Rethinking heterogeneity: the role of attitudes, decision rules and data processing strategies

Stephane Hess*

March 8, 2012

Abstract

The study of heterogeneity across individual decision makers is one of the key areas of activity in the field of behavioural research. However, a disproportionately large share of the research effort focusses on heterogeneity in sensitivities to individual attributes, and in particular how such heterogeneity can be accommodated in a random coefficients framework. While differences in marginal sensitivities clearly play a role in driving behaviour, this paper makes the case that *retrieved differences* in such sensitivities may in fact be caused by a number of different factors. In particular, we look at the possible role of underlying attitudes, differences in decision rules across respondents and the role of information processing strategies. We show evidence from a number of studies that suggest that accounting for such richer behavioural patterns leads to important gains in understanding of behaviour, and may also reduce the level of residual *random* heterogeneity. Conversely, this suggests that not adequately accounting for such additional factors may overstate the degree of unexplained heterogeneity in marginal sensitivities.

Key words: heterogeneity; discrete choice; behavioural mixing; decision rules; information processing; latent attitudes

1 Introduction

The representation of heterogeneity across respondents is a core topic of interest in the field of travel behaviour research. For over two decades after random utility models first became widely used, the focus was mainly on linking such taste heterogeneity to sociodemographic characteristics. However, with the important developments in computational performance and estimation efficiency that took place from the mid 1990s onwards, there

*Institute for Transport Studies, University of Leeds, Leeds LS2 9JT, s.hess@its.leeds.ac.uk, Tel: +44 (0)113 34 36611, Fax: +44 (0)113 343 5334

has been a significant shift towards random coefficient models such as Mixed Multinomial Logit (MMNL) that allow for unexplained variations in sensitivities across respondents.

While a potentially large share of the heterogeneity found in most datasets can indeed only be explained in such a random manner, it should be noted that the growing popularity of random coefficients models has also led to reduced effort in trying to explain taste heterogeneity in a deterministic fashion. This is unfortunate, as from the point of view of interpretation, it is much preferable to explain as much of the heterogeneity as possible by linking it to characteristics of the respondents.

The main focus in choice modelling is on producing estimates of the relative sensitivities to different attributes, and by extension the variation of such relative sensitivities across respondents. However, independently of whether efforts have been made to link heterogeneity to socio-demographics, there is also commonly a lack of recognition that factors other than differences in relative sensitivities may be causing the *retrieved heterogeneity*. The aim of the present paper is to look at three specific factors that may influence our findings in terms of heterogeneity across respondents, namely:

1. the role of attitudes and perceptions;
2. heterogeneous decision rules; and
3. heterogeneous strategies for processing the information describing choice tasks.

This paper argues that not accounting for these factors may lead to misleading interpretations of the heterogeneity findings and reduce insights into individual-level decision making processes.

The remainder of this paper is organised as follows. The following section presents an overview of the three phenomena discussed above, and how they can be modelled in practice. This is followed by a number of case studies. Finally, the conclusions of the work are presented.

2 Behavioural phenomena and modelling approaches

In a typical discrete choice model, we define the modelled utility of alternative i for respondent n in choice situation t as:

$$V_{i,n,t} = f\left(x_{i,n,t}, \beta\right), \tag{1}$$

where $x_{i,n,t}$ is a vector of attributes describing the alternative as well as decision maker n in choice situation t, and β is a vector of parameters to be estimated from the model. Finally, $f\left(\cdot\right)$ gives the functional form of the utility function. We then define $P_{i,n,t}\left(\beta\right)$ to give the probability of respondent n choosing alternative i (out of J) in choice situation t, where this is a function of the estimated parameters β. Looking in particular at the MNL case, we have that:

$$P_{i,n,t}\left(\beta\right) = \frac{e^{V_{i,n,t}}}{\sum_{j=1}^{J} e^{V_{j,n,t}}}, \tag{2}$$

with, as before, $V_{j,n,t} = f(x_{j,n,t}, \beta)$. In the widely used MMNL model (cf. McFadden and Train, 2000; Hensher and Greene, 2003), we now assume that β follows a random distribution with a vector of parameters Ω, say $\beta \sim g(\beta \mid \Omega)$. We generally also assume that tastes vary across respondents but stay constant across choices for the same respondent (cf. Revelt and Train, 1998). The probability of the observed sequence of choices for respondent n is then given by:

$$P_n(\Omega) = \int_\beta \prod_{t=1}^{T_n} P_{i*,n,t}(\beta)\, g(\beta \mid \Omega)\, \mathrm{d}\beta, \tag{3}$$

where $P_{i*,n,t}$ refers to the probability of the actually chosen alternative for respondent n in task t, where respondent n faces T_n choice tasks.

This specification in Equation 3 has been shown to lead to very significant gains in model fit over the specification in Equation 2. However, from an interpretation perspective, it is not satisfying to simply observe a high level of heterogeneity without understanding what may be causing this heterogeneity in our estimates. In this section, we now discuss a number of possible interpretations.

2.1 Attitudes and perceptions

Underlying attitudes, perceptions and personal convictions are potentially a key driver of individual-specific preferences, and surveys routinely collect responses to attitudinal questions. However, the way in which the role of attitudes can best be accommodated in our existing modelling frameworks is not straightforward, a point largely ignored in a number of fields. The main misunderstanding is that responses to attitudinal questions are seen as direct measures of attitudes. However, these answers, often referred to as indicators, are themselves clearly only a function of underlying attitudes, rather than a direct measure of attitudes. The direct incorporation of these indicators in models means that results are possibly affected by measurement error. More importantly however, the responses to attitudinal questions are likely be correlated with other unobserved factors that enter the error term of the models. This thus leads to potential problems with endogeneity bias. Finally, answers to attitudinal questions will clearly not be available into the future, ruling out the possibility of using such a model in forecasting.

A more appropriate approach is to treat the actual attitudes as latent variables (cf. Ben-Akiva et al., 1999; Ashok et al., 2002; Ben-Akiva et al., 2002; Bolduc and Alvarez-Daziano, 2010). In a model looking at the impact of latent variables, we make use of a number of indicators that serve as proxies for these latent variables, typically in the form of responses to attitudinal questions. The value of these indicators is then modelled jointly with the actual choice processes, based on the assumption that both processes are at least in part influenced by the latent attitudes. This approach thus integrates choice models with latent variable models resulting in an improvement in the understanding of preferences as well as explanatory power. A main benefit of using a latent variable approach is to overcome the bias inherent in a direct incorporation of indicators of attitudes (or other subjective measures) in the utility function. The resulting model can also be used in forecasting, as, once the latent variable specification has been calibrated, data on answers to attitudinal questions is not required for model application.

For the sake of simplified notation, we rely on a single latent attitude, which, for respondent n, we define as:

$$\alpha_n = l\left(z_n, \gamma\right) + \eta_n, \tag{4}$$

where $l\left(z_n, \gamma\right)$ represents the deterministic part of α_n, with z_n being a vector of socio-demographic variables of respondent n, and γ being a vector of estimated parameters, and where a decision on the specification of $l\left(\right)$ needs to be taken (e.g. linear). The term η_n is a random disturbance, which we assume follows a Normal distribution (say $h\left(\eta\right)$) across respondents, with a zero mean and a standard deviation of σ_α, which is set to 1 for identification reasons.

This latent variable α_n is then interacted with parameters in our choice model. As an example, we might rewrite Equation 1 as:

$$V_{i,n,t} = f\left(\alpha_n, \tau, \beta, x_{i,n,t}\right), \tag{5}$$

where τ is a vector of parameters that interact α_n with β and x_{int}.

At the same time, α_n is also used to explain the answers to a series of attitudinal questions, say $I_{k,n}$, $k = 1, \ldots, K$. This process is modelled through a set of measurement equations, where the specific functional form depends on the data at hand. As an example, for the typically used Likert-scale ratings, an ordered logit specification is most appropriate, as discussed by Daly et al. (2012), while, for ranking data, we might wish to use a rank exploded logit model. Each time, the latent variables are used as a key explanator in these measurement models.

The use of α_n in the choice model as well as measurement model components means that the estimation of α_n is informed both by the data on choices and the data on responses to attitudinal questions. We now let $L\left(C_n \mid \alpha_n, \beta, \tau\right)$ give the likelihood of the observed sequence of choices for respondent n, and $L\left(I_n \mid \zeta, \alpha_n\right)$ give the probability of observing the specific responses given by respondent n to the various attitudinal questions, with ζ representing a vector parameters for the measurement model.

Both $L\left(C_n \mid \cdot\right)$ and $L\left(I_n \mid \cdot\right)$ are conditional on a specific realisation of the latent variable α_n. Given the random component in α_n, we thus need to integrate over the distribution of η, i.e. $h\left(\eta\right)$. If we have additional random heterogeneity in the β coefficients, additional layers of integration need to be added, and we would have:

$$LL\left(\Omega, \gamma, \tau, \zeta_I, \sigma_I\right) = \sum_{n=1}^{N} \ln \int_\beta \int_\eta L\left(C_n \mid \alpha_n, \beta, \tau\right) L\left(I_n \mid \cdot, \alpha_n\right) h\left(\eta\right) g\left(\beta \mid \Omega\right) \mathrm{d}\beta \mathrm{d}\eta, \tag{6}$$

where this is integrated over the distribution of η, the random component in the latent variable, and β, with $\beta \sim g\left(\beta \mid \Omega\right)$.

2.2 Decision rules

Although structures belonging to the family of random utility models have come to dominate, it is important to recognise that alternative paradigms for decision making have been proposed, for example the elimination by aspects model of Tversky (1972), but also more recent work based on the concepts of happiness by Abou-Zeid and Ben-Akiva (2010) and regret by Chorus et al. (2008). The evidence in the literature is that which paradigm works best is very much dataset specific. However, this misses the crucial point that the variations in decision rules may be across respondents with a single dataset, not just across datasets. Similarly, the question arises whether what is retrieved as random taste heterogeneity in MMNL models may in fact be caused by heterogeneity in decision rules.

Such differences in behavioural models across respondents can be accommodated relatively easily in a latent class (LC) style approach, as recently discussed by Hess et al. (2012). Specifically, while in a standard latent class model for the analysis of discrete choice data, the differences between classes lie in the use of class specific utility parameters aimed at retrieving heterogeneity in sensitivities, this model goes further by allowing for differences across classes in the actual behavioural process.

Let $L\left(C_n \mid \beta_m, m\right)$ give the probability of the observed sequence of choices for respondent n, conditional on using a choice model identified as M, where this uses a vector of parameters β_m. The Hess et al. (2012) framework is based on the idea that M different behavioural processes are used in the data. The choice of decision rule for given respondent is not observed and is thus treated as a latent component. The probability for the sequence of choices observed for respondent n is now given by:

$$L\left(C_n \mid \beta, \pi\right) = \sum_{m=1}^{M} \pi_{n,m} L\left(C_n \mid \beta_m, m\right) \text{ where } \sum_{m=1}^{M} \pi_{n,m} = 1 \text{ and } 0 \leq \pi_m \leq 1, \forall m, \quad (7)$$

where we use different behavioural processes in different classes.

With this model, we need to estimate parameters of the choice models in the individual classes ($\beta_m,\ m = 1, \ldots, M$, grouped together into β), as well as the probabilities for all classes ($\pi_m,\ m = 1, \ldots, M$, grouped together into π). Here, it is possible to link class allocation to respondent characteristics, by formulating a class allocation model.

2.3 Processing strategies

Over the last few years, a substantial amount of research effort has gone into investigating the possibility of individual respondents using different strategies in processing the information describing the choice tasks they face; see Hensher (2010) for a comprehensive overview. In particular, the main emphasis has been on the notion that some respondents may ignore certain attributes, although there has also been some interest in looking at whether specific individuals may process *similar* attributes jointly rather than separately (cf. Layton and Hensher, 2010).

Stated choice surveys now routinely include questions asking respondents whether they ignored certain attributes. Early work in this context deterministically imposed the processing rule on the basis of such stated information, but just as with the deterministic treatment of responses to attitudinal questions (cf. Section 2.1), this leads to issues with endogeneity. Additionally, the question arises as to how reliable this information is (cf. Hess and Hensher, 2010). Later work made use of more reliable approaches that treat the processing strategies as latent components (see e.g. Hensher, 2008).

It should be clear that any heterogeneity across respondents in how the information describing the choice tasks is processed may have substantial impacts on findings in terms of taste heterogeneity when not properly accommodated in the model. This is clearly especially true in the case where some attributes are ignored by certain respondents, as the resulting spike at zero in the *true* distribution may be difficult to accommodate in the *fitted* distribution.

3 Empirical evidence

In this section, we briefly summarise empirical evidence from a number of recent studies. We focus on the possibility that the heterogeneity retrieved in simple MMNL structures may to a certain extent be influenced by other factors, and that allowing for more behaviourally rich behavioural patterns may lead to different insights into heterogeneity. In particular, we look at whether there is evidence that allowing for these richer behavioural patterns leads to reductions in residual purely random heterogeneity.

3.1 Attitudes

Our first case study makes use of data collected by Hess and Stathopoulos (2011) in an online SC survey conducted on UK rail and bus commuters in January 2010. Each respondent was faced with three alternatives (on the current mode) described by travel time, fare, the rate of crowding (out of 10 trains), the rate of delays (out of 10 trains), the average extent of delays, and the provision of a delay information service (via sms text message).

Two different models were estimated on this dataset. In the first model, we allow for random heterogeneity in all five marginal utility coefficients, using Lognormal distributions with appropriate sign changes, where, for crowding and the rate of delays, the attributes were transformed to be on a scale from 0 (i.e. no journeys affected) to 1 (i.e. all journeys affected). In this model, $\mu_{\ln(\beta_{tt})}$ gives the mean of the underlying Normal distribution for the travel time sensitivity, with the corresponding standard deviation being given by $\sigma_{\ln(\beta_{tt})}$. We also use constants for the first two alternatives.

In the second model, identified as the Integrated Choice and Latent Variable (ICLV) model, we incorporate the role of an underlying attitude towards delays, by making use of two specific indicators in a measurement model component, namely respondents' answers to the following two statements:

I_1**:** *Commuters should accept that sometimes delays are inevitable*

I_2**:** *A fast and reliable service should be the absolute priority*

A single latent variable α was specified, given by a standard Normal variate with no socio-demographic interactions. This latent variable was used to explain the answers to the two attitudinal questions, which were on a 1 to 5 scale, where the order of answers for the first question was reversed for consistency across indicators. We used an ordered logit model in the measurement equations, with four thresholds (θ) estimated for each indicator. In addition, the latent variable α was interacted with $\beta_{\text{rate of delays}}$, which is now given as $\beta_{\text{rate of delays}} = -e^{\mu_{\ln\left(\beta_{\text{rate of delays}}\right)}+\sigma_{\ln\left(\beta_{\text{rate of delays}}\right)}\cdot\xi_4+\tau_{\ln\left(\beta_{\text{rate of delays}}\right)}\cdot\alpha}$, where ξ_4 is the standard Normal variate used for the rate of delays coefficient (fourth coefficient).

The results for this study are summarised in Table 1. The model fits for the two models are not comparable given the use of additional data (observations on indicators) in the second model. Both models obtain significant estimates for the means and standard deviations of the underlying Normal distributions for all parameters. In the ICLV model, the interaction term $\tau_{\ln\left(\beta_{\text{rate of delays}}\right)}$ is positive and significant, indicating that as the latent variable α increases, so does the sensitivity to the rate of delays (remembering the sign change on the Lognormal distribution). At the same time, the role of the latent variable in the ordered logit model (with increasing thresholds implying that a higher value for the latent variable leads to a higher value for the indicator) means that a respondent with a higher value for the latent variable is more likely to *disagree* with the statement that *"Commuters should accept that sometimes delays are inevitable"*, remembering that the order for this indicator was reversed. Similarly, a respondent with a higher value for the latent variable is more likely to *agree* with the statement that *"A fast and reliable service should be the absolute priority"*. In conjunction, these findings validate the notion that this latent variable captures an underlying attitude towards delays.

In the context of the present paper, the key interest is in the second part of Table 1, showing the implied willingness-to-pay (WTP) measures. We observe that the inclusion of the latent attitude towards delays has an impact on all five WTP measures, with lower, and arguably more realistic, mean values in the ICLV model, along with reduced levels of heterogeneity (in terms of coefficient of variation, *cv*). This suggests that the ICLV model is able to produce more reliable results, thanks to its greater model flexibility and by making use of the additional data on attitudinal questions.

3.2 Mixing of decision rules

We next turn our attention to evidence from the work of Hess et al. (2012), who discuss the benefits of a modelling approach in which different decision rules are employed within a probabilistic framework, thus allowing for the possibility that different individuals within the same dataset make use of different behavioural processes.

Table 2 summarises the results from two of the case studies discussed in Hess et al. (2012). Each time, a simple MMNL model is compared to a model that combines this MMNL model with models using a different behaviour process.

In the first study summarised in Table 2, this alternative model is a lexicography model. In particular, the models are estimated on data from the Danish Value of Time study, in the form of a simple binary stated choice (SC) experiment with two attributes, travel time and travel cost. The first class in the model is a simple MMNL structure. However, the second class aims to capture respondents who make their choices solely on the basis of the cost attribute, while the third class aims to capture respondents who make their choices solely on the basis of the time attribute. Only respondents who behaved in an apparent lexicographic manner were 'eligible' to be captured by these classes. Importantly, the fact that the first class uses a MMNL model means that this class can capture respondents who behave in an apparently lexicographic manner but whose choices are simply caused by very strong (but not extreme) sensitivities. Thus, apparent lexicographic respondents are not simply excluded from the RUM part of the model.

The results show a statistically significant improvement in model fit by (over the simple MMNL model) by 10.29 units in log-likelihood at the cost of two additional parameters (noting that the lexicography classes are deterministic, i.e. do not use any additional parameters). We see small but statistically significant weights for the two lexicography classes ($\pi_{\text{lex-cost}}$ and $\pi_{\text{lex-time}}$). More importantly, we note a very noticeable impact on the findings in terms of heterogeneity in the MMNL class when comparing the sample level MMNL results to the results for the MMNL class within the joint model, where these impacts are disproportional compared to the small weights for the two lexicography classes. This highlights the strong influence that a small group of respondents may have on our sample level findings in terms of heterogeneity.

The second set of results relate to a case study in which a MMNL model is combined with a elimination by aspects (EBA) model, in which in each step, the choice is made on the basis of just one attribute, where in the case of ties, we move to the next important attribute. This EBA model is thus essentially the same as a lexicography model. This study makes use of data from a survey involving the choice between three rail journeys, described by travel time, fare, the guarantee of a reserved seat, the provision of wifi, and whether the ticket is flexible.

While the weight of the EBA class (π_{EBA}) is modest compared to that of the MMNL class, we once again observe statistically significant gains in log-likelihood, with an improvement by 96 units at the cost of 5 additional parameters. Additionally, we note changes in the heterogeneity patterns of the MMNL class, with increases in the mean WTP measures (comparing the sample level MMNL results to the results for the MMNL class within the joint model), but drops in heterogeneity. This again shows the large impact that a small subset of the data can have on overall results.

3.3 Information processing and its impact on retrieved heterogeneity

Our final example reports evidence from Hess and Hensher (2010) in the context of work looking at heterogeneous information processing strategies, and in particular the case of respondents who ignore specific attributes when making their choices. Rather than focussing on appropriate ways of accommodating such respondents within a model, this work looks at identifying them and quantifying their impact, especially in terms of heterogeneity.

The analysis makes use of data from a toll road study with five attributes; free flow travel time, slowed down travel time, travel time variability, running costs, and tolls, with each respondent facing 16 choice scenarios, each with three alternatives. Direct questioning of respondents after the SC component of the survey revealed high rates of stated attribute ignoring, with 12.7% respondents indicating that they ignored free flow time across all 16 tasks, with corresponding figures of 15.6% for slowed down time, 29.8% for travel time variability, 28.8% for running costs, and 8.8% for tolls.

A first analysis by Hess and Hensher (2010) estimates separate coefficients for respondents who indicate that they ignored a given attribute. Crucially, the results show that the coefficients in the *ignoring* group are still significantly different from zero, albeit smaller than in the *non-ignoring* group, casting some doubt on the appropriateness of the stated information on processing strategies.

Rather than using this information, Hess and Hensher (2010) attempt to infer the actual processing strategies from the data, on the basis of individual specific conditional parameter distributions obtained from MMNL models. They assign respondents to the *ignoring* and *non-ignoring* classes on the basis of the probability that the sensitivity for a given attribute is zero, and then estimate group specific models. This thus equates to a two-stage approach.

The results in Table 3 show very visible differences in group allocation in comparison with the stated ignoring strategies, along with lower rates, most notably for tolls (2% instead of 9%). Unlike with the results using the stated data on ignoring, the models show that the sensitivities in the inferred *ignoring* group are indeed equal to zero (see details in Hess and Hensher 2010).

For the present paper, the key interest lies in the impact on the findings in terms of heterogeneity, comparing the sample level degree of heterogeneity (in terms of coefficient of variation) to the heterogeneity in the inferred non-ignoring group. Here, the results in Table 3 show that for some of the attributes, notably free flow time and running costs, respondents who consistently ignore a given attribute have a very large influence on the findings in terms of heterogeneity, which is a direct result of standard distributions not being able to capture a spike at zero that occurs in the *true* distribution.

4 Conclusions

A growing number of travel behaviour studies now make use of random coefficients models, in particular Mixed Logit. These models are generally observed to lead to significant gains in model fit, along with often fundamentally different results in terms of key outputs such as the value of time or other willingness-to-pay indicators. However, there is little recognition in the field that *retrieved* patterns of heterogeneity may in fact be caused by a diverse set of phenomena.

The present paper has looked at three specific behavioural phenomena that could potentially play a role in influencing our findings in terms of heterogeneity in "standard" random coefficients models. In particular, we have discussed:

- the role that underlying attitudes and perceptions play in driving taste heterogeneity;
- the impact of different behavioural rules; and
- the role of information processing strategies in SC surveys.

The paper has summarised empirical evidence from three separate case studies each showing that significant gains in understanding behaviour can be obtained by allowing for these diverse phenomena in our models. Crucially, the results show how, if not properly accounted for in our models, these behavioural traits may have a large influence on our findings in terms of random heterogeneity, even when they only apply to a small share of respondents in the sample.

In summary, while Mixed Logit has allowed us to move away from an approach assuming taste homogeneity, there is now growing evidence that differences between respondents go beyond simple variations in marginal utilities. We need to acknowledge how the choice is approached and how decisions are actually made in real life. The emphasis should now be on understanding heterogeneity, and modelling it in flexible frameworks. The recommendation from this paper is that analysts need to invest considerably more effort in understanding the drivers of heterogeneity in their data, and make appropriate provisions in model specification, rather than just relying on a crude albeit powerful Mixed Logit specification.

As always, there is a need for the findings from this paper to be confirmed in other research, using different datasets. An important further area for work is also to test the impact that the representation of heterogeneity has on forecasting performance, for example using hold out samples.

Acknowledgments

This paper draws extensively on past collaborations with Amanda Stathopoulos, David Hensher and Andrew Daly. Without them, this work would not have been possible - any mistakes in the present paper remain the responsibility of the author.

References

Abou-Zeid, M., Ben-Akiva, M. (2010). A model of travel happiness and mode switching. In: Hess, S., Daly, A. (Eds.), Choice Modelling: The State-of-the-Art and the State-of-Practice. Emerald Publishing, Bingley, UK, 289–305.

Ashok, K., Dillon, W. R., Yuan, S. (2002). Extending discrete choice models to incorporate attitudinal and other latent variables. *Journal of Marketing Research*. 39 (1), 31–46.

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

Ben-Akiva, M., Walker, J., McFadden, D., Grling, T., Gopinath, D., Bolduc, D., Brsch-Supan, A., Delqui, P., Larichev, O., Morikawa, T., Polydoropoulou, A., Rao, V. (1999). Extended framework for modeling choice behavior. *Marketing Letters*. 10 (3), 187–203.

Bolduc, D., Alvarez-Daziano, R. (2010). On estimation of hybrid choice models. In: Hess, S., Daly, A. (Eds.), Choice Modelling: The State-of-the-Art and the State-of-Practice: Proceedings from the Inaugural International Choice Modelling Conference. Emerald Publishing, Bingley, UK, 259–288.

Chorus, C., Arentze, T., Timmermans, H. (2008). A random regret minimization model of travel choice. *Transportation Research Part B*. 42 (1), 1–18.

Daly, A. J., Hess, S., Patruni, B., Potoglou, D., Rohr, C. (2012). Using ordered attitudinal indicators in a latent variable choice model: A study of the impact of security on rail travel behaviour. *Transportation*. 39 (2), 267–297.

Hensher, D. A. (2008). Joint estimation of process and outcome in choice experiments and implications for willingness to pay. *Journal of Transport Economics and Policy*. 42 (2), 297–322.

Hensher, D. A. (20100. Attribute processing, heuristics and preference construction in choice analysis. In: Hess, S., Daly, A. J. (Eds.), State-of Art and State-of Practice in Choice Modelling: Proceedings from the Inaugural International Choice Modelling Conference. Emerald, Bingley, UK, Ch. 3, pp. 35–70.

Hensher, D. A., Greene, W. H. (2003). The Mixed Logit Model: The State of Practice. *Transportation*. 30 (2), 133–176.

Hess, S., Hensher, D. A. (2010). Using conditioning on observed choices to retrieve individual-specific attribute processing strategies. *Transportation Research Part B*. 44 (6), 781–790.

Hess, S., Stathopoulos, A. (2011). Linking response quality to survey engagement: a combined random scale and latent variable approach. ITS working paper. Institute for Transport Studies, University of Leeds.

Hess, S., Stathopoulos, A., Daly, A. J. (2012). Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. *Transportation*. forthcoming.

Layton, D., Hensher, D. A. (2010). Aggregation of common-metric attributes in preference revelation in choice experiments and implications for willingness to pay. *Transportation Research Part D*. 15 (7), 394–404.

McFadden, D., Train, K. (2000). Mixed MNL Models for discrete response. *Journal of Applied Econometrics*. 15 (5), 447–470.

Revelt, D., Train, K. (1998). Mixed Logit with repeated choices: households' choices of appliance efficiency level. *Review of Economics and Statistics*. 80 (4), 647–657.

Tversky, A. (1972). Elimination by aspects: a theory of choice. *Psychological Review*. 79, 281–299.

Table 1: Results for latent attitudes study

	MMNL	ICLV
Observations (choices)	3,680	3,680
Observations (indicators)	0	736
Final LL	-3,004.52	-3,847.21

	MMNL est.	MMNL asy. t-rat	ICLV est.	ICLV asy. t-rat
δ_1	0.7917	9.65	0.8057	10.03
δ_2	0.3109	4.09	0.3100	4.23
$\mu_{\ln(\beta_{\text{tt}})}$	-2.7885	-20.74	-2.7394	-22.08
$\mu_{\ln(\beta_{\text{fare}})}$	1.2704	10.57	1.3156	14.68
$\mu_{\ln(\beta_{\text{crowding}})}$	0.7078	3.27	0.8013	3.83
$\mu_{\ln(\beta_{\text{rate of delays}})}$	1.0385	6.88	0.9336	5.74
$\mu_{\ln(\beta_{\text{av. delay}})}$	-3.4746	-9.52	-3.4773	-13.00
$\mu_{\ln(\beta_{\text{delay sms}})}$	-1.0292	-4.10	-0.6908	-4.48
$\sigma_{\ln(\beta_{\text{tt}})}$	1.1586	9.03	0.9431	7.83
$\sigma_{\ln(\beta_{\text{fare}})}$	1.5809	16.93	1.4599	22.05
$\sigma_{\ln(\beta_{\text{crowding}})}$	1.5806	9.74	1.4891	9.81
$\sigma_{\ln(\beta_{\text{rate of delays}})}$	1.1748	8.21	0.8473	8.95
$\sigma_{\ln(\beta_{\text{av. delay}})}$	1.8159	9.55	1.7520	15.52
$\sigma_{\ln(\beta_{\text{delay sms}})}$	1.0796	6.07	0.7565	9.29
$\tau_{\ln(\beta_{\text{rate of delays}})}$	-		0.9028	10.97
$\theta_{1,1}$	-		-2.6918	-13.17
$\theta_{1,2}$	-		0.8081	6.11
$\theta_{1,3}$	-		2.2993	13.15
$\theta_{1,4}$	-		3.5547	12.87
$\theta_{2,1}$	-		-5.2037	-8.84
$\theta_{2,2}$	-		-4.3408	-11.06
$\theta_{2,3}$	-		-1.5505	-10.71
$\theta_{2,4}$	-		1.0020	7.66

	MMNL			ICLV		
	μ	σ	cv	μ	σ	cv
WTP for travel time reductions (£/hr)	7.07	47.76	6.75	4.71	20.80	4.42
WTP for one less crowded journey (£)	0.69	8.41	12.13	0.53	4.60	8.74
WTP for one less delayed journey (£)	0.55	3.80	6.88	0.28	1.14	4.03
WTP for reductions in average delays (£/hr)	9.47	171.50	18.12	6.70	89.97	13.43
WTP for sms delay service (£)	0.63	3.87	6.17	0.52	1.94	3.73

Table 2: Summary results from behavioural mixing study by Hess et al. (2012)

Results for study combining MMNL with a lexicography model

	MMNL	MMNL & lexicography
Observations	13,408	13,408
Log-likelihood	-7,360.62	-7,350.33
par.	5	7
adj. ρ^2	0.2075	0.2084

				asy. t-rat.	
	est.	asy. t-rat.	est.	vs 0	vs $\frac{1}{3}$
π_{trading}	100%	-	90.45%	65.46	41.34
$\pi_{\text{lex-cost}}$	0%	-	6.92%	6.12	-23.39
$\pi_{\text{lex-time}}$	0%	-	2.63%	3.91	-45.58

WTP

	mean	median	st.dev.	cv	cv (β_{TT})	cv (β_{TC})
MMNL	90.66	28.21	261.94	2.89	5.45	17.80
MMNL & lexicography	70.61	34.60	123.07	1.74	3.47	7.98
change	-22.12%	+22.65%	-53.02%	-39.68%	-36.31%	-55.16%

Results for study combining MMNL with EBA model

	MMNL	MMNL & EBA
Observations	7,968	7,968
Log-likelihood	-5,453.85	-5,357.85
par.	11	16
adj. ρ^2	0.3757	0.3861

				asy. t-rat.	
	est.	asy. t-rat.	est.	vs 0	vs $\frac{1}{4}$
π_{MMNL}	100.00%	-	88.06%	46.60	20.14
π_{EBA}	0.00%	-	11.94%	6.32	-20.14

WTP (at fare of 40)

	MMNL			MMNL & EBA			change		
	mean	std.dev.	cv	mean	std.dev.	cv	mean	std.dev.	cv
time reduction (£/hr)	12.50	11.89	0.95	13.50	11.87	0.88	8.05%	-0.15%	-7.59%
reserved seat (£)	7.76	9.76	1.26	8.90	10.09	1.13	14.70%	3.40%	-9.85%
wifi (£)	2.58	5.96	2.31	2.98	6.80	2.28	15.79%	13.94%	-1.60%
flexible ticket (£)	3.34	6.80	2.04	3.50	7.20	2.05	4.92%	5.78%	0.82%

Table 3: Summary results from information processing study by Hess and Hensher (2010)

	Base model	Inferred IPS
adjusted ρ^2	0.4492	0.4696

	Base model *cv.*	Inferred IPS *cv.*	Reduction in *cv*	Respondents in *ignoring* class
β_{FFT}	0.98	0.75	-23.65%	15.61%
β_{SDT}	0.72	0.68	-5.47%	2.44%
β_{RC}	0.78	0.63	-19.39%	5.37%
β_{TOLL}	0.57	0.57	-0.01%	1.95%
β_{VAR}	3.97	2.76	-30.49%	29.27%