1 is smaller than one, the distribution of the marginal utility of price is skewed to the left. If Min\_1 is larger than one it is skewed to the right. If it equals one, then the distribution is symmetric.



### 3.3 Results of the dataset simulations with non-symmetric heterogeneity

The first part of Table 3 shows that  $\beta_{2q}$  is given a symmetric triangular distribution. The marginal utility of  $X_{1q}$  (i.e.  $\beta_{1q}$ ) is given an asymmetric distribution. The *spread to the right* equals  $0.999 \cdot \text{mode}$  (except in the most left column) and the *spread to the left* varies from 0 to 3. In the most left column, all simulated marginal utilities are fixed and there is no heterogeneity. In the other columns, if  $\text{Min}_1$  is smaller than one, the distribution of the marginal utility of price is skewed to the left. If  $\text{Min}_1$  is larger than one it is skewed to the right. If it equals one, then the distribution is symmetric. Both marginal utilities remain negative because the maximum is  $(1-0.999) \cdot \text{mode}_k$  and we use a negative mode. The second part of the table gives, for each version of heterogeneity, the means of the estimated WTP's for variable  $X_{2q}$ , the design level of this WTP and the relative size of the two. The third part does the same for the WTP of  $X_{3q}$ .

Figure 2 depicts the relative sizes of the estimated WTP's to their design levels. It also shows the 90% interval of the estimates, that is the interval within which 90% of the WTP estimates lie. This range gives an indication of the accuracy of the simulation and helps in determining whether a mean estimated WTP is really different from its design value.

How skewed the distributions are seems to have no effect on the mean of the WTP's of the 2000 datasets per version. For relatively strong asymmetry, the estimated WTP's are approximately equal to the design values, and for small amounts of asymmetry this is also the case. The mean WTP's vary a bit over the table. However, there is no clear pattern of effect of (non-symmetric) heterogeneity. Further, heterogeneity in itself also does not seem to bias the MNL estimates. The heterogeneity does cause the estimates to have a larger spread and makes the standard errors of the estimates larger.

This exercise suggests that the estimated WTP's of a MNL estimation are not affected by heterogeneity in a marginal utility, even if this heterogeneity is asymmetric. These results are consistent with Horowitz (1980), who finds that heterogeneity in itself does not necessarily bias WTP's from estimations that ignore the heterogeneity (such as MNL)

## 4 Correlation between two heterogeneous marginal utilities

The previous section showed that heterogeneity in itself should not bias MNL estimates, even if the heterogeneity is non-symmetric. This section studies the effect of correlated heterogeneous marginal utilities. For this we perform MNL estimates on 4000 simulated datasets per version of the strength of the relation between the two heterogeneous marginal utilities. Thereafter, we compare the means and 90% intervals of the estimated WTP's to the design levels. We first consider some examples of correlated marginal utilities. Thereafter, we examine the set-up of the dataset simulations and then the results of the simulations.

### 4.1 Examples of correlated heterogeneity in marginal utilities

Consider the value of statistical life. Suppose that for richer people the marginal utility of money is lower, because of a decreasing marginal utility of income. Further, suppose that the marginal disutility of the risk of an accident increases with income, because richer people perceive that they have more to lose. Then, of course, these two marginal utilities are correlated. Similarly, the marginal utility of travel time could be correlated with the marginal utility of price.

In a study on the purchase of appliances (see Revelt and Train (1998) for a study on refrigerator purchase) the marginal utility of the monetary saving due to the efficiency of the appliance might be correlated with the marginal utility of the efficiency level of the appliance. Train (2007) studies the choice of electricity supplier and finds significant covariance between

the marginal utility of whether the rates are differentiated over the day and the marginal utilities of price, whether the electricity supplier is a known company and if the electricity price differs over the year. Train (1998) notes that, for his dataset, the WTP's from mixed logit with correlated marginal utilities (i.e. free-covariance mixed logit) are lower than the WTP's from regular (no-covariance) mixed logit and MNL. Hence, not controlling for correlation can have an effect.

#### 4.2 Set-up dataset simulations with correlated heterogeneity

This section analyses the effects on a MNL estimate of unobserved heterogeneity in the marginal utilities when there is correlation between the marginal utilities. For this we use the same basic set up as in Section 3 and the utility functions in (3a-c). The random elements of the parameters are again determined by Table 1. The means of the marginal utilities are the same as before. However, now we only use symmetric distributions for the marginal utilities, with the design spreads equal to the mean. The difference with Section 3 is that now the heterogeneous marginal utilities of  $X_{1iq}$  and  $X_{2iq}$  are related. In particular they are given by (4) and (5).

$$\beta_{1q} = a_1 + Spread_1 * T_{1q} \quad (4)$$

$$\beta_{2q} = \left( a_2 * \frac{\rho}{(\beta_{1q} * \text{mean}(1/\beta_{1q}))} \right) + (1 - \rho) * (a_2 + Spread_2 * T_{2q}) \quad (5)$$

The  $a_2$  is the design mean of the marginal utility of  $X_{2iq}$  (i.e.  $\beta_{2q}$ ) and  $spread_{2q}$  gives the spread of the random element of this marginal utility. The strength of the relationship between  $\beta_{2q}$  and  $\beta_{1q}$  is given by  $\rho$ , which is between zero and one. The  $T_{kq}$  is a random variable with a triangular distribution, with  $-1 \leq T_{kq} \leq 1$  and a zero mean<sup>1</sup>. Formula (5) shows that the larger  $\rho$ , the stronger the negative correlation between the marginal utility of  $X_{1iq}$  and  $X_{2iq}$ .

#### 4.3 Results with an decreasing relation between the simulated marginal utilities

To facilitate interpretation,  $X_{1iq}$  is again seen as the price. To analyse the effect of the correlated heterogeneous marginal utilities, we compare the estimated WTP's from MNL with the design WTP's. The design levels of the WTP's are found by dividing the dataset mean of  $\beta_{iq}$  by the mean  $\beta_{1q}$ . Table 5 shows the results of the 4000 datasets per level of  $\rho$ . There are eight different values for  $\rho$ , hence there are eight versions. Thus, we performed 32000 dataset simulation and MNL estimates for Table 5. In each dataset there are 500 individuals. Figure 5 shows the relative values of the estimated WTP's for  $X_{2iq}$  to the design values of the WTP. The dashed lines in the figure give the area in which 90% of the estimations lie. Figure 6 does this for the WTP of  $X_{32q}$ .

Table 5 also gives the average correlation between  $\beta_{1q}$  and  $\beta_{2q}$  per value of  $\rho$ . Even for  $\rho=1$  the linear correlation coefficient is not  $-1$ , as  $\beta_{2q}$  is related to the inverse of  $\beta_{1q}$ . The strongest correlation we simulate is  $-0.6$ . Revelt and Train (1998) find correlations between the heterogeneous marginal utilities as large as  $+0.59$  and  $-0.68$ . Train (2007) finds correlations between  $0.1$  and  $0.94$ , with three correlation coefficients being larger than  $0.9$ . This indicates that the strength of the correlation patterns we simulate can occur in the empirical real world.

The negative effect on the accuracy of MNL estimates of the correlated heterogeneous marginal utilities is not limited to the non-linear (inverse) relation. When we reran the simulation

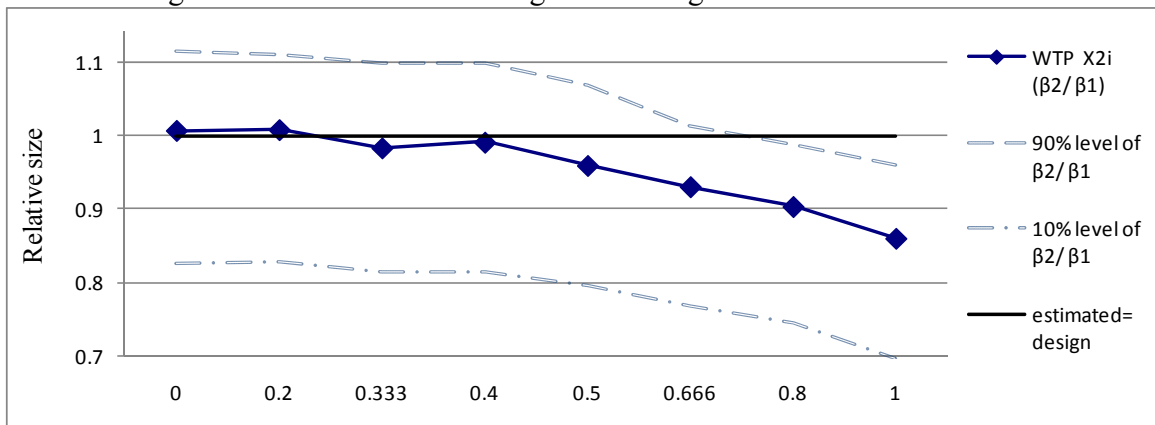
<sup>1</sup> The  $Spread_2$  (i.e. the spread of the random triangular of  $\beta_{2q}$ ) is different for each level of  $\rho$ . This to compensate for the effect of the correlation with  $\beta_{1q}$  on the total spread of  $\beta_{2q}$  and to keep the total spread the same for each version.

using a linear decreasing relationship between  $\beta_{1q}$  and  $\beta_{2q}$ , we found that this could also bias the WTP's from MNL.

Table 5: MNL estimations when the heterogeneous marginal utilities are negatively correlated

	1	2	3	4	5	6	7	8
<b>Design variables</b>								
Strength of relation between $\beta_{1q}$ and $\beta_{2q}$ ( $\rho$ )	0	0.2	0.333	0.4	0.5	0.666	0.8	1
Resulting correlation between $\beta_{1q}$ and $\beta_{2q}$	-0.00	-0.32	-0.46	-0.51	-0.55	-0.57	-0.57	-0.60
<b>WTP variable (<math>X_{2iq}</math>) (<math>\beta_2/\beta_1</math>)</b>								
Estimated	0.168	0.168	0.164	0.165	0.160	0.155	0.151	0.143
Design value	0.167	0.167	0.167	0.167	0.167	0.167	0.167	0.167
Relative value	1.007	1.008	0.983	0.991	0.959	0.930	0.903	0.860
<b>WTP variable <math>X_{32q}</math> (<math>\beta_3/\beta_1</math>)</b>								
Estimated	0.152	0.143	0.143	0.142	0.137	0.137	0.139	0.150
Design value	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133
Relative value	1.138	1.070	1.070	1.068	1.029	1.031	1.045	1.126

Figure 5: Relative sizes estimated WTP's for  $X_{2iq}$  to their design values for different strengths of the decreasing relation between the heterogeneous marginal utilities



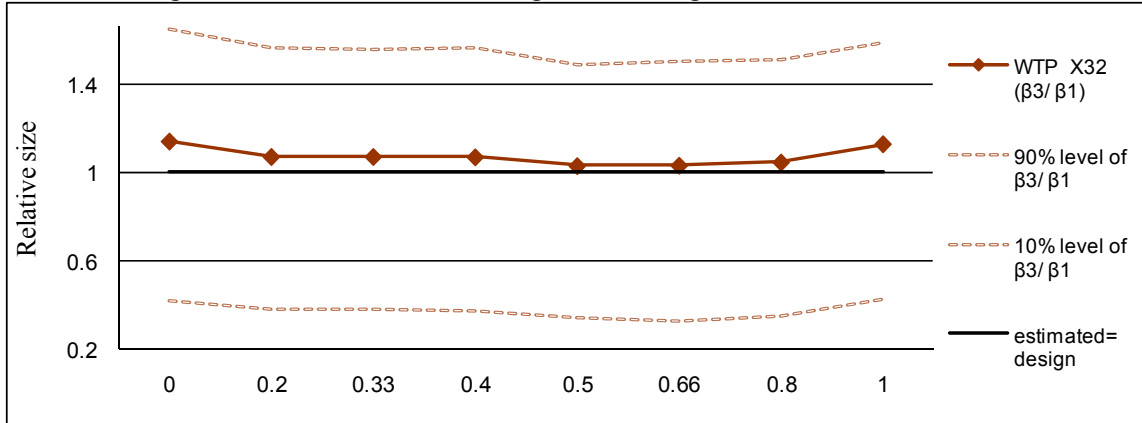
The estimated WTP's for the third attribute  $X_{32q}$  in Figure 6 and Table 5 seem to be somewhat above the design levels. However, for this WTP there is no clear pattern in the effect of the correlation between the two marginal utilities. Furthermore, the unity line, at which estimated divided by the design is one, is well inside the 90% interval. Hence, we cannot conclude that this WTP is affected by the correlation between the heterogeneous marginal utilities. The fixed marginal utility of  $X_{32q}$  is of course uncorrelated with the marginal utility of price. Again, the 90% interval for the WTP's of  $X_{32q}$  is larger than the 90% interval for the WTP's of  $X_{2iq}$ . This seems to be caused by the larger variance in the estimates of the coefficient of  $X_{32q}$ .

The stronger the relationship between the two random marginal utilities, the larger the difference between the design values and the MNL estimates of the WTP's of  $X_{2iq}$  seems to be and the larger estimated WTP. Figure 5 shows that if the relationship between the two heterogeneous marginal utilities is not that strong (say  $\rho < 0.4$ ) then the MNL estimates are not that strongly affected. Conversely, for larger values of  $\rho$  there is a clear underestimation.

However, the relationship between the two marginal utilities has to be rather strong ( $\rho \geq 0.8$ ) for the entire 90% interval of relative values of estimated WTP's to be below the

“estimated=design” line (i.e. where the relative value is 1). This indicates that correlated heterogeneous marginal utilities can bias WTP’s estimated with MNL, where the problem is most likely to occur if the relationship between the two marginal utilities is strong.

Figure 6: Relative sizes estimated WTP’s for  $X_{32q}$  to their design values for different strengths of the decreasing relation between the heterogeneous marginal utilities



#### 4.4 Results with an increasing relation between the simulated marginal utilities

Table 5 shows that if the marginal utility of  $X_{2iq}$  is, following formula (5), a decreasing function of the marginal utility of  $X_{1iq}$ , a MNL can result in an underestimation of the WTP of  $X_{2iq}$  (i.e.  $\beta_2/\beta_1$ ). Conversely, if  $\beta_{2q}$  is an increasing function of  $\beta_{1q}$ , following (6), this can result in an overestimation by a MNL of this WTP.

$$\beta_{2q} = \left( a_2 * \rho * \left( 2 - \frac{1}{(\beta_{1q} * \text{mean}(1/\beta_{1q}))} \right) \right) + (1 - \rho) * (a_2 + \text{Spread}_2 * T_{2q}) \quad (6)$$

Table 6: MNL estimations when the heterogeneous marginal utilities are positively correlated

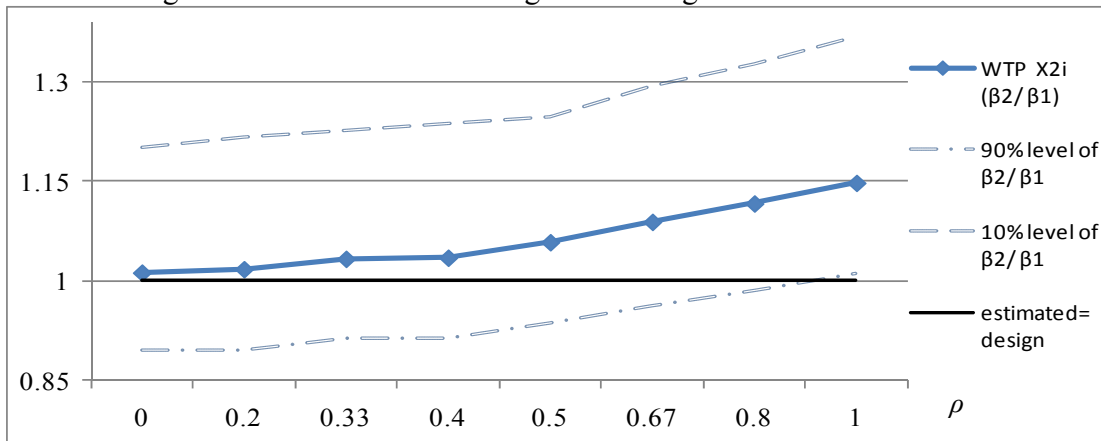
	1	2	3	4	5	6	7	8
Strength of the relationship between $\beta_{1q}$ and $\beta_{2q}$ ( $\rho$ )	0	0.2	0.333	0.4	0.5	0.666	0.8	1
Resulting correlation between $\beta_{1q}$ and $\beta_{2q}$	-0.00	0.32	0.47	0.51	0.54	0.59	0.59	0.60
<b>WTP variable <math>X_{2iq}</math> (<math>\beta_2/\beta_1</math>)</b>								
Estimated	0.169	0.170	0.172	0.172	0.176	0.181	0.186	0.191
Design value	0.167	0.167	0.167	0.167	0.167	0.167	0.166	0.167
Relative value	1.012	1.017	1.033	1.034	1.058	1.089	1.116	1.147
<b>WTP variable <math>X_{32q}</math> (<math>\beta_3/\beta_1</math>)</b>								
Estimated	0.152	0.157	0.158	0.153	0.161	0.156	0.161	0.158
Design value	0.133	0.133	0.133	0.133	0.133	0.133	0.133	0.133
Relative value	1.139	1.178	1.184	1.148	1.211	1.167	1.208	1.186

The results of this simulation, with 4000 different created datasets per level of  $\rho$ , are shown in Table 6 and Figure 7. A stronger (increasing) effect of  $\beta_{1q}$  on  $\beta_{2q}$  raises the estimated WTP of  $X_{2iq}$ , even though the design value of the WTP is always the same. Unlike the previous

Subsection 4.2, the 90% interval of the estimated values is now only above the design value line (i.e. one) for the strongest relationships between the marginal utilities ( $\rho=1$ ). Thus, the correlation now has to be rather strong for the estimated WTP to substantially differ from its design value.

A solution to the problem caused by correlated marginal utilities might be to estimate a mixed logit that allows the marginal utilities to be correlated (i.e. with free-covariance). For future research it would be interesting to study how mixed logit without free covariance performs when the heterogeneous marginal utilities are correlated.

Figure 7: Relative sizes estimated WTP's for  $X_{2i}$  to their design values for different strengths of the increasing relation between the heterogeneous marginal utilities



#### 4.5 Conclusions on the effect of correlated marginal utilities

This section found that if two heterogeneous marginal utilities are correlated with each other, this can bias the WTP's from MNL estimations. The absolute value of this bias seems to increase with the strength of the correlation. If the heterogeneous marginal utility of the attribute to be valued (e.g. travel time) is an increasing function of the marginal utility of the monetary variable, this seems to cause a positive bias in the WTP from MNL. Conversely, if there is a decreasing relationship between the two heterogeneous marginal utilities, this can result in a negative bias in the WTP. This is for both linear and non-linear relationships between the marginal utilities. The strengths of the correlations between the marginal utilities we simulated, with maximum correlation coefficients of  $\pm 0.60$ , have also been observed in empirical studies using free-covariance mixed logit.

## 5 Conclusion

This paper studied three situations in which using MNL models might result in biased Willingness-To-Pay (WTP) estimates. First, random heterogeneity in the parameters. Second, non-symmetric heterogeneity. Third, correlation between the heterogeneous marginal utilities.

To study the effect of heterogeneous marginal utilities on the WTP estimates from MNL, we compare the MNL estimates of the WTP's with the design WTP's of a large number of simulated datasets. We find that if two heterogeneous marginal utilities are related, this can result in biased WTP estimates by MNL. In contrast, uncorrelated heterogeneity in the marginal utilities seems to have no detrimental effect on the MNL estimations. This is the case for both symmetric and non-symmetric heterogeneity. We find that the stronger the relationship between two heterogeneous marginal utilities is, the larger the bias in the WTP from MNL. This is the case for both linear

and non-linear relations between the marginal utilities. If the relationship between the two marginal utilities is increasing (i.e. if the two marginal utilities are positively correlated), the resulting bias is positive. If the relation is decreasing, the bias is negative.

This could explain the different results reported in the empirical literature as to whether MNL estimates give different WTP estimates to mixed logit. Our results suggest that if the MNL estimates are biased, the sign and size of this bias could be determined by the correlation pattern of the heterogeneous marginal utilities in the population. Thus, if the correlation pattern is different, one will find a different bias. If there is no correlation between the heterogeneous marginal utilities the MNL estimates of the WTP's might be unbiased. An interesting extension of this paper is testing if using a no-covariance or free-covariance mixed logit eliminates the bias caused by correlated marginal utilities.

### **Acknowledgements**

This paper is financially supported by the project *Betrouwbaarheid van transportketens* of Transumo. Transumo (TRANSition SUSTainable MOBility) is a Dutch platform for companies, governments and knowledge institutes that cooperate in the development of knowledge with regard to sustainable mobility. The authors are grateful for the comments of Muhammad Sabir and Eva Gutiérrez-i-Puigarnau.

### **Literature**

- Algers, S., Bergstrom, M., Dahlberg, M. and Dillen, J., 1999. Mixed logit estimation of the value of travel time. Uppsala University working paper, retrieved on 23 October 2009 from <http://www.nek.uu.se/Pdf/1998wp15.pdf>
- Bhat, C. R., 1998a. Accommodating variations in responsiveness to level-of-service variables in travel mode choice models. *Transportation Research A* 32, 455–507.
- Bhat, C. R., 1998b. Accommodating flexible substitution patterns in multi-dimensional choice modeling: formulation and application to travel mode and departure time choice. *Transportation Research Part B* 32(7), 455-466.
- Bhat, C. R., 2000. Incorporating observed and unobserved heterogeneity in urban work travel mode choice modeling. *Transportation Science* 34(2), 228-239.
- Daly, A., Hess, S. and Train, K., 2009. Assuring finite moments for willingness to pay in random coefficient models. University of California working paper, retrieved on 23 October 2009, from [http://elsa.berkeley.edu/~train/DHT\\_WTP.pdf](http://elsa.berkeley.edu/~train/DHT_WTP.pdf)
- Hensher, D. A., Rose, J. M. and Greene, W. H., 2008. Combining RP and SP data: biases in using the nested logit trick - contrasts with flexible mixed logit incorporating panel and scale effects. *Journal of Transport Geography* 16 (2), 126-133.
- Hess, S., Axhausen, K.W. and Polak, J.W., 2006. Distributional Assumptions in Mixed Logit Models, paper presented at the 85<sup>th</sup> Annual Meeting of the Transportation Research Board, Washington D.C. Retrieved on 23 October 2009 from [http://www.stephanehess.me.uk/files/Hess\\_Axhausen\\_Polak\\_2006.pdf](http://www.stephanehess.me.uk/files/Hess_Axhausen_Polak_2006.pdf)
- Horowitz, J., 1980. The accuracy of the multinomial logit model as an approximation to the multinomial probit model of travel demand. *Transportation Research B* 14, 331-342.
- Rizzi, L. I. and Ortúzar, J de D., 2006. Road safety valuation under a stated choice framework. *Journal of Transport Economics and Policy* 40(1), 69–94.
- Revelt, D. and Train, K.E., 1998. Models with taste differences over people. *Land Economics* 74(2), 230-239.
- Sillano M., Ortúzar, J. de D., 2005. Willingness-to-pay estimation with mixed logit models: some

new evidence. *Environment and Planning A*, 37, 3, 525-550

Train, K.E., 2001. A comparison of hierarchical Bayes and maximum simulated likelihood for mixed logit. University of California working paper, retrieved on 7 April 2008, from <http://elsa.berkeley.edu/~train/compare.pdf>

Train, K.E., 2007. A recursive estimator for random coefficient models. University of California working paper, retrieved on 15 April 2008, from <http://elsa.berkeley.edu/~train/re.pdf>

Van den Berg, V., Kroes, E. and Verhoef E.T., 2009. Choice of train ticket: a study on Dutch travellers. VU University working paper; available on request.