RE-ESTIMATING UK APPRAISAL VALUES FOR NON-WORK TRAVEL TIME SAVINGS USING RANDOM COEFFICIENT LOGIT MODEL

Jeff Tjiong

CH2M Hill, London UK

1 INTRODUCTION

For many years, the high computing cost of estimating models have created significant obstacles for discrete choice analysts to apply more flexible model specifications including multinomial probit (MNP) and mixed multinomial logit (MMNL) model. Amongst these two advanced models, MMNL is gaining more attention in particular and has become the current state-of-the-art approach (Hess, 2005, Hess et al., 2005), largely facilitated by advances in computing power and simulation techniques over the past 10 years. Its advantage also lies in the capability to replicate the correlation structure of any type of GEV model (McFadden and Train, 2000).

In the field of VTTS valuation, MMNL model has been applied to capture random taste heterogeneity across respondents (i.e. variations of VTTS amongst respondents) using a mixture logit model. This MMNL model assumes a mixing distribution to represent the distribution of VTTS, which is estimated as the ratio between the marginal utilities of travel time and cost within a preference space framework (Algers et al., 1998, Hess et al., 2005, Hess et al., 2008). This research was motivated by these preceding MMNL model applications and aimed to compare VTTS estimation outcomes generated by the advanced choice model and basic choice model.

Two non-working trip purposes, namely, commuting trips and other trips were modelled in both AHCG (1999) and Mackie et al. (2003) with similar model specifications. For this research, only commuting trips were modelled to narrow the focus such that different MMNL techniques can be exploited to demonstrate the change of VTTS estimates generated by the advanced choice models. Note that this research solely focuses on the comparison of choice modelling results and is independent of the on-going VTTS updates being undertaken by the UK Department for Transport (DfT).

The layout of the paper is as follows: **Chapter 2** first presents an overview of discrete choice theories and model concepts. **Chapter 3** summarizes the MNL model specification and findings from the 2001 UK VTTS study. **Chapter 4** provides a summary of findings consolidating all the model results as well as issues involved throughout the modelling process. **Chapter 5** concludes**.**

2 THEORETICAL UNDERPINNINGS

2.1 MMNL model

The central idea of the probability mixture model is to mix up different standard restrictive functions to generate more flexible functions. Essentially, MMNL probability distribution takes integrals of logit probabilities over a density imposed by researcher. It is expressed as:

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta \quad (1)$$

Where $L_{ni}(\beta)$ is the logit probability evaluated at parameter β while $f(\beta)$ is the mixing distribution. Assuming the systematic part of utility takes into a linear-form of $V_{ni}(\beta) = \beta' x_{ni}$, the MMNL probability thus becomes:

$$P_{ni} = \int \left(\frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}} \right) f(\beta) d\beta \quad (2)$$

The MMNL logit probability is effectively a weighted average of the logit formula weighted by mixing density $f(\beta)$ (Train, 2009). There are two types of MMNL model specifications, namely, random coefficient logit (RCL) model and error component (ECL) model. They are mathematically equivalent in modelling terms but differ in their functional forms. RCL model is a more commonly used structure which is more suitable for this research to capture random taste variation by assuming the parameter vector β randomly distributed according to a mixing distribution. On the other hand, ECL models are mostly specified to model appropriate substitution pattern by allowing inter-alternative correlation.

The mixing function $f(\beta)$ could be a continuous or discrete distribution. Continuous distribution is commonly used for discrete choice analysis using MMNL but the integrals in choice probability are not in closed-form and thus simulation is required which could potential require very long computing time (Ben-Akiva and Bolduc, 1996, Hensher and Greene, 2003, Hess, 2005, Revelt and Train, 1998, Revelt and Train, 2000). Furthermore, the process of finding the “true” mixing distribution is not straightforward, which is described further in Chapter 4. The use of discrete mixing distribution is emerging as a substitute of the continuous mixture model, which has primary advantage over continuous logit mixture model as its probability distribution does not require an integral though it is still not widely used currently (Hess, 2012).

MMNL also accommodates correlation between repeated choices (i.e. panel data) in *a prior* treatment. This is rather important for this research since individuals always provide multiple responses in stated choice (SC) experiment. MMNL incorporates calculations for panel data effect by involving a product of logit formulas for each time period (i.e. all the SC games participated by each respondent), instead of simply one logit formula at time (Train, 2009). This panel data treatment is superior to the re-sampling techniques (e.g. jack-knifing, bootstrapping) that were used in the past for *ex-post* correction of repeated choices. It is noted that AHCG has incorporated the first-ever 'jack-knife' error correction for VTTS valuation in 1999 (Daly et al., 2013).

2.2 Empirical valuation of VTTS

In discrete choice analysis, VTTS can be computed by taking the ratio of the partial derivatives of the marginal utility of time and cost:

$$VOT = \frac{\partial V / \partial TT}{\partial V / \partial TC} \tag{3}$$

Where V is the systematic part of utility, TT is the travel time attribute and TC is the travel cost attribute. In a typical utility function with linear-in-attributes and fixed-taste coefficients: $V = \beta_{TT}TT + \beta_{TC}TC$, VTTS can be computed as β_{TT}/β_{TC}, which is the marginal rate of substitution between travel time and cost at constant utility, or in other words, the willingness-to-pay to increase travel time in one unit. It has been assumed that the derivative of the random error with respect to travel time TT and travel cost TC is zero, which means all the effects of travel time and cost are captured in the systematic part of utility (Hess et al., 2005).

In choice analysis using RCL model, time and/or cost coefficients are now random coefficients with a distribution which means willingness-to-pay (WTP) indicator is also a distribution itself. Problem arises when cost coefficient enters as denominator in the WTP calculation. Typically simulation technique is employed to estimate the WTP when both time and cost coefficients are random coefficients. Depending on the distributional assumption of the cost coefficient, the WTP may or may not have finite moments (Daly et al., 2012).

3 CURRENT UK VTTS

3.1 1994 stated choice survey design

The current UK VTTS were estimated based on the 1994 stated choice experiment conducted by Accent Marketing and Research and Hague Consulting Group (AHCG), which surveyed 766 respondents in total. These respondents were segregated into 12 different groups, separating the respondents into three road user types and four different distance bands. 8 pairwise comparisons between time and cost variables were set out for each questionnaire based on the relative changes to current journey (i.e. Δc, Δt). In one of the alternatives, each of Δc, Δt is set to zero. 9 attribute levels were allocated for Δt (-20,-15,-10,-5,-3,+5,+10,+15,+20min) while a wide range of values between -300p and +300p were distributed with majority within the range between -100p to 100p. A detailed summary of the experiment design specification can be found in Appendix A of Bates and Whelan (2001).

3.2 1999 and 2003 VTTS valuations

An international seminar was held in 1996 to discuss the findings of the VTTS valuation conducted by AHCG following the data collection effort. The output of the VTTS valuation by AHCG was published in 1999 alongside independent reviews by John Bates, Tony Fawkes, Mark Wardman and Stephen Glaister (AHCG, 1999). In 2001, Leeds ITS was commissioned to review the original work conducted by AHCG. As a product of this review work, a total of six working papers were produced and a new set of VTTS was re-estimated using MNL model with the same set of discrete choice data. It should be noted that although findings from the 1999 AHCG report have been significant updated by Leeds ITS in 2001, the SC design and VTTS valuation techniques were considered as state-of-the-art at the time (Nellthorp, 2001). The final choice model recommended by Leeds ITS is called "Elasticity Model" (Mackie et al., 2003, Table E14a), which has the following utility formulation:

$$V = \underbrace{\beta_C \cdot \Delta C \cdot \left(\frac{Cost}{AvgCost}\right)^{\lambda_{Cost}} \cdot \left(\frac{Inc}{AvgInc}\right)^{\lambda_{Inc}}}_{Cost\ Utility} + \underbrace{\beta_T \cdot \Delta\tau}_{Time\ Utility} + \underbrace{\beta_{Inertia} \cdot Inertia_{Dummy}}_{Inertia\ Term} \tag{4}$$

Where:

$\Delta\tau = sign(\Delta t) \cdot \left\{ |\Delta t| \cdot [|\Delta t| \geq \theta] + \theta \left(\frac{|\Delta t|}{\theta}\right)^m \cdot [|\Delta t| < \theta] \right\}$

$Inertia_{Dummy} = 1 \cdot [\Delta t = 0\ and\ \Delta c = 0]$

$AvgCost = 100p\ and\ AvgInc$=£35K

$\lambda_{cost} = Cost\ Elasticity$

$\lambda_{Inc} = Income\ Elasticity$

Based on this utility function, value of time can be derived using **Equation 5**:

$$VOT = \frac{\partial V/_{\partial \Delta\tau}}{\partial V/_{\partial \Delta C}} = \frac{\beta_\tau}{\beta_C} \cdot \left(\frac{Cost}{AvgCost}\right)^{-\lambda_{Cost}} \cdot \left(\frac{Inc}{AvgInc}\right)^{-\lambda_{Inc}} \tag{5}$$

This utility function was formulated in particular to address the three long standing issues with respect to the VTTS analysis, which are further described in the following sections. These three effects are:

- Sign of the VTTS – WTA (willingness-to-accept) is much higher than WTP
- Size of the VTTS – VTTS is much smaller for small travel time savings
- Interactions between VTTS and journey covariates

3.2.1 Sign effect of VTTS

The nature of the sign and size effects of VTTS and their implications in appraisal evaluation have been examined by researchers for many years (e.g. Welch and Williams, 1997). These effects create practical challenges in conventional appraisal system, as significant portion of the time savings perceived by travellers due to a typical road improvement scheme will need to be excluded if small travel time savings are valued lower than standard VTTS. The advantage of an appraisal system this is reversibility (i.e. gains equal losses) had also been reiterated in Fowkes (2010) and Mackie et al. (2003).

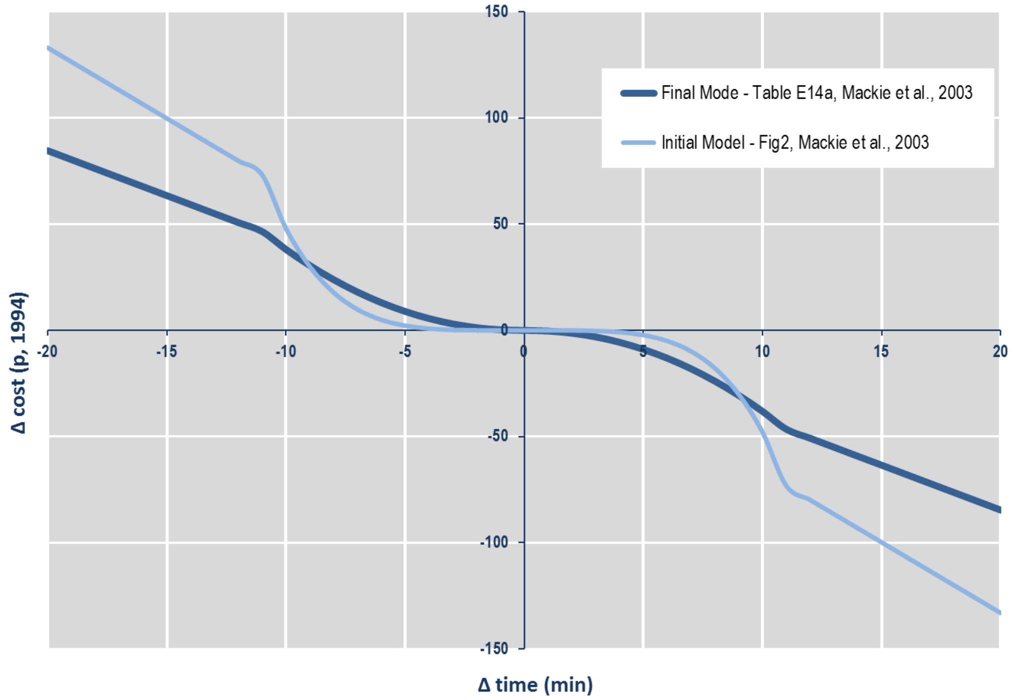
With respect to the sign effect, AHCG reported that losses were valued almost twice as high as gains, and time changes were valued more than small changes. This study attributed the size and sign effects to the short-run nature of the SC experiment and expected these effects will vanish in the longer run hence averages of the coefficients were recommended. AHCG's finding about the sign effect was challenged by Leeds ITS in 2001, in which it was found that ratios between gains and losses for quadrants 2+4 (i.e. tradeoffs that involve comparison with current position) are substantially higher than ratios for quadrants 1+3 (i.e. tradeoffs do not involve comparison with current position). By specifying a dummy "inertia" term for options that referenced to current journey, the ratios between gains and losses are substantially reduced and became fairly symmetrical (Bates and Whelan, 2001). To certain extent, this is an solution that explains the task-simplifying nature of SC respondents, supported by model results which are similar to an earlier Swedish VTTS study (Dillén and Algers, 1999). This finding was later challenged by Gunn and Burge (2001) who stated that the inertia term did not completely remove the asymmetry issue by using the AHCG dataset. Similarly, Van de Kaa (2005) also found that inertia bias itself could not explain the loss aversion effect modelled in the Dutch VTTS study, which produced very similar experimental design and outcomes as in the UK VTTS study.

Recently, researchers in the field of choice modelling have generally accepted the Prospect Theory (Kahneman and Tversky, 1979) which states that the gains–losses asymmetry in the value function are attributed by departures of decision-making from reference points instead of states (Stathopoulos and Hess, 2012). There are recent methodological advances developed by De Borger and Fosgerau (2008) to recover a 'reference-free' value of time as an average of WTP and WTA, based on an essential assumption of symmetric damping rates of time and cost in the value functions. While this new approach provides new direction in explaining loss aversion in VTTS estimation, it is also noted by Daly et al. (2013) that the averaging method to recover the 'reference-free' value of time is somewhat similar to the averaging method for gain and loss coefficients incorporated in the past studies. In retrospect, the inertia term specified in by Leeds ITS clearly is not convincing enough to fully explain the gain-loss behaviour based on recent consensus amongst researchers. This research, however, has retained the model specification so as to have a consistent comparison of VTTS results between basic and advanced choice models.

3.2.2 Size effect of VTTS

Regarding the size of the VTTS, review conducted by Leeds ITS did find similar results as AHCG, that VTTS is increasingly "discounted" below 10min. An "adjusted" indifference curve was then estimated to represent the level of discount for time changes between a threshold θ, and was coined as "perception effect". A grid search was undertaken to estimate the parameter θ and m, which were found to be 11 minutes and 2.09 respectively. The indifference curve for commuting trips that incorporated the perception effect is shown below:

Figure 1 – Indifference curves with perception effect (Mackie et. al, 2003)



It is stressed in Mackie et al. (2003) that the issue with small time savings is due to the SC nature, as respondents are not able to visualize the use of the small time savings in the short-run nature of SC experiment. In longer term however, people will be able to adjust schedule to accommodate the time savings. Daly et al. (2013) recommends researchers to also exploit rich RP dataset available brought by advances in data collection methods. To summarize, there is no particular straightforward solution to deal with the size effect of VTTS to date.

3.2.3 Journey cost covariates

The final recommended model includes continuous interaction between cost and the journey cost covariates income level and journey distance. The original idea was to isolate the “budget effects” on VTTS, based on observations of coefficients in discrete segmentation that higher income households show much higher VTTS. In the actual formulation, current journey cost is used as proxy of journey distance since the SC experiment did not collect distance information. This elasticity model allows deterministic taste heterogeneity, which is not only desirable for interpretation issues but also could also be advantageous in model estimation to random coefficient approaches for large dataset (Hess et al., 2008). It appears that the only drawback is concerning the non-linearity that might cause model estimation issue.

4 MODEL RE-ESTIMATION

4.1 Methodology

The modelling tool used in research is the continuous MMNL model within random coefficient framework. This is not the only advanced choice model as there are a few other options that could be more desirable (e.g. latent class model, see Hess et al. (2011)). However, this continuous RCL model is the most commonly used advanced choice model for understanding taste heterogeneity (Hess, 2012). By implementing RCL model, the VTTS valuation is estimated parametrically in this case. Alternatively, modeller could also estimate VTTS using a non-parametric model, in which no specific distributional assumption on the random error term is assumed by making hypothesize that disturbances follow very general properties (Ben-Akiva and Lerman, 1985). This non-parametric procedure had been applied in VTTS studies in Denmark and Sweden recently (Börjesson et al., 2012, Fosgerau, 2006, Fosgerau, 2007). One of the greatest benefits of applying non-parametric estimation is to ease the drawback of parametric analysis that misleading conclusions might be drawn when there is not enough data supporting the range of distribution (e.g. estimation primarily based on the tail of a distribution). However, non-parametric estimation is relatively cumbersome when the model specification includes several journey covariates simultaneously (Börjesson and Eliasson, 2012). Therefore, non-parametric estimation is not suitable for the VTTS valuation if we follow the model set out in Mackie et al. (2003) that includes both income and distance as journey cost covariates.

4.2 MMNL in preference space

4.2.1 Distributional assumptions

The pitfalls of incorporating inappropriate distributional assumptions of the random coefficients have been discussed at length in Hensher and Greene (2003) and Hess et al. (2005). For VTTS valuation, the occurrence of counter-intuitive sign of the time or cost coefficients (i.e. positive sign) is problematic in particular as it contradicts to most microeconomics theories. That being said, this is less of an issue for this research since the SC questionnaire guaranteed that VTTS is positive by design by enforcing trading-offs of time and cost respondents such that respondents could only choose between time increase or cost decrease (Bates and Whelan, 2001). This implies that time or cost coefficients must be negative, and hence any *a priori* statistical distributions that generate positive time or cost sensitivities must be rejected. Amongst the standard statistical distributions, lognormal and Johnson Sb are the best options to impose strict sign assumptions for the choice analysis. Nevertheless, standard normal, symmetrical triangular and uniform distributions for random coefficients were also tested to examine the sign of coefficients estimated. The distributional assumptions that were applied in Biogeme (BIerlaire's Optimization package for GEV Models Estimation) for this research are briefly described below:

- **Lognormal** – it facilities strict (negative) sign assumption by placing negative sign before exponential but the long tail is not desirable; there might be computational and model convergence issues (Hess et al., 2006); *c*, which is a draw from lognormal distribution can be expressed as:

  For $\xi \sim N(\mu, \sigma^2)$: $$c = e^{\xi} \quad (6)$$

- **Johnson Sb** – it can be bounded and thus overcomes the long tail issue with lognormal distribution while maintaining the tight sign control; it is also the most flexible distribution in terms of the shape, which is controlled by 4 parameters in total (with two additional parameters for range and offset):

  For $\xi \sim N(\mu, \sigma^2)$: $$c = a + (b - a) \cdot \frac{e^{\xi}}{e^{\xi}+1} \quad (7)$$

  Where *c*, a random draw from Sb distribution, is bounded between *a* and *b* (Hess et al., 2006)

- **Normal** – there is problem with the sign interpretation of the coefficients for unbounded distribution
- **Uniform** – specified with lower endpoint *a* and range parameter *b*; sign issue when b is greater or equal to zero
- **Symmetrical triangular** – specified as summation of two uniform distribution; bounded to both side and hence avoid the long tail issue; sign issue when b is greater than zero

Unbounded distributions were not tested due to software limitation at the time of study. In terms of VTTS calculation, Monte Carlo simulation was required for most statistical distributions when both numerator and denominator are randomly distributed except for lognormal distributions. 1 million random Monte Carlo draws were assigned each time for computing the VTTS ratio. Since division between lognormal distributions result in lognormal distribution, thus VTTS can be computed analytically (de Dios Ortuzar and Willumsen, 2011) instead of carrying out Monte Carlo simulation:

$$VOT \sim logN\left(\mathrm{e}^{\left((u_x-u_y)+\frac{\sigma x^2+\sigma y^2}{2}\right)}, e^{\left((u_x-u_y)+\frac{\sigma x^2+\sigma y^2}{2}\right)}\sqrt{e^{\sigma x^2+\sigma y^2)}-1}\right) \quad (8)$$

4.2.2 Model convergence issue

There was another further question concerning model estimation regarding how many draws to be used in the simulation. There is no clear answer and the amount of draws required depends on dataset. It appears that bottom-line seems to be at least 200 Halton draws indicated by Chiou and Walker (2007). Multiple runs with different number of draws were undertaken to check the model convergence (i.e. stable Log-likelihood (LL) and VTTS) in this study.

Results of the model convergence assuming lognormal distribution for both time and cost coefficients are very worrying however as shown in **Table 1.** A range of draws from 100 to 1,250 Halton draws were used. It can be seen that there is a very large range of VTTS estimated from 5.7p/min to 9.25p/min. Final LL estimates are relatively stable with a range between -2,425 and -2,430, except for model run with 750 Halton draws that result in LL of -2,456. Clearly, the model did not converge within 1250 Halton draws. It was also experienced that Bison version of Biogeme could not complete any model runs with more than 1,250 Halton draws. As such, model runs were undertaken using PythonBiogeme, which is a python version of Biogeme that allows parallel computing and more flexible model specification (e.g. likelihood function) (Bierlaire and Fetiarison, 2009).

Table 2 shows that even pushing number of Halton draws from 1000 to 7000 using the PythonBiogeme, the RCL model with log-normally distributed random time and cost coefficients still could not converge as VTTS varies between 5.6p/min to 8.3p/min. There is an interesting "cyclic" pattern emerged: LL's are the highest at 1000, 4000 and 7000 draws that averaged at -2,430 approximately and VTTS estimates revolve around 7p/min to 8p/min. LL's are the lowest at 3000 and 6000 draws with and average LL around -2,455 and VTTS around 5.5p/min. This problem could be related to the complicated non-linear formulation that produces log-likelihood functions which are not globally concaved. Amongst these model estimated using different number of draws, the lognormal RCL with 5000 draws was adopted as it results in fairly good LL (at one of the "LL peaks") as well as a relatively averaged VTTS of 7.13p/min.

Table 1 – Model convergence test for RCL-Lognormal (100 to 1250 draws)

Bison 2.2 (CFSQP)	logn~(time) logn~(cost) 100 draws		logn~(time) logn~(cost) 250 draws		logn~(time) logn~(cost) 500 draws		logn~(time) logn~(cost) 750 draws		logn~(time) logn~(cost) 1000 draws		logn~(time) logn~(cost) 1250 draws	
β-N(Time)-μ	-1.9500	-23.3	-1.8100	-21.5	-1.9800	-18.5	-1.9700	-12.0	-1.9800	-22.1	-1.8500	-21.6
β-N(Time)-σ	0.8880	13.0	-0.6640	-8.6	-0.9090	-3.2	0.4710	1.6	0.9870	10.2	0.6770	8.8
β-N(Cost)-μ	-3.2300	-38.9	-3.1300	-34.4	-3.2100	-35.2	-3.2600	-33.6	-3.2200	-36.6	-3.1100	-35.4
β-N(Cost)-σ	-1.0300	-12.0	0.9280	14.0	0.9040	9.2	-0.8200	-4.4	-0.9980	-11.3	-0.9360	-13.6
WTP-μ (Avg, p/min)	9.0677	-	7.1781	-	7.7815	-	5.681	-	9.2543	-	6.8703	-
WTP-σ (Avg, p/min)	20.986	-	11.744	-	15.896	-	6.8301	-	22.991	-	11.492	-
λ-Cost	-0.3000	-4.8	-0.3340	-5.0	-0.2850	-4.3	-0.3110	-3.3	-0.3050	-4.5	-0.3160	-5.0
λ-Income	-0.4040	-5.2	-0.4090	-4.7	-0.3920	-4.4	-0.4330	-3.9	-0.3860	-4.6	-0.3980	-4.2
m	1.1600	11.1	1.1100	11.5	1.1400	9.1	1.0300	9.6	1.1800	11.3	1.0800	11.4
ϑ	11.0000	-	11.0000	-	11.0000	-	11.0000	-	11.0000	-	11.0000	-
β-Inertia	1.2400	18.5	1.2100	18.5	1.2300	18.1	1.1700	18.0	1.2500	18.4	1.2200	18.4
Parameters	8		8		8		8		8		8	
Observations	4737		4737		4737		4737		4737		4737	
Individuals	695		695		695		695		695		695	
Initial LL	-41477.683		-40162.51		-40810.18		-2965.705		-3053.662		-3051.902	
Final LL	-2426.853		-2430.96		-2431.462		-2456.695		-2425.891		-2430.656	
adj ρ^2	0.258		0.257		0.257		0.249		0.259		0.257	

Table 2 - Model convergence test for RCL-Lognormal (1000 to 7000 draws)

Python 2.2 (CFSQP)	logn~(time) logn~(cost) 1000 draws		logn~(time) logn~(cost) 2000 draws		logn~(time) logn~(cost) 3000 draws		logn~(time) logn~(cost) 4000 draws		logn~(time) logn~(cost) 5000 draws		logn~(time) logn~(cost) 6000 draws		logn~(time) logn~(cost) 7000 draws	
β-N(Time)-μ	-1.9000	*-21.5*	-1.8800	*-21.1*	-1.9900	*-14.2*	-1.9200	*-21.6*	-1.9200	*-21.5*	-1.9900	*-14.5*	-1.8800	*-20.7*
β-N(Time)-σ	-0.8200	*-8.8*	0.8590	*12.5*	0.5150	*2.5*	0.7460	*8.5*	0.7550	*7.6*	0.5120	*2.5*	0.8370	*9.8*
β-N(Cost)-μ	-3.1700	*-36.9*	-3.1900	*-37.8*	-3.2600	*-34.3*	-3.1600	*-37.1*	-3.1700	*-36.1*	-3.2600	*-34.3*	-3.1800	*-37.9*
β-N(Cost)-σ	0.9400	*13.1*	-0.9400	*-14.3*	-0.7920	*-5.5*	-0.9350	*-12.4*	-0.9270	*-12.2*	-0.7920	*-5.6*	-0.9390	*-13.9*
WTP-μ (Avg, p/min)	7.7524	-	8.3373	-	5.5636	-	7.0665	-	7.1326	-	5.5550	-	8.0942	-
WTP-σ (Avg, p/min)	14.9920	-	16.8004	-	6.6791	-	12.6050	-	12.7113	-	6.6514	-	15.9151	-
λ-Cost	-0.3100	*-4.8*	-0.3440	*-5.7*	-0.3150	*-3.6*	-0.2890	*-4.5*	-0.2870	*-4.2*	-0.3170	*-3.6*	-0.3400	*-5.6*
λ-Income	-0.4040	*-4.7*	-0.4050	*-5.1*	-0.4230	*-4.2*	-0.3940	*-4.5*	-0.3960	*-4.2*	-0.4220	*-4.1*	-0.4010	*-5.0*
m	1.1400	*11.4*	1.1900	*11.6*	1.0300	*9.9*	1.0900	*11.2*	1.0900	*11.1*	1.0300	*10.0*	1.1700	*11.5*
ϑ	11.0000	-	11.0000	-	11.0000	-	11.0000	-	11.0000	-	11.0000	-	11.0000	-
β-Inertia	1.2300	*18.4*	1.2300	*18.4*	1.1700	*18.4*	1.2200	*18.5*	1.2200	*18.5*	1.1700	*18.5*	1.2300	*18.4*
Parameters	8		8		8		8		8		8		8	
Observations	4737		4737		4737		4737		4737		4737		4737	
Individuals	N/A		N/A		N/A		N/A		N/A		N/A		N/A	
Initial LL	-2430.375		-2432.126		-2475.563		-2432.264		-2432.909		-2475.703		-2431.847	
Final LL	-2428.466		-2427.803		-2456.084		-2431.666		-2432.22		-2455.904		-2428.597	
adj ρ^2	0.258		0.258		0.250		0.257		0.257		0.250		0.258	

4.2.3 Model results

Table 3 summarizes all the model results for MMNL models in preference space assuming different mixing distributions. All parameters generated are significant. Comparing to the base MNL model, LL's have been significantly improved by 250 units approximately in applying any of the five standard distributions for random coefficients. Similarly, $\bar{\rho}^2$ values are also improved substantially from 0.179 in MNL to over 0.25 approximately using MMNL. This finding of significant LL improvement with MMNL model aligns with past literature (Hensher and Greene, 2003, Hess et al., 2005) and thus again it highlights the importance of explaining taste heterogeneity. Both lognormal and Johnson Sb RCL models produce much higher VTTS compared to MNL model, with VTTS 69% and 81% higher than 4.22p/min estimated by original MNL model. It indicates that the MNL understates VTTS significantly by ignoring variation in the time and cost sensitivities. Both standard normal, symmetrical triangular distributed RCL model display positive cost sensitivities and thus the resulting VTTS estimates that have infinite moments are not reported. Although cost sensitivity is negative for uniformly distributed RCL model, time sensitivity has 20% chance of being positive and thus the corresponding VTTS is rejected also.

Comparing the RCL models with lognormal and Johnson Sb distributions, Johnson Sb distributions provide a clear advantage in providing better model fit in both LL and $\bar{\rho}^2$. This model goodness-of-fit improvement came at very high computing cost though. The model estimation for 100 Halton draws using Johnson Sb distributions took 4.5 hours approximately while it took simply 3 hours to complete a lognormal RCL model with 1250 Halton draws. Also, attempts were made to add more draws for RCL model with Johnson Sb distributions; however, model did not seem to converge at all, indicating significant computational challenges in implementing Johnson Sb distributions for complex non-linear formulation.

An additional step to test whether the lognormal distribution is a "true" distribution through a semi-nonparametric (SNP) estimation was also carried out. This SNP approach, developed by Fosgerau and Bierlaire (2007), first defined a generalized mixing distribution to transform the *a prior* distributional assumption (i.e. lognormal) using a Legendre series polynomial. Where more Legendre polynomials are added, the approximation is supposed to increase. Likelihood ratio tests were then carried out to test the base distribution against the generalized distribution in this case. Results from SNP show that both the time and cost coefficients have shown to be the "true" mixing distributions in this case.

Table 3 – MMNL in preference space

	base MNL (with panel effect)		logn~(time) logn~(cost) 5000 draws		Sb~(time) Sb~(cost) 100 draws		Norm~(time) Norm~(cost) 500 draws		unif~(time) unif~(cost) 500 draws		triang~(time) triang~(cost) 500 draws	
β-N(Time)-μ	-0.1030	-12.8	-1.9200	-21.5	4.4400	9.7	-0.1530	-11.9	-0.1540	-12.0	-0.1140	-9.0
β-N(Time)-σ			0.7550	7.6	-1.6200	-5.0	0.1620	11.5	0.2570	13.0	0.1530	8.7
β-N(Cost)-μ	-0.0244	-11.7	-3.1700	-36.1	4.2600	6.5	-0.0486	-14.1	-0.0487	-14.2	-0.0265	-12.8
β-N(Cost)-σ			-0.9270	-12.2	4.6800	4.2	-0.0264	-10.8	-0.0438	-12.3	0.0323	9.0
Sb-Time-λ					6.2300	-3.4						
Sb-Cost-λ					0.3340	-5.3						
Sb-Time-ξ					-6.2900	3.4						
Sb-Cost-ξ					-0.3570	5.7						
WTP-μ (Avg, p/min)	4.2213	9.6	7.1326	-	7.6442	-	-	-	5.1705	-	-	-
WTP-σ (Avg, p/min)			12.7113	-	13.8975	-	-	-	8.6136	-	-	-
λ-Cost	-0.4090	-6.3	-0.2870	-4.2	-0.3700	-5.6	-0.3290	-5.5	-0.3310	-6.1	-0.4050	-7.3
λ-Income	-0.3660	-5.4	-0.3960	-4.2	-0.4180	-4.6	-0.2540	-3.3	-0.3260	-4.4	-0.3840	-5.0
m	2.0900	5.3	1.0900	11.1	1.2300	10.1	0.9940	9.6	0.9540	9.5	1.1700	9.7
ϑ	11.0000	-	11.0000	-	11.0000	-	11.0000	-	11.0000	-	11.0000	-
β-Inertia	0.9000	19.9	1.2200	18.5	1.2500	18.0	1.2500	18.6	1.2400	18.5	1.1800	17.2
Parameters	6		8		12		8		8		8	
Observations	4737		4737		4737		4737		4737		4737	
Individuals	695		N/A		695		695		695		695	
Initial LL	-3283.438		-2432.909		-3283.438		-3256.649		-3171.325		-2538.695	
Final LL	-2690.983		-2432.22		-2412.007		-2443.635		-2447.045		-2469.678	
adj ρ^2	0.179		0.257		0.262		0.253		0.252		0.245	
% positive - time	0%		0%		0%		14%		20%		3%	
% positive - cost	0%		0%		0%		11%		0%		2%	
									** σ - half spread*		** σ - half spread*	
									** U~ (time,cost)*		** U~ (time,cost)*	

Modelling results also indicate that parameter *m*, which controls the curvature of the VTTS "discount" for small travel time savings, have dropped significantly from 2.1 in MNL model to 1.1 and 1.2 in MMNL models with lognormal and Johnson Sb distributions for random coefficients respectively. By definition the *m* parameter becomes 1 when there is no VTTS "discount" at all. **Figure 2** and **Table 4** demonstrate the significant reduction of VTTS "discount" for travel time savings smaller than 11min. For example, VTTS is estimated to be discounted by 58% at 5min time savings using basic MNL model but this VTTS "discount" almost got completely wiped out (7% discount only) as estimated by MMNL model. It implies that that the size effect of VTTS is almost negligible in this setup, or is much smaller than what modellers used to think in the past using this UK SC dataset. Further research is required to investigate such reduced size effect.

Figure 2 – Indifference curves with perception effect – MMNL (Lognormal, preference space)

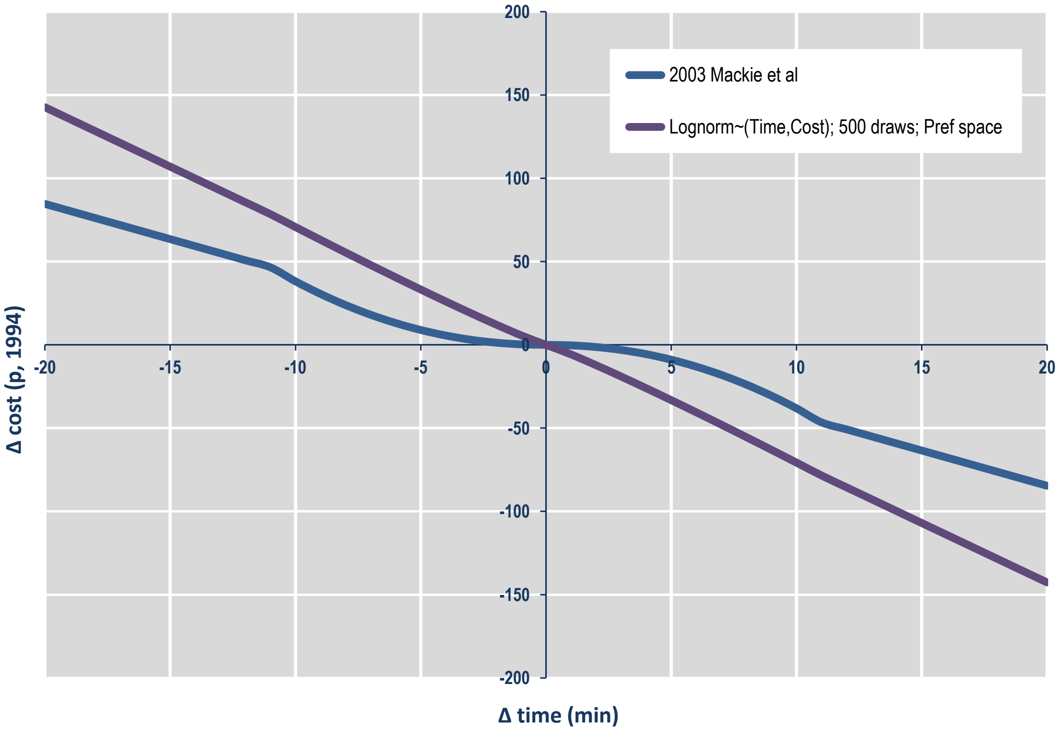


Table 4 – VTTS "discount" for small travel time savings

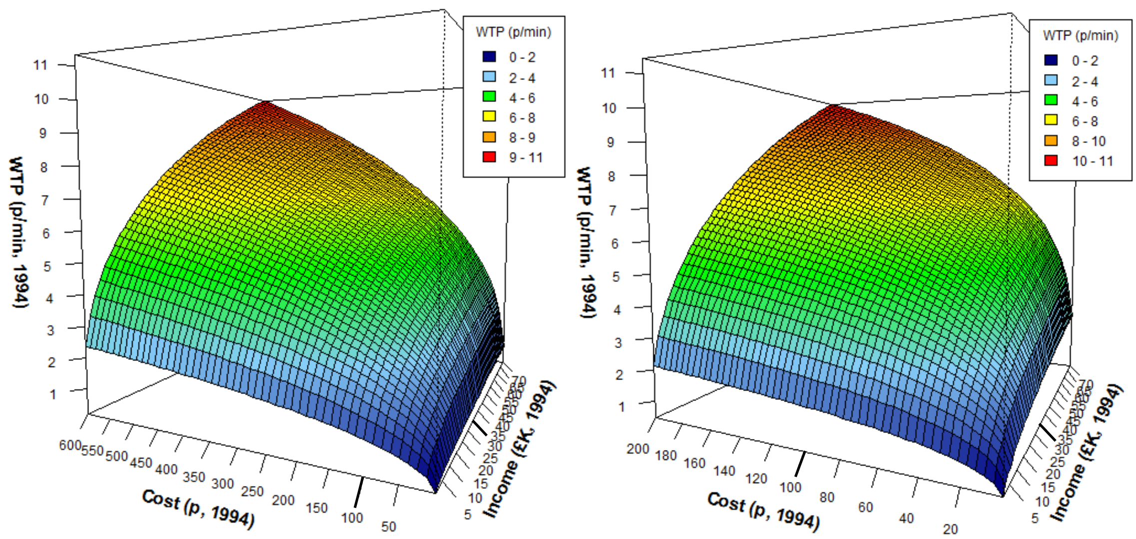
Time Savings (min)	MNL Model	MMNL Model Lognormal 5000 Draws
10	-10%	-1%
9	-20%	-2%
8	-29%	-3%
7	-39%	-4%
6	-48%	-5%
5	-58%	-7%
4	-67%	-9%
3	-76%	-11%
2	-85%	-14%
1	-93%	-19%

5 CONCLUSIONS

This study shows strong empirical evidence of taste variation amongst respondents, as exemplified by the improved LL and goodness-of-fit using RCL models. Rigorous tests have been undertaken to examine a range of possible high and low estimates of VTTS. This work recommends the RCL model with lognormal distribution assumed for time and cost coefficients in preference space, which produces good model fit yet providing more superior computing advantage over RCL model with Sb distributional assumptions. Hence, the recommended VTTS is 7.13p/min, which is 69% higher than the current base VTTS of 4.22p/min. A graphical representation of the VTTS changes in relation to journey cost and income is presented below.

Figure 3 – VTTS distribution (Lognormal-RCL in preference space)

MNL Model | Lognormal-RCL Model



Since this new un-weighted VTTS for non-working trips presented in this paper is substantially different to the value estimated back in the 2001 VTTS study, it is certainly more desirable to update the current appraisal VTTS using a more advanced choice model. Furthermore, with the VTTS estimated by RCL model now becomes a distribution itself, new modelling approach could therefore better utilize this distributional VTTS through micro-simulation approach within a activity-based modelling framework or enforcing very refined VTTS segmentation that closely corresponds to the VTTS distributional profile so as to generate more realistic travel behavioural modelling, in particular for any price-sensitive transport schemes.

It is also shown that RCL model reduces the *m* parameter of the perception effect to value around one and hence significantly reduces the VTTS discount for small travel time savings. This particular finding challenges the need to discount or even completely ignored the small travel time savings for appraisal (e.g. as in Germany). Lastly, the significant difference of VTTS outcomes resulted by varying model specifications, distributional assumptions and number of draws demonstrated in this study also encourages decision-makers to set out a more stringent guideline to ensure advanced choice modelling approaches adopted for appraisal evaluation are specified appropriately and applied consistently amongst choice modellers for fair assessment of travel benefits.

ACKNOWLEDGEMENTS

This paper is largely based on research conducted as part of my dissertation submitted for the MSc Transport Planning at the University of Leeds. I would like to take this opportunity to sincerely thank my supervisor Professor Stephane Hess for his support and guidance throughout this research. Any mistakes in the present paper remain the responsibility of the author.

REFERENCES

AHCG, A. M. A. R., HAGUE CONSULTING GROUP, DEPARTMENT OF THE ENVIRONMENT, TRANSPORT AND REGIONS 1999. *The Value of Time on UK Roads,* Netherlands, The Hague.

ALGERS, S., BERGSTRÖM, P., DAHLBERG, M. & LINDQVIST DILLÉN, J. 1998. Mixed logit estimation of the value of travel time, Working Paper.

BATES, J. & WHELAN, G. 2001. Size and sign of time savings. *Working Paper 561.* Institute of Transport Studies, Univeristy of Leeds.

BEESLEY, M. E. 1965. The value of time spent in travelling: some new evidence. *Economica,* 32**,** 174-185.

BEN-AKIVA, M. & BOLDUC, D. 1996. *Multinomial probit with a logit kernel and a general parametric specification of the covariance structure*, D'epartement d''economique, Universit'e laval with Department of Civil and Environmental Engineering, Massachusetts Institute of Technology.

BEN-AKIVA, M. E. & LERMAN, S. R. 1985. *Discrete choice analysis: theory and application to predict travel demand*, The MIT press.

BIERLAIRE, M. & FETIARISON, M. Estimation of discrete choice models: extending biogeme. Proceedings of the 9th Swiss Transport Research Conference. Ascona, Switzerland, 2009.

BÖRJESSON, M. & ELIASSON, J. 2012. Experiences from the Swedish Value of Time study. CTS-Centre for Transport Studies Stockholm (KTH and VTI).

BÖRJESSON, M., FOSGERAU, M. & ALGERS, S. 2012. Catching the tail: Empirical identification of the distribution of the value of travel time. *Transportation Research Part A: Policy and Practice,* 46**,** 378-391.

CHIOU, L. & WALKER, J. L. 2007. Masking identification of discrete choice models under simulation methods. *Journal of Econometrics,* 141**,** 683-703.

DALY, A., HESS, S. & TRAIN, K. 2012. Assuring finite moments for willingness to pay in random coefficient models. *Transportation,* 39**,** 19-31.

DALY, A., TSANG, F. & ROHR, C. 2013. The Value of Small Time Savings for Non-business Travel *Journal of Transport Economics and Policy (JTEP),* to be published.

DE BORGER, B. & FOSGERAU, M. 2008. The trade-off between money and travel time: A test of the theory of reference-dependent preferences. *Journal of Urban Economics,* 64**,** 101-115.

DE DIOS ORTUZAR, J. & WILLUMSEN, L. G. 2011. *Modelling Transport*, Wiley.

DILLÉN, J. L. & ALGERS, S. Further research on the national Swedish value of time study. World Transport Research: Selected Proceedings of the 8th World Conference on Transport Research, 1999.

FOSGERAU, M. 2006. Investigating the distribution of the value of travel time savings. *Transportation Research Part B: Methodological,* 40**,** 688-707.

FOSGERAU, M. 2007. Using nonparametrics to specify a model to measure the value of travel time. *Transportation Research Part A: Policy and Practice,* 41**,** 842-856.

FOSGERAU, M. & BIERLAIRE, M. 2007. A practical test for the choice of mixing distribution in discrete choice models. *Transportation Research Part B: Methodological,* 41**,** 784-794.

FOWKES, A. 2010. The value of travel time savings. *In:* DE BOECK , A. (ed.) *Applied Transport Economics: A Management and Policy Perspective.*

GUNN, H. & BURGE, P. The value of travel time savings: some new evidence. PROCEEDINGS OF THE AET EUROPEAN TRANSPORT CONFERENCE, HELD 10-12 SEPTEMBER, 2001, HOMERTON COLLEGE, CAMBRIDGE, UK-CD-ROM, 2001.

HENSHER, D. A. & GREENE, W. H. 2003. The mixed logit model: the state of practice. *Transportation,* 30**,** 133-176.

HESS, S. 2005. *Advanced discrete choice models with applications to transport demand.* University of London.

HESS, S. 2012. Rethinking heterogeneity: the role of attitudes, decision rules and information processing strategies. *Transportation Letters: the International Journal of Transportation Research,* 4**,** 105-113.

HESS, S., BEN-AKIVA, M., GOPINATH, D. & WALKER, J. 2011. Advantages of latent class over continuous mixture of logit models. *unpublished:* http://www. *stephanehess. me. uk/papers/Hess_Ben-Akiva_Gopinath_Walker_May_2011. pdf [accessed August 2013]*.

HESS, S., BIERLAIRE, M. & POLAK, J. W. 2005. Estimation of value of travel-time savings using mixed logit models. *Transportation Research Part A: Policy and Practice,* 39**,** 221-236.

HESS, S., ERATH, A. & AXHAUSEN, K. W. 2008. Estimated value of savings in travel time in Switzerland: analysis of pooled data. *Transportation Research Record: Journal of the Transportation Research Board,* 2082**,** 43-55.

HESS, S., POLAK, J. W. & AXHAUSEN, K. W. Distributional Assumptions in Mixed Logit Models. Transportation Research Board 85th Annual Meeting, 2006.

KAHNEMAN, D. & TVERSKY, A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society***,** 263-291.

MACKIE, P., WARDMAN, M., FOWKES, A., WHELAN, G., NELLTHORP, J. & BATES, J. 2003. Values of travel time savings UK.

MCFADDEN, D. 1974. Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics***,** 105-142.

MCFADDEN, D. & TRAIN, K. 2000. Mixed MNL models for discrete response. *Journal of applied Econometrics,* 15**,** 447-470.

MVA, ITS, L. & TSU, O. 1987. The Value of Travel Time Savings. *Policy Journals*.

NELLTHORP, J. 2001. Valuation Conventions for UNITE.

STATHOPOULOS, A. & HESS, S. 2012. Revisiting reference point formation, gains–losses asymmetry and non-linear sensitivities with an emphasis on attribute specific treatment. *Transportation Research Part A: Policy and Practice,* 46**,** 1673-1689.

TRAIN, K. 2009. *Discrete choice methods with simulation*, Cambridge university press.

REVELT, D. & TRAIN, K. 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics,* 80**,** 647-657.

REVELT, D. & TRAIN, K. 2000. Customer-specific taste parameters and Mixed Logit: Households' choice of electricity supplier.

VAN DE KAA, E. J. Heuristic Judgement, Prospect Theory and Stated Preference Surveys Aimed to Elicit the Value of Travel Time. European Transport Conference, 2005, 2005.

WALKER, J. L. & BEN-AKIVA, M. 2011. Advances in discrete choice: mixture models. *In:* DE PALMA, A., LINDSEY, R., QUINET, E. & VICKERMAN, R. (eds.) *A Handbook of Transport Economics.* Edward Elgar Publishing, Inc.

WELCH, M. & WILLIAMS, H. 1997. The sensitivity of transport investment benefits to the evaluation of small travel-time savings. *Journal of transport economics and policy***,** 231-254.