I T L S

WORKING PAPER

ITLS-WP-10-15

Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models

By

William H Greene[1] and David A Hensher

[1]Department of Economics
Stern School of Business
New York University

August 2010

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in
Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-10-15

TITLE: **Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models**

ABSTRACT: There is growing interest in establishing a mechanism to account for scale heterogeneity across individuals (essentially the variance of a variance term or the standard deviation of utility over different choice situations), in addition to the more commonly identified taste heterogeneity in mixed logit models. A number of authors have recently proposed a model that recognizes the relationship between scale and taste heterogeneity, and investigated the behavioural implications of accounting for scale heterogeneity in contrast to a term in the utility function, itself. In this paper we present a general model that extends the mixed logit model to explicitly account for scale heterogeneity in the presence of preference heterogeneity, and compare it with models that assume only scale heterogeneity (referred to as the scale heterogeneous multinomial logit model) and only preference heterogeneity. Our empirical assessment suggests that accommodating scale heterogeneity in the absence of accounting for preference heterogeneity may be of limited empirical interest, resulting in a statistically inferior model, despite it being an improvement over the standard MNL model. Scale heterogeneity in the presence of preference heterogeneity does garner favour, with the generalized mixed logit model an improvement over the standard mixed logit model. The evidence herein suggests, however, that compared to a failure to account for preference heterogeneity that is consequential, failure to account for scale heterogeneity may not be of such great empirical consequence in respect of behavioural outputs such as direct elasticities and willingness to pay. However additional studies are required to establish the extent to which this evidence is transferable to a body of studies.

KEY WORDS: ***Scale heterogeneity, taste heterogeneity, generalized mixed logit, commuting mode choice, stated choice experiment, elasticities***

AUTHORS: William H Greene and David A Hensher

CONTACT:
Institute of Transport and Logistics Studies (C37)
The Australian Key Centre in Transport Management
The University of Sydney NSW 2006 Australia

Telephone: +61 9351 0071
Facsimile: +61 9351 0088
E-mail: itls@sydney.edu.au
Internet: http://www.sydney.edu.au/business/itls

DATE: August 2010

Acknowledgements The comments of Riccardo Scarpa, John Rose and Chandra Bhat and three referees on earlier versions are appreciated.

1. Introduction

Choice analysis has often been described as a way of explaining variations in the behaviour of a sample of individuals. The consequence of this view is that a key focus of model development has been the search for increasing sources of variance, or heterogeneity, in the candidate observed and unobserved influences on choice making.

Recent emphasis has been given to the treatment of scale, in particular recognition of variance in utility over different choice situations. This is referred to as scale heterogeneity. Scale heterogeneity is a relatively old problem (see Louviere 1999, 2002, Hensher et al. 1999 for the historical context), but it is only in recent years that we have seen a concerted effort to develop estimation capability within the family of logit models to account for it at the respondent level. Fiebig et al. (2009) is the most recent paper, formalizing the campaign led by Louviere and colleagues (1999, 2002, 2006, 2008) to recognize this claimed important source of variability that has been neglected by a focus on revealing preference heterogeneity (now aligned with the mixed logit model). Papers by Breffle and Morey (2000) and Hess et al. (2009) are other contributions.

It is early days to be definitive in the empirical implications of the role of scale and the extent which preference and scale heterogeneity are independent or proportional. What is clear however is that the specification of a model that allows for both sources of heterogeneity induces correlation amongst the observed attributes, and this should be accounted for (see Train and Weeks 2005).

In investigating the potential role of scale heterogeneity, we need to estimate a number of models that accommodate mixtures of preference and scale heterogeneity. The set of interest include, in addition to the basic multinomial logit model, the standard mixed logit model (with random parameters), the multinomial logit model extended to allow for scale heterogeneity, and a generalized mixed logit model in which both random parameters (to account for preference heterogeneity) and variation on the variance condition associated with the random component (known as scale heterogeneity) are included.

We utilize an existing stated choice data set for the choice amongst existing and potentially new modal alternatives for the commuter trip in Sydney in 2003, to estimate the four model forms, contrasting them on their overall goodness-of-fit and associated mode-specific direct elasticities for travel time and cost.

The paper is organised as follows. The next section sets out the generalised logit model that explicitly defines the alternative specifications for identification of heterogeneity of tastes and scale. The empirical data are briefly overviewed, followed by model estimation and findings.

2. Accounting for scale and taste heterogeneity

The generalized mixed logit model employed here builds on the specifications of the mixed logit developed in Train (2003), Hensher and Greene (2003) and Greene (2007), amongst others, and the "generalized multinomial logit model" proposed in Fiebig et al. (2009). The mixed multinomial logit model is

$$\text{Prob}(choice_{it} = j \mid \mathbf{x}_{it,j}, \mathbf{z}_i, \mathbf{v}_i) = \frac{\exp(V_{it,j})}{\sum_{j=1}^{J_{it}} \exp(V_{it,j})} \tag{1}$$

where

$V_{it,j}$ = $\beta_i' x_{it,j}$

β_i = $\beta + \Delta z_i + \Gamma v_i$

$x_{it,j}$ = the K attributes of alternative j in choice situation t faced by individual i,

z_i = a set of *M* characteristics of individual *i* that influence the mean of the taste parameters; and

v_i = a vector of *K* random variables with zero means and known (usually unit) variances and zero covariances.

The multinomial choice model thus far embodies both observed and unobserved heterogeneity in the preference parameters of individual *i*. Observed heterogeneity is reflected in the term $\mathbf{\Delta z}_i$ while the unobserved heterogeneity is embodied in $\mathbf{\Gamma v}_i$. Structural parameters to be estimated are the constant vector, $\boldsymbol{\beta}$, the $K \times M$ matrix of parameters $\mathbf{\Delta}$ and the nonzero elements of the lower triangular Cholesky matrix, $\mathbf{\Gamma}$.

A number of interesting special cases are straightforward modifications of the model. Specific nonrandom parameters are specified by rows of zeros in $\mathbf{\Gamma}$. A pure random parameters MNL model results if $\mathbf{\Delta} = \mathbf{0}$ and $\mathbf{\Gamma}$ is diagonal. The basic multinomial logit model results[1] if $\mathbf{\Delta} = \mathbf{0}$ and $\mathbf{\Gamma} = \mathbf{0}$.

A growing number of authors have stated that the mixed logit model, and multinomial choice models more generally, do not adequately account for scale heterogeneity (e.g., Feibig et al. 2009 and Keane 2006). Scale heterogeneity across choices is easily accommodated in the model already considered by random alternative-specific constants. As in the earlier implementation, we accommodate both observed and unobserved heterogeneity in the model. The preceding is modified accordingly as equation (2).

$$\boldsymbol{\beta}_i = \sigma_i[\boldsymbol{\beta} + \mathbf{\Delta z}_i] + [\gamma + \sigma_i(1-\gamma)]\,\mathbf{\Gamma v}_i \qquad (2)$$

where

σ_i = $\exp[\bar{\sigma} + \boldsymbol{\delta}'\mathbf{h}_i + \tau w_i]$, the individual specific standard deviation of the idiosyncratic error term

$\mathbf{h}_i$ = a set of *L* characteristics of individual *i* that may overlap with $\mathbf{z}_i$,

$\boldsymbol{\delta}$ = parameters in the observed heterogeneity in the scale term

w_i = the unobserved heterogeneity, standard normally distributed

$\bar{\sigma}$ = a mean parameter in the variance

τ = the coefficient on the unobserved scale heterogeneity

γ = a weighting parameter that indicates how variance in residual preference heterogeneity varies with scale, with $0 \leq \gamma \leq 1$;

and

The weighting parameter, γ, is central to the generalized model. It controls the relative importance of the overall scaling of the utility function, σ_i, versus the scaling of the individual preference weights contained in the diagonal elements of $\mathbf{\Gamma}$. Note that if σ_i equals one, (i.e., $\tau = 0$), then γ falls out of the model and (2) reverts back to the base case random parameters model. A nonzero γ cannot be estimated apart from $\mathbf{\Gamma}$ when σ_i equals one. When σ_i is not equal to one, then γ will spread the influence of the random components between overall scaling and the scaling of the preference weights. In addition to the useful special cases of the original mixed model, some useful special cases arise in this model. If $\gamma = 0$, then a scaled mixed logit model emerges, given in (3).

[1] One can however allow for deterministic taste heterogeneity via interaction terms with respondent-specific characteristics.

$$\boldsymbol{\beta}_i = \sigma_i[\boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i + \boldsymbol{\Gamma}\mathbf{v}_i] \quad (3)$$

If, further, $\boldsymbol{\Gamma} = \mathbf{0}$ and $\boldsymbol{\Delta} = \mathbf{0}$, a "scaled multinomial logit model" model is implied;

$$\boldsymbol{\beta}_i = \sigma_i\boldsymbol{\beta}. \quad (4)$$

This generalized mixed model also provides a straightforward method of reparameterizing the model to estimate the taste parameters in willingness to pay (WTP) space, which has recently become a behaviourally appealing alternative way of directly obtaining an estimate of WTP (See Train and Weeks 2005, Fosgerau 2007, 2007, Scarpa et al. 2008, 2008a, Sonnier et al. 2007, and Hensher and Greene 2009). If $\gamma = 0$, $\boldsymbol{\Delta} = \mathbf{0}$ and the element of $\boldsymbol{\beta}$ corresponding to the price or cost variable is normalized to 1.0 while a nonzero constant is moved outside the brackets, the following reparameterized model emerges:

$$\boldsymbol{\beta}_i = \sigma_i\beta_c\begin{bmatrix} 1 \\ \left(\frac{1}{\beta_c}\right)(\boldsymbol{\beta}+\boldsymbol{\Gamma}\mathbf{v}_i) \end{bmatrix} = \sigma_i\beta_c\begin{bmatrix} 1 \\ \boldsymbol{\theta}_c+\boldsymbol{\Gamma}_c\mathbf{v}_i \end{bmatrix}. \quad (5)$$

In the simple multinomial logit case ($\sigma_i = 1$, $\boldsymbol{\Gamma} = \mathbf{0}$), this is a one to one transformation of the parameters of the original model. Where the parameters are random, however, the transformation is no longer that simple. We, as well as Train and Week (2005), have found, in application, that this form of the transformed model produces generally much more reasonable estimates of willingness to pay for individuals in the sample than the model in the original form in which WTP is computed using ratios of parameters (Hensher and Greene 2009)[2].

The full model, in the unrestricted form or in any of the modifications, is estimated by maximum simulated likelihood (see Greene 2007). Fiebig et al. (2009) note two minor complications in estimation. First, the parameter $\overline{\sigma}$ in σ_i is not separately identified from the other parameters of the model. We will assume that the variance heterogeneity is normally distributed. Neglecting the observed heterogeneity (i.e., $\boldsymbol{\delta}'\mathbf{h}_i$) for the moment, it will follow from the general result for the expected value of a lognormal variable that $E[\sigma_i] = \exp(\overline{\sigma} + \tau^2/2)$. That is, $\sigma_i = \exp(\overline{\sigma})\exp(\tau w_i)$ where $w_i \sim N(0,1)$, so $E[\sigma_i] = \exp(\overline{\sigma})E[\exp(\tau w_i)] = \exp(\overline{\sigma})\exp(E[\tau w_i] + ½\,\mathrm{Var}[\tau w_i]) = \exp(\overline{\sigma} + \tau^2/2)$. It follows that $\overline{\sigma}$ is not identified separately from τ, which appears nowhere else in the model. Some normalization is required. A natural normalization would be to set $\overline{\sigma} = 0$. However, it is more convenient to normalize σ_i so that $E[\sigma_i^2] = 1$, by setting $\overline{\sigma} = -\tau^2/2$ instead of zero.

A second complication concerns the variation in σ_i during the simulations. The lognormal distribution implied by $\exp(-\tau^2/2 + \tau w_i)$ can produce extremely large draws and lead to overflows and instability of the estimator. To accommodate this concern, we have truncated the standard normal distribution of w_i at -1.96 and +1.96. In contrast to Fiebig et al. who propose an acceptance/rejection method for the random draws, we have used a one draw method, $w_{ir} = \Phi^{-1}[.025 + .95U_{ir}]$ where $\Phi^{-1}(t)$ is the inverse of the standard normal *cdf* and U_{ir} is a random draw from the standard uniform population. This will maintain the smoothness of the estimator in the random draws. The acceptance/rejection approach requires, on average, 1/.95 draws to obtain an acceptable draw, while the inverse probability approach always requires exactly one.

Finally, in order to impose the limits on γ (equation 2), γ is reparameterized in terms of α, where $\gamma = \exp(\alpha)/[1 + \exp(\alpha)]$ and α is unrestricted. Likewise, to ensure $\tau \geq 0$, the model is fit in terms of λ, where $\tau = \exp(\lambda)$ and λ is unrestricted. Restricted versions in which it is desired to restrict $\gamma = 1$ or 0

[2] The paper by Hensher and Greene (2009), like Train and Weeks (2005) supports the WTP space framework for estimating WTP distributions given that the evidence on the range is behaviourally more plausible, despite the overall goodness-of-fit being inferior to the utility space specifications.

and/or $\tau = 0$ are imposed directly during the estimation, rather than using extreme values of the underlying parameters, as in previous studies. Thus, in estimation, the restriction $\gamma = 0$ is imposed directly, rather than using, for example, $\alpha = -10.0$ or some other large value.

The fully specified model is given in equation (6). Combining all terms, the simulated log likelihood function for the sample of data is shown in equation (6).

$$\log L = \sum_{i=1}^{N} \log\left\{\frac{1}{R}\sum_{r=1}^{R} \prod_{t=1}^{T_i} \prod_{j=1}^{J_{it}} P(j, \mathbf{X}_{it}, \boldsymbol{\beta}_{ir})^{d_{it,j}}\right\} \tag{6}$$

where

$\boldsymbol{\beta}_{ir} = \sigma_{ir}[\boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i] + [\gamma + \sigma_{ir}(1-\gamma)]\,\boldsymbol{\Gamma}\mathbf{v}_{ir}$,

$\sigma_{ir} = \exp[-\tau^2/2 + \boldsymbol{\delta}'\mathbf{h}_i + \tau w_{ir}]$,

v_{ir} and w_{ir} = the R simulated draws on v_i and w_i,

d_{itj} = 1 if individual i makes choice j in choice situation t and 0 otherwise,

and

$$P(j, \mathbf{X}_{it}, \boldsymbol{\beta}_{ir}) = \frac{\exp(\mathbf{x}'_{it,j}\boldsymbol{\beta}_{ir})}{\Sigma_{j=1}^{J_{it}} \exp(\mathbf{x}'_{it,j}\boldsymbol{\beta}_{ir})} \tag{7}$$

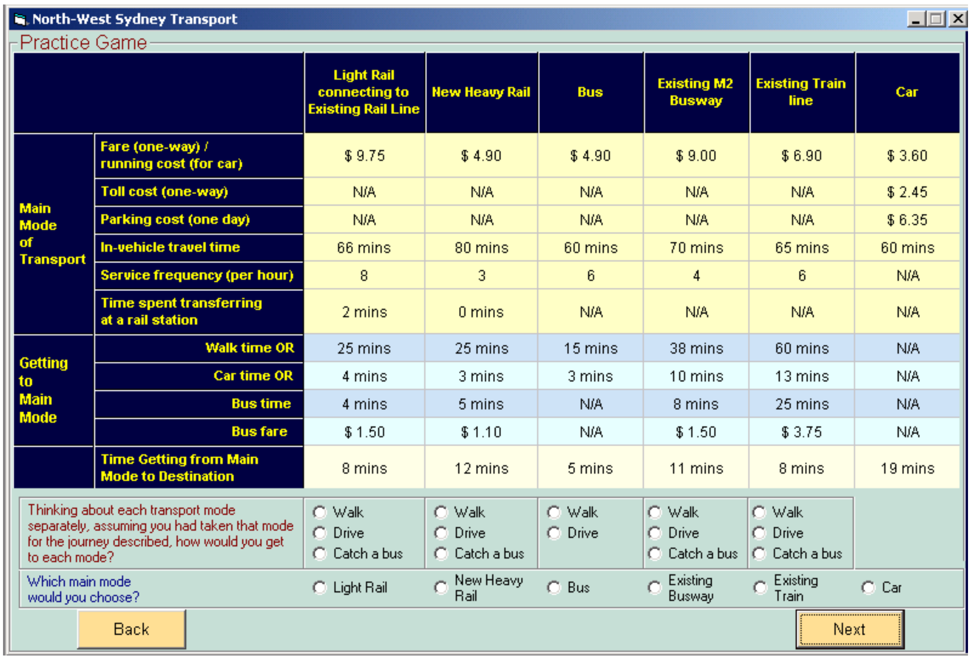
3. Empirical application

To illustrate the empirical implications of the four model forms, we use a stated preference data set on commuting mode choice collected in July 2003 in Sydney. The primary objective of the survey was to establish the preferences of a sample of residents of the study area for alternative modes of transport for commuting and to use this information in forecasting patronage levels for a number of 'new' transport modes – especially the extension of the heavy rail system versus light rail or a busway along the same corridor.

Details of the survey are given in Hensher and Rose (2007). Using the computer aided personal interview (CAPI) technology, sampled residents were invited to review a number of alternative main and access modes, in terms of levels of service and costs within the context of a recent trip, and to indicate which main mode and access mode would be the most preferred. This process of review was repeated 10 times under alternative scenarios of attribute levels, each time requiring the individual to indicate the preferred main and access mode. The choice sets comprised all existing available main modes (i.e., subsets of bus, heavy rail, car, busway) and access modes (i.e., subsets of walk, bus, car), plus two of the new modal options from the full set of three evaluated across the entire sample (i.e., new heavy rail, new light rail, new busway).[3] Respondents were sampled to cover travel both within and outside of the region.

An example of a stated choice screen is shown as Figure 1, derived as one row of a D-optimal stated choice design. See Rose and Bliemer (2008) and Rose et al. (2008) for details of these design methods. All design attributes had four levels, including those that are mode-specific. These were chosen as the following variations around the base level: -25%, 0%, +25%, +50%.

[3] Extensive development work was undertaken in the design of the CAPI instrument followed by a pre-pilot of 80 respondents. The pre-pilot data was used to estimate a series of multinomial and nested logit models for the pooled data. On the basis of the review of the pilot output, minor changes to the survey instrument were made.


North-West Sydney Transport

Practice Game

		Light Rail connecting to Existing Rail Line	New Heavy Rail	Bus	Existing M2 Busway	Existing Train line	Car
Main Mode of Transport	Fare (one-way) / running cost (for car)	$ 9.75	$ 4.90	$ 4.90	$ 9.00	$ 6.90	$ 3.60
	Toll cost (one-way)	N/A	N/A	N/A	N/A	N/A	$ 2.45
	Parking cost (one day)	N/A	N/A	N/A	N/A	N/A	$ 6.35
	In-vehicle travel time	66 mins	80 mins	60 mins	70 mins	65 mins	60 mins
	Service frequency (per hour)	8	3	6	4	6	N/A
	Time spent transferring at a rail station	2 mins	0 mins	N/A	N/A	N/A	N/A
Getting to Main Mode	Walk time OR	25 mins	25 mins	15 mins	38 mins	60 mins	N/A
	Car time OR	4 mins	3 mins	3 mins	10 mins	13 mins	N/A
	Bus time	4 mins	5 mins	N/A	8 mins	25 mins	N/A
	Bus fare	$ 1.50	$ 1.10	N/A	$ 1.50	$ 3.75	N/A
	Time Getting from Main Mode to Destination	8 mins	12 mins	5 mins	11 mins	8 mins	19 mins
Thinking about each transport mode separately, assuming you had taken that mode for the journey described, how would you get to each mode?		Walk / Drive / Catch a bus	Walk / Drive / Catch a bus	Walk / Drive	Walk / Drive / Catch a bus	Walk / Drive / Catch a bus	
Which main mode would you choose?		Light Rail	New Heavy Rail	Bus	Existing Busway	Existing Train	Car

Back Next

Figure 1: Example stated choice screen

4. Empirical analysis

Four models of interest are summarised in Table 2[4]. Model 1 (M1) is the standard multinomial logit (MNL) model, Model 2 (M2) is the base random parameter (or mixed logit) model (MXL) in utility space, Model 3 (M3) is the generalized random parameter or mixed logit model (GMXL) that accounts for taste and scale heterogeneity, and Model 4 (M4) is the scale heterogeneity model (SMNL)without taste heterogeneity.

All random parameters are specified with unconstrained triangular distributions and correlation amongst the set of random parameters. The correlation is accommodated by an unrestricted lower triangular matrix, **Γ**. All random parameters are estimated using a panel specification. We ran a series of models (MXL, GMXL, SMNL) with varying numbers of intelligent draws (50 through to 1,000). The results stabilised at 500 draws. The fixed and random parameter estimates associated with the trip time and cost attributes are of the expected sign and statistically significant[5]. Personal income (only significant for GMXL) appears in the utility expression for public transport, indicating that a person on a higher income has a lower probability of choosing public transport (compared to car use).

The overall goodness-of-fit (pseudo R^2) varies from 0.410 for GMXL3 to 0.295 for MNL. MXL and GMXL that allow for taste heterogeneity are a substantial improvement over the multinomial logit model, whose log-likelihood at convergence is -2422.49. In contrast scale MNL is marginally improved over MNL. The Akaike Information Criterion (AIC)[6] clearly indicates that one should choose GMXL over the other models.

The elements of the Cholesky matrix (shown in Table 2 as the diagonal and below diagonal values) show strong evidence of correlated attributes, which makes an uncorrelated specification

[4] All models are estimated using (pre-release) Nlogit5.

[5] We did not find any statistically significant ‘h’ effects as per equations (2) and (6).

[6] AIC = 2*k*-2Ln(*L*) where *k* is the number of parameters in the model, and *L* is the maximised value of the likelihood function for the estimated model.

inappropriate. Of particular note is the statistically significant variance parameter for scale (or τ in equation 2), equal to 0.4109 in the GMXL model and 1.418 for SMNL. This suggests that scale heterogeneity is present even after accounting for correlated random parameters. The estimate of γ in GMXL, which governs how the variance of residual taste heterogeneity varies with scale, is 0.00028, but is statistically not significantly different from zero. What this suggests is that, in equation (2), $\boldsymbol{\beta}_i = \sigma_i[\boldsymbol{\beta} + \boldsymbol{\Delta}\mathbf{z}_i + \boldsymbol{\Gamma}\mathbf{v}_i]$.

Table 1: Summary of model results

Note: All public transport is new heavy rail, light rail, and busway; and existing bus, train, and busway; time is in minutes and cost is in dollars ($2003). T-values are in brackets.

Attribute	Alternatives	M1: Multinomial Logit (MNL)	M2: Mixed Logit (MXL)	M3: Generalised Mixed Logit (GMXL)	M4: Scale MNL (SMNL)
Random Parameters: Mean		All non-random parameters			
Main mode in-vehicle time	All public modes	-0.0481 (23.67)*	-0.0537 (12.6)	-0.0735 (9.95)	-0.0576 (22.10)*
Wait time	All public modes	-0.0270 (2.11)*	-0.0747 (3.17)	-0.0660 (2.36)	-0.0306 (3.39)*
Access time	All public modes	-0.0592 (12.6)*	-0.1064 (8.25)	-0.1087 (7.97)	-0.0666 (14.53)*
Egress travel time	All public transport	-0.0150 (3.1)*	-0.1127 (6.45)	-0.0985 (4.11)	-0.0196 (6.17)*
Main mode in-vehicle cost	All public modes	-0.1845 (13.5)*	-0.2947 (7.70)	-0.3164 (8.48)	-0.2358 (18.66)*
Non-Random Parameters:					
New light rail constant	New light rail	2.4098 (6.44)	1.9450 (4.38)	3.3733 (6.84)	2.4026 (13.70)
New busway constant	New busway	1.249 (3.56)	1.628 (4.37)	2.8672 (6.08)	1.2493 (6.92.)
Existing bus constant	Bus	1.8142 (5.87)	1.8458 (5.31)	3.1042 (7.33)	1.8140 (11.61)
Train constant	Existing and new Train	2.1039 (6.63)	2.215 (5.89)	3.4937 (7.87)	2.1132 (13.85)
Existing busway constant	busway	1.6058 (5.07)	1.8235 (4.99)	3.0240 (6.82)	1.6088 (10.7)
Access bus mode fare	Where bus is access mode	-0.07673 (2.41)	-0.0547 (1.50)	-0.0321 (1.02)	-0.0735 (3.14)
Car cost	Car	-0.1128 (4.05)	-0.2044 (4.53)	-0.1634 (3.09)	-0.1367 (5.51)
Car invehicle time	Car	-0.0340 (8.80)	-0.0480 (8.16)	-0.0482 (7.62)	-0.0307 (10.21)
Car parking cost	Car	-0.0139 (1.97)	-0.0429 (3.25)	-0.06278 (4.27)	-0.0675 (7.07)
Egress travel time	Car	-0.0561 (4.07)	-0.0957 (5.96)	-0.1206 (5.13)	-0.0902 (6.65)
Personal income ($'000s)	Public transport	-0.0026 (1.4)	-0.0016 (1.56)	-0.0099 (2.72)	-0.0003 (1.71)
Random Parameters: Standard deviation					
Main mode in-vehicle time	All public modes	-	0.0753 (8.35)	0.1030 (7.68)	-
Wait time	All public modes	-	0.2795(7.07)	0.4318 (8.52)	-
Access time	All public modes	-	0.1937 (5.97)	0.2230 (6.07)	-
Egress travel time	All public transport	-	0.3012 (8.55)	0.3974 (9,87)	-
Main mode in-vehicle cost	All public modes	-	0.6502 (6.45)	0.6961 (7.08)	-
Cholesky matrix: diagonal values					
Main mode in-vehicle time	All public modes	-	0.0753 (8.35)	0.1030 (7.68)	-
Wait time	All public modes	-	0.2243(7.65)	0.3274 (7.31)	-
Access time	All public modes	-	0.0995 (2.98)	0.1919 (5.28)	-
Egress travel time	All public transport	-	0.2447 (7.58)	0.2741 (6.63)	-
Main mode in-	All public	-	0.5325 (6.91)	0.4146 (3.38)	-

vehicle cost	modes				
Cholesky matrix: Below diagonal values					
Wait: In-vehicle time	All public modes	-	0.1667 (4.16)	-0.2814 (4.93)	-
Access: In-vehicle time	All public modes	-	0.0104 (0.36)	-0.0893 (2.08)	-
Access: Wait time	All public modes	-	-0.1658 (5.44)	0.0703 (2.27)	-
Egress: In-vehicle time	All public transport	-	0.1381 (3.41)	-0.2505 (4.66)	-
Egress: Wait time	All public modes	-	-0.1077 (2.67)	0.1240 (2.33)	-
Egress: Access Time	All public modes	-	-0.0129 (0.30)	0.0684 (1.09)	-
In-vehicle cost: In-vehicle time	All public modes	-	-0.0865 (0.75)	0.0613 (0.47)	-
In-vehicle cost: Wait time	All public modes	-	0.2490 (2.32)	-0.4509 (3.79)	-
In-vehicle cost: Access time	All public transport	-	0.1192 (1.22)	0.3006 (2.88)	-
In-vehicle cost: Egress time	All public modes	-	0.2358 (2.48)	-0.1238 (1.02)	-
Variance Parameter in Scale (τ):			-	0.4109 (7.39)	1.1418 (12.11)
Weighting Parameter γ:		-	-	0.00028 (0.007)	-
Sigma:				-	
Sample Mean		-	-	0.9758	0.8185
Sample Standard deviation		-	-	0.3504	0.8347
Model Fit:					
Log-likelihood at zero		-3580.48			
Log-likelihood at convergence		-2522.49	-2156.88	-2111.62	-2415.54
McFadden Pseudo-R^2		0.295	0.398	0.410	0.325
Info. Criterion: AIC		5076.97	4375.75	4289.25	4865.07
Sample Size		1840			
VTTS ($/person hr)					
Main mode in-vehicle time	All public modes	15.64	10.92 (16.92)	12.60 (6.58)	14.66
Wait time	All public modes	8.78	17.09 (33.84)	13.01 (48.9)	7.79
Access time	All public modes	19.25	18.94 (19.94)	20.94 (4.60)	16.95
Egress travel time	All public transport	4.88	18.80 (30.79)	15.55 (30.25)	4.99
In-vehicle time	Car	18.08	14.09	17.60	13.48

* fixed parameters

A useful behavioural output to compare models is the mean estimates of direct elasticity (Table 3 and Figure 2), since these provide direct evidence on the relative sensitivity of each model in respect of modal shares associated with a change in the level of a specific trip attribute. The formula for calculating the mean elasticities for models MXL and GMXL is given in equation (8).

$$Est.Avg.\frac{\partial \log P_j}{\partial \log x_{k,l}} = \frac{1}{N}\sum_{i=1}^{N} \int_{\boldsymbol{\beta}_i} [\delta_{j,l} - P_l(\boldsymbol{\beta}_i, \mathbf{X}_i)]\beta_k x_{k,l,i} d\boldsymbol{\beta}_i \quad (8)$$

where *j* and *l* index alternatives, x indexes the attribute and *i* indicates the individual. The integrals cannot be computed directly, so they are simulated in the same fashion (and at the same time) as the log likelihood function. Using R simulated draws from the distribution of β_i, we obtain the simulated values of the means of the elasticities (equation 9):

$$Est.Avg.\frac{\partial \log P_j}{\partial \log x_{k,l}} = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{R}\sum_{r=1}^{R} [\delta_{j,l} - P_l(\boldsymbol{\beta}_{i,r}, \mathbf{X}_i)]\beta_{k,i,r} x_{k,l,i} \tag{9}$$

Although some models are capable of producing elasticity distributions, the scale heterogeneity model SMNL is of the MNL form, and hence only mean estimates for each person are meaningful. The best way to compare the evidence across the four models is to take the Mixed Logit model (essentially the 'reference' given the focus on contrasting GMXL and SMNL with MXL) and difference the mean estimates for each other model against this model (Figure 3).

Table 2: Direct time and cost elasticities

(Note: uncalibrated models, standard deviations in brackets)

Attribute	Alternative	M1: Multinomial Logit*	M2: Mixed Logit	M3: Generalised Mixed Logit	M4: Scale MNL*
In-vehicle time	New light rail (invt-NLR)	-1.674 (1.021)	-1.421 (0.758)	-1.481 (0.796)	-1.106 (0.462)
	New heavy rail (invt-NHR)	-1.595 (0.945)	-1.399 (0.684)	-1.533 (0.752)	-1.172 (0.530)
	New busway (invt-NBWY)	-2.133 (0.976)	-1.744 (0.652)	-1.936 (0.747)	-1.415 (0.465)
	Bus (invt-Bus)	-1.773 (0.995)	-1.356 (0.581)	-1.475 (0.650)	-1.260 (0.456)
	Busway (invt-Bway)	-1.540 (0.880)	-1.317 (0.703)	-1.465 (0.809)	-1.188 (0.530)
	Train (invt-Train)	-1.344 (0.752)	-1.227 (0.609)	-1.340 (0.731)	-1.035 (0.469)
	Car (invt-Car)	-1.215 (0.709)	-0.894 (0.648)	-0.763 (0.441)	-0.847 (0.853)
Cost	New light rail (cost-NLR)	-0.699 (0.446)	-0.883 (0.512)	-0.775 (0.475)	-0.493 (0.236)
	New heavy rail (cost-NHR)	-0.704 (0.452)	-0.756 (0.391)	-0.733 (0.389)	-0.547 (0.272)
	New busway (cost-NBWY)	-1.143 (0.496)	-0.917 (0.507)	-0.943 (0.468)	-0.806 (0.319)
	Bus (cost-Bus)	-0.942 (0.486)	-0.826 (0.384)	-0.815 (0.389)	-0.770 (0.326)
	Busway (cost-Bway)	-0.646 (0.414)	-0.758 (0.396)	-0.739 (0.427)	-0.522 (0.264)
	Train (cost-Train)	-0.832 (0.483)	-0.713 (0.368)	-0.686 (0.351)	-0.626 (0.281)
	Car (cost-Car)	-0.580 (0.339)	-0.537 (0.387)	-0.363 (0.209)	-0.528 (0.530)

* The standard deviations are an artifact of different choice probabilities and not a result of preference heterogeneity

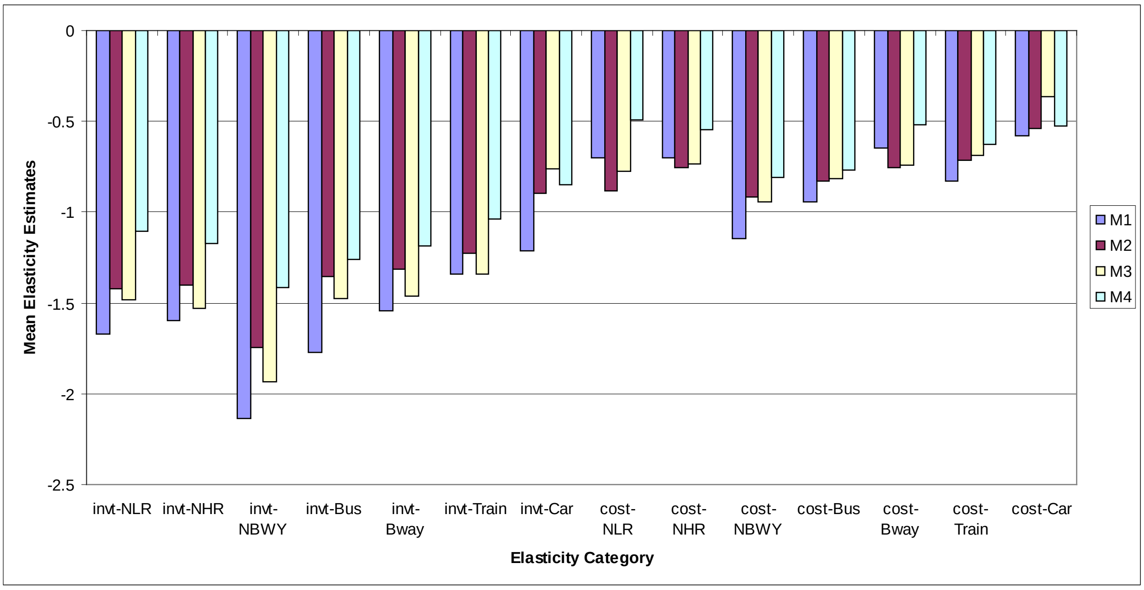


Figure 2: Contrasts of direct elasticities

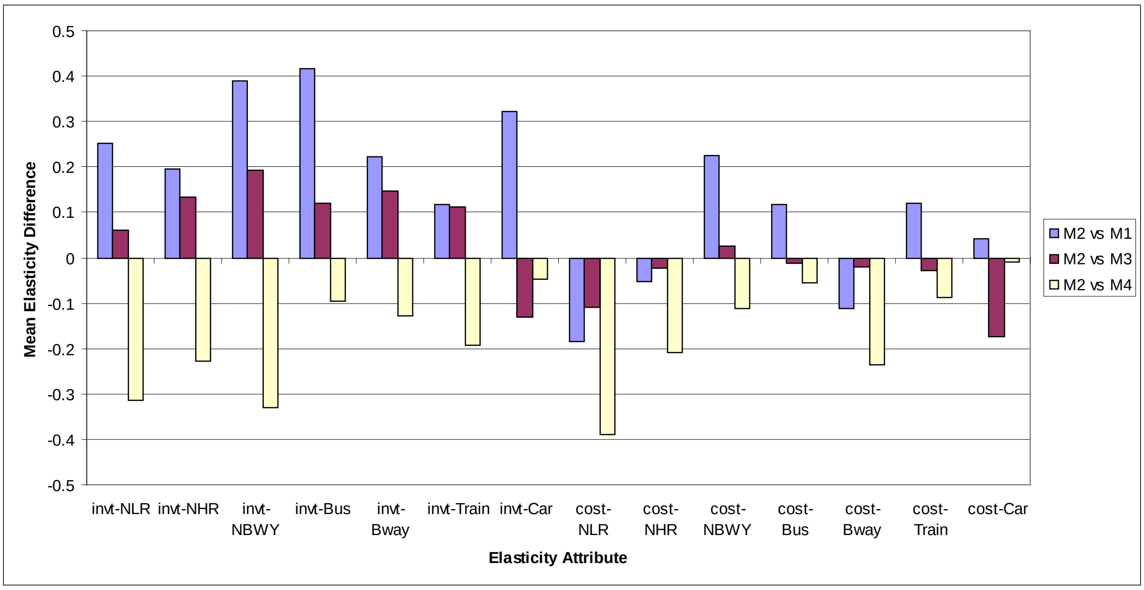


Figure 3: Contrasts of direct elasticities: MXL (M2) vs. MNL (M1), GMXL (M3) and SMNL (M4)

Beginning with the in-vehicle time mean elasticities[7] (the left half of Figure 3), the SMNL model (M4) has the greatest consistently negative difference[8] relative to MXL (M2), remaining directionally negative for all modal alternatives, indicating that all mean elasticities are higher for mixed logit compared to scale MNL. In contrast, mixed logit has higher mean elasticity estimates than MNL (M1) and GMXL (M3) with the one exception of car in-vehicle time for GMXL.

This evidence, albeit from one study, suggests that the SMNL model, that excludes consideration of attribute preference heterogeneity, produces noticeably lower mean estimates of the elasticities for in-

[7] The elasticities are based on uncalibrated models and as such the numerical magnitudes are only valid in the comparisons across models. These models cannot be used to forecast patronage without calibration using revealed preference shares on existing modes.

[8] Since all elasticities are negative, a lower value is an absolute lower value (e.g. -0.435 is lower than -0.650).

vehicle travel time. For the cost attribute, the same findings apply for mixed logit compared to SMNL; however the directional implication is not clear in comparisons of MXL with MNL and GMXL.

When we undertake a statistical test of differences (using the mean and standard deviation) between various model pairs (see Table 4), we find on the t-ratio of differences test, that there is no statistically significant difference between the mean estimates, without exception[9]. Hence the extension from MNL to mixed logit to generalized mixed logit, and the focus only on scale heterogeneity, does not impact materially on the evidence on direct elasticities, despite the actual mean estimates that are typically used in practice being different in absolute terms.

This empirical evidence suggests that although recognition of preference and scale heterogeneity through observed attributes improves on the goodness-of-fit of the models (Table 2), and aligns the mean elasticity estimates 'closer' to those of the popular mixed logit model (which assumes scale homogeneity), the differences are not statistically significant. However, despite this evidence, practitioners tend to focus on applying the mean estimates, and hence when only scale heterogeneity is accommodated, the mean elasticity estimates are, with a few exceptions, noticeably lower than both mixed logit and generalized mixed logit.

Table 4: Tests of statistical significance between elasticity estimates

Attribute	Alternative	MXL vs. MNL	MXL vs. GMXL	MXL vs. SMNL	GMXL vs. SMNL
In-vehicle time	New light rail (invt-NLR)	0.199	-0.055	0.355	0.407
	New heavy rail (invt-NHR)	0.168	-0.131	0.193	0.393
	New busway (invt-NBWY)	0.331	-0.194	0.411	0.592
	Bus (invt-Bus)	0.362	-0.137	0.130	0.271
	Busway (invt-Bway)	0.198	-0.138	0.147	0.286
	Train (invt-Train)	0.121	-0.119	0.250	0.351
	Car (invt-Car)	0.334	0.167	0.044	0.087
Cost	New light rail (cost-NLR)	-0.197	0.155	0.692	0.532
	New heavy rail (cost-NHR)	0.087	0.042	0.439	0.392
	New busway (cost-NBWY)	0.319	-0.038	0.185	0.242
	Bus (cost-Bus)	0.187	0.020	0.111	0.089
	Busway (cost-Bway)	-0.196	0.033	0.496	0.432
	Train (cost-Train)	0.196	0.053	0.188	0.133
	Car (cost-Car)	0.084	0.396	0.014	-0.290

This evidence, admittedly from a single study, raises doubts about the substantive empirical merits of allowing for scale heterogeneity in the absence of the influence of preference heterogeneity, given that Model 3 is the preferred model. When both sources of heterogeneity are captured, the statistical fit of the GMXL model is superior (with a two degrees of freedom difference), suggesting that accounting for both preference and scale heterogeneity is a significant improvement over the standard mixed logit model. In terms of the behavioural implications associated with mean direct elasticities, however, this tends to result in slightly lower travel time estimates and slightly higher travel cost estimates; however given the standard deviations, the difference is not statistically significant.

Finally, we report the mean estimates of values of travel time savings (Table 2)10. We calculated the mean WTP (and standard deviation where appropriate) using the unconditional estimates, and we used the generic in-vehicle cost parameter to obtain VTTS. There are differences in the mean estimates for all time attributes; however the differences between ML and GMXL are not statistically significant on a test of differences, given the standard errors. What does appear notable is the presence of lower mean estimates for SMNL compared to MNL (recognising the value for egress time is very similar). The behavioural implications are far from clear other than that the mixed logit and generalized mixed

[9] We also undertook a bootstrap calculation for two of the variables to ensure that the t-ratio test was a useful approximation. The resulting standard errors confirm that the t-ratios are a good approximation.

[10] In this paper all models are estimated in preference space. We have estimated a GMXL model using the same data in WTP space in Hensher and Greene (2009).

logit models appear to produce higher mean estimates than the models that assume preference homogeneity. This finding is known from other studies (see Hensher 2010).

Conclusions

This paper has set out a number of discrete choice model forms that are able to account for sources of observed and unobserved heterogeneity across individuals. The particular focus is on identification of preference (or taste) heterogeneity and scale heterogeneity, and the incremental contribution of allowing for preference heterogeneity, scale heterogeneity, and both sources of heterogeneity, with scale heterogeneity suggested by a number of authors as a neglected source of heterogeneity.

Our empirical assessment suggests that accommodating scale heterogeneity in the absence of accounting for preference heterogeneity may be of limited empirical interest, resulting in a statistically inferior model form, despite it being an improvement over the standard MNL model, as might be expected. Scale heterogeneity in the presence of preference heterogeneity does garner favour, with the generalised mixed logit model an improvement over the standard mixed logit model. These findings accord with the evidence in Fiebig et al. (2009) where, across 10 data sets, on the AIC test, the generalized mixed logit model performs better than mixed logit in eight data sets; however the SMNL model performs relatively poorly in all cases.

Compared to a failure to account for preference heterogeneity that is consequential, failure to account for scale heterogeneity may not be of such great empirical consequence in respect of behavioural outputs such as direct elasticities and willingness to pay.

Clearly, a library of empirical evidence from other data sets is required before we can make any definitive statements about the extent to which analysts should routinely allow for scale heterogeneity across individuals in the presence and/or absence of taste heterogeneity.

References

Breffle, W.S., and Morey, E.R. (2000) Investigating preference heterogeneity in a repeated discrete-choice recreation demand model of Atlantic salmon fishing, Marine Resource Economics, 15, 1-20.

Fiebig, D., Keane, M., Louviere, J., and Wasi, N. (2009) The generalized multinomial logit: accounting for scale and coefficient heterogeneity, Marketing Science, published online before print July 23, DOI:10.1287/mksc.1090.0508

Fosgerau, M. (2007) Using nonparametrics to specify a model to measure the value of travel time, Transportation Research A, 41(9), 842-856.

Fosgerau, M. (2006) Investigating the distribution of the value of travel time savings, Transportation Research B, 40(8), 688-707.

Greene, W.H. (2007) Nlogit 4, Econometric Software, New York and Sydney.

Hensher, D.A. and Greene, W.H. (2003) Mixed logit models: state of practice, Transportation, 30 (2), 133-176.

Hensher, D.A. (2010) Attribute Processing, Heuristics and Preference Construction in Choice Analysis. Invitational Keynote Paper for Choice Modelling Conference, Leeds UK. March 30-April 1 2009, in Hess, S. and Daly, A. (eds.) Choice Modelling, Emerald Press, UK.

Hensher, D.A. and Rose, J.M (2007) Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: a case study, Transportation Research A, 41 (5), 428-433.

Hensher, D.A. and Greene, W.H. (2009) Valuation of Travel Time Savings in WTP and Preference Space in the Presence of Taste and Scale Heterogeneity, Institute of Transport and Logistics Studies, University of Sydney, November.

Hensher, D.A., Louviere, J.J. and Swait, J. (1999) Combining Sources of Preference Data, Journal of Econometrics, 89, 197-221.

Hess, S., Rose, J.M. and Bain, S. (2009) Random scale heterogeneity in discrete choice models, mimeo, June 23.

Keane, M. (2006) The Generalized Logit Model: Preliminary Ideas on a Research Program, presentation at Motorola-CenSoC Hong Kong Meeting, October 22, 2006.

Louviere, J. and T. Eagle (2006) Confound it! that pesky little scale constant messes up our convenient assumptions, Proceedings, 2006 Sawtooth Software Conference, 211-228, Sawtooth Software, Sequem, Washington, USA.

Louviere, J. J., R. J. Meyer, D. S. Bunch, R. Carson, B. Dellaert, W. M. Hanemann, D.A. Hensher and J. Irwin (1999) Combining Sources of preference data for modelling complex decision processes, Marketing Letters, 10:3, 205-217.

Louviere, J.J., R.T. Carson, A. Ainslie, T. A. Cameron, J. R. DeShazo, D. A. Hensher, R. Kohn, T. Marley and D.J. Street (2002) Dissecting the random component of utility, Marketing Letters, 13, 177-193.

Louviere, J. J., D. Street, L. Burgess, N. Wasi, T. Islam and A. A. J. Marley (2008) Modelling the choices of individuals decision makers by combining efficient choice experiment designs with extra preference information," Journal of Choice Modelling, 1:1, 128-163.

Rose, J.M. and Bleimer, M.C.J. (2008) Stated preference experimental design strategies, in Hensher, D.A. and Button, K.J. (eds.) Handbook of Transport Modelling, Elsevier, Oxford, Ch 8, 151-180.

Rose, J.M., Bliemer, M.C., Hensher and Collins, A. T. (2008) Designing efficient stated choice experiments in the presence of reference alternatives, Transportation Research Part B 42 (4), 395-406.

Scarpa, R., Campbell, D. Hutchinson, W. G. (2007) Benefit estimates for landscape improvements: sequential Bayesian design and respondents' rationality in a choice experiment study. Land Economics. (November) 83(4):617-634.

Scarpa, R., Thiene, M. and Train, K. (2008) Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps, American Journal of Agricultural Economics, 90(4):994-1010. (also seeAppendix: Utility in WTP space: a tool to address confounding random scale effects in destination choice to the Alps. Available at: http://agecon.lib.umn.edu/).

Scarpa, R., M. Thiene, F. Marangon (2008a) Using flexible taste distributions to value collective reputation for environmentally-friendly production methods. Canadian Journal of Agricultural Economics, 56:145-162

Sonnier, G., Ainslie, A. and Otter, T. (2007) Heterogeneity distributions of willingness to-pay in choice models, Quantitative Marketing Economics 5(3), 313–331.

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

Train, K. and Weeks, M. (2005) Discrete choice models in preference space and willing to-pay space, in R. Scarpa and A. Alberini, eds., Applications of Simulation Methods in Environmental and Resource Economics, Springer Publisher, Dordrecht, ch. 1, 1–16.