

4. Workshop report: working with repeated choice data

**Andrew Daly, Stephane Hess and
Christine Eckert**

1. INTRODUCTION

In recent years the study of data containing multiple responses from each individual has become the approach of preference for choice modelling, mainly due to increasing reliance on data from stated choice surveys. Multiple responses offer analysts the opportunity to distinguish within-respondent from between-respondent heterogeneity, to investigate behaviour by the same respondent under a range of attribute level combinations and to study behaviours that may not exist in current markets; moreover there are often cost savings in this approach compared with alternative data capture procedures.

Notwithstanding these advantages, the analysis of data containing multiple responses from individuals presents specific issues. It is necessary to make allowance for the correlation of individuals' repeated responses arising from their specific unmeasured attributes in order to avoid biases in the inferences drawn from the data. The issues involved have been known for some years and methods have been developed to deal with them, but work continues and two schools of thought can be distinguished, each offering distinct insights.

The first school takes the approach of modelling responses at the aggregate (sample) level while making allowances for the repeated choice nature of the data either explicitly or through correction approaches. The second approach is to develop a model for each individual respondent, thus bypassing the issue of separating out individual-specific effects.

The objectives of the workshop, chaired by the first author, were:

- to consider the contributions of the two schools, based on presentations by the other two authors;
- to discuss whether there was potential for cross-fertilisation between the two schools;

- to summarise the current situation; and
- to suggest lines for future work.

This workshop report summarises in the following two sections the current position in the two schools and goes on to discuss the main current issues.

2. SAMPLE-LEVEL MODELS

The first part of the workshop looked at important decisions that need to be taken on how to analyse data containing multiple choices for individual respondents, and in particular what steps need to be taken to move away from a purely cross-sectional specification of the models. Three topics were included in the presentation, and these are now looked at in turn.

2.1 Fatigue

Respondent fatigue has received substantial attention in the literature on stated preference methods, centred around the notion that long surveys may lead to fatigue, boredom or disengagement with the survey, and result in lower response quality, measured in the form of reduced model scale (i.e. less deterministic choice processes). The impact on substantive (i.e. relative) model outputs has received significantly less attention.

A review of the extensive body of work in this area is given in Hess et al. (2012), tracing research efforts back to early evidence in Bradley and Daly (1994), who, with data from a survey making use of up to sixteen binary choice tasks, showed that scale decreased significantly throughout the duration of the experiment. The evidence in Bradley and Daly (1994) has received a disproportionately large weight in the literature, with some recent work (Brownstone et al., 2012) going as far as suggesting that model results are not reliable if they fail to account for the increasing error variance. While some papers support the evidence in Bradley and Daly (1994), others do not, for example Adamowicz et al. (1998), Hanley et al. (2002) and Risa Hole (2004). Crucially, such evidence is almost always based on just a single dataset.

Hess et al. (2012) go on to present more extensive new evidence on the topic. The study made use of five datasets, with a total of nine subsamples, with data collected in different countries, making use of different choice contexts in terms of number of alternatives as well as number of attributes, and different underlying designs. Four of the surveys used computer aided personal interviews while the fifth was conducted online. Figure 4.1 shows

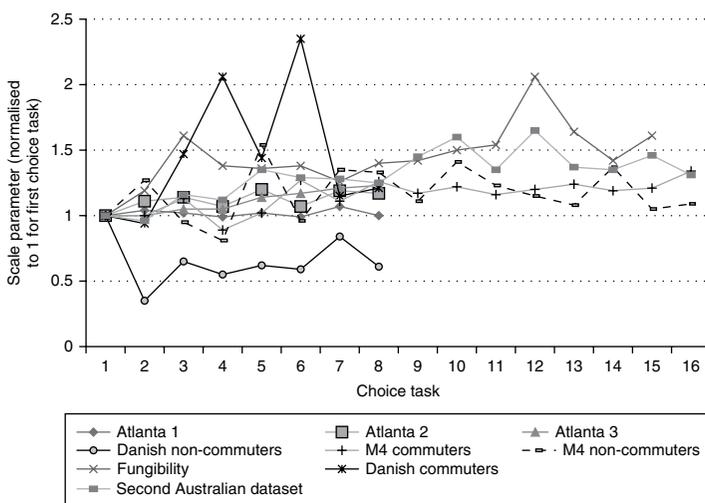


Figure 4.1 Scale differences across choice tasks: evidence from Hess et al. (2012)

the evolution of the scale parameters across choice tasks in the different datasets. It is clear that there is no overall evidence of a downward trend in scale as respondents progress through the experiments. In fact, the opposite seems to apply in some of the datasets, providing some support to the notion that response quality increases as respondents become more familiar with the choice scenario setup. The findings also confirm the mixed results seen in the literature review.

Figure 4.2 shows the changes in the valuation of travel time produced when we estimate a model repeatedly, at each step adding data from later choice tasks. The findings suggest that these changes are somewhat more pronounced than those for the scale differences, with some evidence of increasing willingness-to-pay as respondents go through the survey, which could be evidence of respondents *learning* their true sensitivities (cf. Plott, 1996) – the same could obviously be said if the opposite trend was observed.

2.2 Treatment of heterogeneity

With the increasing use of Mixed Multinomial Logit (MMNL) models for the representation of random taste heterogeneity across respondents, the additional information content of repeated choice data is generally seen as an advantage. However, important questions arise as to how to

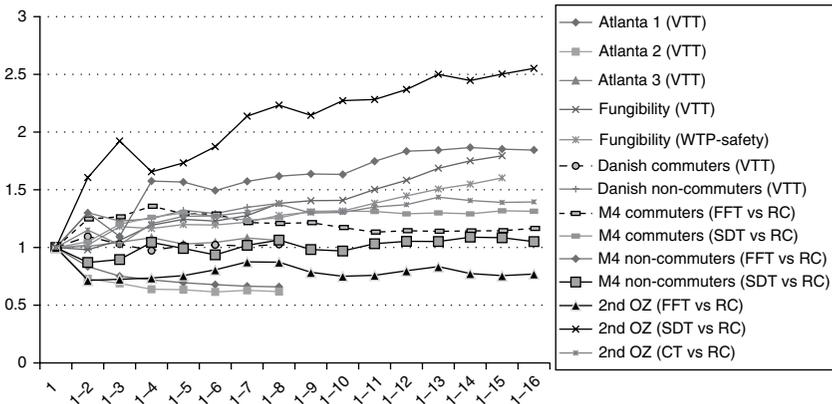


Figure 4.2 Changes in willingness-to-pay measures across choice tasks: evidence from Hess et al. (2012)

specify heterogeneity, and the complexities that may arise in estimating this heterogeneity.

Let us assume that we have N respondents and T_n choice scenarios for respondent n . In the most widely used specification, we allow tastes to vary over respondents but we assume that the tastes of each respondent are constant over choice situations. We then have the log-likelihood expressed as a function of the parameters Ω of the distribution g of the choice model parameters β :

$$LL(\Omega) = \sum_{n=1}^N \ln \left(\int_{\beta} \left[\prod_{t=1}^{T_n} (P_{n,t}(j_{n,t}|\beta)) \right] g(\beta|\Omega) d\beta \right), \tag{4.1}$$

where $P_{n,t}(j_{n,t}|\beta)$ is the probability of individual n choosing the actual observed alternative in choice occasion t , i.e. $j_{n,t}$. This log-likelihood is approximated in simulation as:

$$SLL(\Omega) = \sum_{n=1}^N \ln \left(\frac{1}{R} \sum_{r=1}^R \left[\prod_{t=1}^{T_n} (P_{n,t}(j_{n,t}|\beta_{r,n})) \right] \right), \tag{4.2}$$

where $\beta_{r,n}$ gives one of R draws from $g(\beta|\Omega)$ for respondent n .

An alternative would be to use cross-sectional estimation on repeated choice data, with:

$$LL(\Omega) = \sum_{n=1}^N \sum_{t=1}^{T_n} \ln \left(\int_{\beta} P_{n,t}(j_{n,t}|\beta) g(\beta|\Omega) d\beta \right) \tag{4.3}$$

and

$$SLL(\Omega) = \sum_{n=1}^N \sum_{t=1}^{T_n} \ln \left(\frac{1}{R} \sum_{r=1}^R P_{n,t}(j_{n,t} | \beta_{r,t,n}) \right). \quad (4.4)$$

Recent empirical evidence by Hess and Train (2011) shows that retrieving *true* patterns of heterogeneity can be considerably less precise on cross-sectional data, confirming the advantages of repeated choice data. However, while, with inter-personal variation only, Hess and Train (2011) show that maximum simulated likelihood (Eqn. 4.2) performs well, there is a large error and bias in some cases with small samples sizes, pointing to the value of larger samples (possibly larger than typically used). Using cross-sectional estimation (Eqn. 4.4) results in fairly accurate estimates with large samples, but with a loss of efficiency as large as a factor of four or five. There have also been efforts in the literature (cf. Paag et al., 2001) to capture correlation across choices by reusing the same draws in all the choice situations for the same respondent in Equation 4.4, but Hess and Train (2011) show that this does not lead to any greater ability to capture correlation across tasks.

There is also increasing interest in allowing for additional variation in sensitivities across choices for the same respondent (Bhat and Castelar, 2002; Bhat and Sardesai, 2006; Hess and Rose, 2009; Cherchi et al., 2009). Looking at a simple case, we let $\beta_{n,t} = \alpha_n + \gamma_{n,t}$, where α_n is distributed across respondents but not over choice situations for a given respondent, and $\gamma_{n,t}$ is distributed over choice situations as well as respondents. We then have:

$$LL(\Omega) = \sum_{n=1}^N \ln \left[\int_{\alpha} \left(\prod_{t=1}^{T_n} \left(\int_{\gamma} P_{n,t}(j_{n,t} | \alpha, \gamma) h(\gamma) d\gamma \right) \right) f(\alpha) d\alpha \right] \quad (4.5)$$

and

$$SLL = \sum_{n=1}^N \ln \left[\frac{1}{R} \sum_{r=1}^R \left(\prod_{t=1}^{T_n} \frac{1}{K} \sum_{k=1}^K (P_{n,t}(j_{n,t} | \alpha_{r,n}, \gamma_{k,t,n})) \right) \right]. \quad (4.6)$$

The estimation of such a model is computationally very demanding, with each draw from γ being used for a given draw from α , before averaging across γ . Hess and Train (2011) show that the recovery of such patterns of heterogeneity is possible but requires large samples and large amounts of computation. A potential approximation to Equation 4.6 is based on using averaging only at the level of a respondent (rather than the additional averaging at the choice task level), but with different draws for γ across tasks. This is implemented in BIOGEME (Bierlaire, 2003) and has been used for example by Yáñez et al. (2011) and Cherchi et al. (2009). However, results by Hess and Train (2011) show that this approximation

does not provide accurate estimates, especially for the intra-personal heterogeneity, and that Equation 4.6 must be used instead.

2.3 Correction of standard errors

It has long been recognised (see e.g. Louviere and Woodworth, 1983) that individual choices for the same respondent are not independent, and that the information content of N respondents each giving T responses is usually considerably less than that of NT respondents each giving a single response. As a result, the estimates of accuracy given by cross-sectional modelling are incorrect, i.e. such models are likely to produce biased standard errors, with the expectation being a downwards bias.

Alongside ad-hoc correction approaches, resampling methods such as Bootstrap and Jack-knife have been used to correct the standard errors from cross-sectionally estimated models (see e.g. Cirillo et al., 2000). A disadvantage of these methods is the need to determine the number of samples, and the use of different samples or different numbers of samples gives different results. An alternative approach is to use a panel specification of the ‘sandwich’ or robust estimator, which is compared with resampling by Daly and Hess (2011). The sandwich estimator is defined as:

$$S = (-H)^{-1}B(-H)^{-1} \quad (4.7)$$

where H is the Hessian matrix, i.e. the matrix of second derivatives of the log likelihood function with respect to the model parameters to be estimated, and B is the Berndt-Hall-Hall-Hausman (BHHH) matrix (Berndt et al., 1974), defined as the matrix which has in cell jk the value

$$B_{j,k} = \sum_n L_{j,n} L_{k,n} \quad (4.8)$$

where $L_{j,n}$ is the first derivative with respect to model parameter j of the contribution to the log-likelihood function from observation n . When there is no explicit retention of information between observations, H will be identical whether the data is considered as a panel or not. However, although the first derivative components used in calculating B are the same for the panel and the non-panel cases, $B_j = \sum_{n,t} L_{j,n,t} = \sum_n (\sum_t L_{j,n,t})$, the calculation procedure for S is different for the panel case and the non-panel cases, because the specification of an “observation” is different. Specifically, if we treat the data as being panel observations, we make the calculation:

$$B_{j,k} = \sum_n L_{j,n} L_{k,n} = \sum_n \left(\sum_t L_{j,n,t} \right) \left(\sum_t L_{k,n,t} \right) \quad (4.9)$$

whereas, if we treat panel data as being independent, the calculation is:

$$B_{j,k} = \sum_n \sum_t L_{j,n,t} L_{k,n,t} \quad (4.10)$$

which is clearly different and will typically give a matrix with larger components on the diagonal, so that inverting the matrix will (incorrectly) indicate smaller estimation errors.

In purely cross-sectional estimation, the sandwich estimator corrects the standard errors for general mis-specification of the model. In a panel specification, it additionally corrects the standard errors to accommodate the panel nature. The advantage over re-sampling methods is that the correction is performed alongside estimation, meaning that there is no need for additional re-sampling runs, and always gives the same solution. Daly and Hess (2011) show that the correction obtained is similar to that given by resampling methods.

In recent years, the increasing use of Mixed Logit means that many studies now directly account for the repeated choice nature of data, typically through means of random taste heterogeneity using the methods discussed in Section 2.2. In some studies Mixed Logit may be used solely with a view to capturing the repeated choice nature of the data, rather than due to a desire to e.g. allow for random heterogeneity. This leads to significant computational and identification issues, resulting in studies taking shortcuts, e.g. using just a single random coefficient, as well as numerous studies making mistakes in computing willingness-to-pay measures from random coefficients.

An alternative way of using Mixed Logit in this context is through an error components specification, aiming to capture an individual-specific effect that creates correlation across choices for the same respondent. It is clearly impossible for identification reasons to add the same error components to all J of the alternatives. As for example reported by Yáñez et al. (2011), it has become common practice to instead include the same error component in $J - 1$ of the utility functions, an approach also suggested in Bierlaire. (2003). Unfortunately, such a specification now introduces correlation across those $J - 1$ alternatives, as well as heteroskedasticity. Adding an error component to just a single alternative is no better, while randomly varying the omitted error component across respondents (cf. Yáñez et al., 2011) is similarly misguided, as it creates random variations in the correlation and/or heteroskedasticity structure across individuals. An alternative is to use an approach first put forward by Hess et al. (2008) which consists of adding *iid* (across respondents but not observations) error components to *all* of the alternatives, where this model is identified in panel data given the independent nature of the error

components. However, in theory this approach is not suitable for truly unlabelled data.

3. INDIVIDUAL-LEVEL MODELS

The purpose of this part of the workshop, which described a yet unpublished work by the third author with Jordan Louviere, Bart Frischknecht, and Tiago Ribeiro at the Centre for the Study of Choice, was to show that one can use simple methods to model single person choices by extending the Louviere et al. (2008) approach to estimation methods familiar to most academics and practitioners, such as ordinary least squares regression and weighted least squares regression. These two methods yield biased estimates of the choice probabilities, but we demonstrate that one can improve these estimates rather simply. More specifically, we show with Monte-Carlo simulations and empirical choice data sets that we can refine simple parameter estimates to produce well-behaved in-sample and out-of-sample predictions of choices.

3.1 Background

Early efforts to model individuals outside of psychology labs can be found in Beggs et al. (1981) and Chapman (1984), who show how individual-level preference rank order information on options in choice sets can be expanded to estimate individual-level models (ILMs). However, apart from those early attempts, there has been little research on estimating individual level choice models since then.

The main challenge in the estimation of ILMs is data sparsity, i.e. the estimation often fails due to an insufficient number of observations to estimate parameters reliably. One way to overcome this limitation is to expand the available full or partial ranking information to create implied choice sets and choice observations. That is, data expansion methods use full or partial preference rankings of choice options to infer chosen options in implied choice sets. Rank order expansion of full or partial rankings of choice options in choice sets to implied choices in implied choice sets was discussed by Luce and Suppes (1965), and used in practical applications by Beggs et al. (1981) and Chapman and Staelin (1982). Louviere et al. (2008) show how to use expanded observations from Best-Worst response tasks in choice experiments to estimate individual preference parameters with Weighted Least Squares (WLS) regression, which has spurred renewed interest in the estimation of ILMs. Marshall et al. (2011), for example, describe and discuss a comparison of the WLS

approach with standard Hierarchical Bayes (HB) estimation of a Mixed Logit model, and report that both predict aggregate choice shares similarly well in holdout choice tasks with out-of-sample individuals.

3.2 Proposed modelling approaches

Two ways to estimate ILMs are proposed. The first approach is motivated by the first step of the Newton-Raphson algorithm to maximize the individual's likelihood function with respect to the model parameters, whereas the second approach is motivated by differences in consumers' choice consistency which need to be taken into consideration when making out-of-sample predictions.

In the following let i represent the individual, r represent the choice set, j represent the alternative in the choice set, and k represent other alternatives in the choice set, and let the utility of alternative j for respondent i be the sum of a deterministic part depending on the attributes X of the alternatives, and a Gumbel-distributed stochastic part ε_{ij} :

$$u_{ij} = \beta_{i1}x_{1rj} + \beta_{i2}x_{2rj} + \dots + \varepsilon_{ij} = X_{rj}\beta_i + \varepsilon_{ij}$$

Further, y_{ij} is an indicator variable taking on the value 1 if the respondent chose alternative j in the r -th choice set and 0 otherwise. Thus, the dependent variable is the best (first) choice of the respondent in each choice set and it is assumed that this best choice satisfies the MNL model.

3.2.1 Approximation via Newton-Raphson algorithm

The first approach to estimate ILMs is the use of the first step of the Newton-Raphson algorithm to maximize the individual's likelihood function with respect to the model parameters. That is, we are using the first approximation for finding the zeroes of the derivative of the individual's likelihood function via the Newton-Raphson method of successive approximations (Bonnans et al., 2006).

This approach can be applied both to the 1/0 best choice indicators as well as to expanded data stemming from a full or partial ranking of the choices in the choice set. Applying this approach to these data sources yields the following ILM parameter estimates:

Ordinary least squares estimation for 1/0 choice indicators When applying the Newton-Raphson approximation to the 1/0 best choice indicators, one obtains individual-level parameter estimates by regression of the 1/0 choice indicators $Y_i = [y_{ij}]$ on $X = [X_{rj}]$ and choice set dummy variables C coefficients multiplied by a constant K_1 :

$$\hat{\beta}_i = K_1 \left(\begin{pmatrix} X \\ C \end{pmatrix}' \begin{pmatrix} X \\ C \end{pmatrix} \right)^{-1} \begin{pmatrix} X \\ C \end{pmatrix}' Y_i.$$

Ordinary least squares for expanded choice data When applying the Newton-Raphson approximation to the expanded choice data, one obtains individual-level parameter estimates by regression of the expected frequency counts $Y_i^e = [y_{ij}^e]$ from the full- or partial rankings on $X = [X_{ij}]$ and choice set dummy variables C coefficients multiplied by a constant K_2 :

$$\hat{\beta}_i = K_2 \left(\begin{pmatrix} X \\ C \end{pmatrix}' \begin{pmatrix} X \\ C \end{pmatrix} \right)^{-1} \begin{pmatrix} X \\ C \end{pmatrix}' Y_i^e.$$

3.2.2 Reweighting based on choice consistency

The second approach is motivated by differences in choice consistency across respondents: people whose choices are less well predicted by the specified model and thus have a higher mean squared error, which we interpret as being less consistent in their choices, should have predicted choice probabilities closer to $1/K$ with K being the number of alternatives in each choice set. Predicted probabilities closer to $1/K$ are produced by smaller magnitude parameters while predicted probabilities closer to 0 and 1 are produced by larger magnitude parameters. In this approach, instead of using a constant scalar across all respondents, one uses an individual rescaling factor proportional to the respondent's choice consistency. Specifically, one estimates an OLS on the 1/0 choice indicators or a weighted least squares regression with the natural log of the implied choices as the dependent variable and the implied choices as the weights (see also Louviere et al., 2008) to model a single person and uses the model to predict the observed dependent variable for each choice option in each choice set of interest. Then one calculates the associated residual mean squared error (MSE_i) for each respondent i , and multiplies all OLS or WLS model estimates for an individual by $1/MSE_i$. The revalued parameters are used to make choice predictions for each person using a Logit link function.

3.3 Simulation studies and empirical examples

We demonstrate the usefulness of the proposed ILMs using both simulated and empirical data (see also Ebling et al., 2010; Frischknecht et al., 2011a, b).

In the Monte Carlo simulation study we test the performance of the four proposed ILMs against a more commonly used "top-down" approach to

obtain individual level parameters, namely the means of the individual level posteriors of a Mixed Logit model estimated with hierarchical Bayes. In this simulation study we varied both the amount of information per individual available to the researcher (i.e. number of choice sets observed), as well as the randomness in respondents' choices (i.e. the scale of the Gumbel distribution in the Logit function). We compare the performance of the five methods using measures of individual-level parameter recovery, and individual and aggregate choice share prediction success both for in-sample as well as hold-out choices. Results from the simulations show that for small numbers of observations the ILMs perform better than the HB Mixed Logit approach. Individual and aggregate choice share prediction success is also very high for ILMs for in-sample choices, but prediction success for hold-out choices reveals that this in-sample success is partly due to overfitting particularly if only a very small number of observations per respondent is used for estimation. Still all models perform at a satisfactory level in these simulations, with the only exception of the reweighted WLS approach, which shows rather weak hold-out prediction performance.

In the empirical analyses we use the five methods discussed above to estimate model parameters and predict observed choices in four empirical data sets with varying amounts of information per respondent. The results show that individual level prediction performances are fairly similar across methods and the four data conditions. Except for the reweighted WLS estimates it appears that all ILMs perform very similarly to individual-level estimates obtained from HB Mixed Logit. However, as also observed in the simulations, the reweighted WLS estimates seem to overfit the data as indicated by high in-sample and low hold-out prediction success. Regarding aggregate choice share predictions, even though individual level HB Mixed Logit estimates best reproduce the true choice shares for all four datasets, the ILMs also produce very accurate results.

4. SUMMARY AND DISCUSSION

Considering the sample-level models, the findings from the three separate topics discussed above can be summarised succinctly as follows. Firstly, respondent fatigue seems to be less of an issue than is commonly assumed, and too much weight is still given to the evidence in Bradley and Daly (1994). Secondly, the data requirements of Mixed Logit models are substantial, and the small sample issues are arguably more important than is commonly assumed. This is especially true when one seeks to recover both between-respondent and within-respondent heterogeneity. To guarantee

recovery of the true patterns of heterogeneity, analysts should make use of the correct specification of the (simulated) log-likelihood function and avoid any shortcuts, however computationally attractive they may be. Thirdly, analysts should beware of using Mixed Logit solely with a view to capturing the repeated choice nature of their data and should instead, where possible, rely on correction approaches. This is of course only possible if the estimates from the naïve model are consistent, which is for example not the case when the correlations across choices or utility variances are not independent of the attributes.

Considering the individual-level models, we reported simulation studies as well as the empirical analyses in four different categories which indicate that ILMs can be a reliable way to estimate parameters as well as predict individual and aggregate choices particularly in the event of sparse data. The advantage of these ILMs lies hereby in the fact that only very limited technical skills are required to apply them. Whereas in the work done so far on ILMs, we focussed on parameter recovery and choice share prediction, future research should also investigate the performance of these models to recover other quantities of interest, such as Willingness-to-Pay estimates.

The workshop discussion that ensued achieved a surprising degree of consensus, with representatives of the two schools each seeking to learn from the other. For fruitful discussion it was necessary to define the objectives of particular analyses, which could be:

- understanding the reasons underlying behaviour; and/or
- forecasting the total numbers choosing each alternative or the numbers of respondents switching between specific alternatives; and/or
- measuring willingness-to-pay or consumer surplus.

Once the objectives of the work had been defined it was possible to consider the specific contributions that each school might deliver.

In a useful digression, workshop members raised the important issue of the simplicity of approaches that might be adopted. For each topic of research, an appropriate level of simplicity should be selected which need not be the most sophisticated, provided the essence of the topic was covered. In particular, the reviewers of academic papers had a duty to balance the use of the most recently published methods against the benefits of simplicity. Clearly, the use of advanced methods imposes a requirement for readers to have specific skills that may not be widely available and this issue needs to be taken into account to ensure wide impact for papers. In the present context, it was apparent that the individual-level models used considerably less sophisticated methods than the sample-level

models, but the contribution in the form of insight given by individual-level modelling should not be underestimated because of its use of more widely familiar methods; in fact, the use of simpler methods means that the skill set required is more widely available and analyses can be conducted in more studies.

It appeared that the key difference between the two approaches was that sample-level modelling was necessarily strong with respect to statistical theory, while individual-level modelling gave more insight into variations in behaviour between individuals. This distinction parallels a classical debate in statistical modelling between the need to impose a structure so that data can be understood and the need to allow the data to 'speak'. The implication is clearly that the two schools can learn from each other: sample-level modelling can profit from insight into the distributions, including correlations, of consumer values derived by individual-level modelling, as these may not be found by conventional testing, given, for example, the problems with sample size requirements noted above; individual-level modelling can profit from the theoretical progress made in sample-level modelling.

A particular issue in individual-level modelling is that the lack of a full theoretical basis means that parameter estimation errors cannot be calculated and therefore significance tests cannot be made.

However, an important point to note is that sample-level models are very computationally intensive when we try to model both within-respondent and between-respondent variations. In individual-level models, within-respondent variation can be modelled but between-respondent variation can be treated entirely empirically and this may be a productive way forward.

In summary, the crucial issue is how much the behaviour of one person can tell us about the behaviour of another. On the answer to that question depends the approach that should be taken to modelling the variation of preferences in the population.

5. DIRECTIONS FOR RESEARCH

Because of the limited time available and the interesting discussions on the two approaches, the workshop was not able to devote much time to the consideration of research needs. In particular, a systematic programme could not be discussed.

With respect to the individual-level models, an important step will be to look at the way in which willingness to pay and other trade-off ratios vary in the data, rather than the separate coefficients.

More importantly, work needs to be done on developing theory for individual-level models. This need not necessarily be very sophisticated and indeed it seemed that a consideration of older literature might yield useful methods. Modified maximum likelihood methods may offer a way forward. Consideration also needs to be given to collecting more extensive data sets to allow this type of analysis to be undertaken.

With respect to sample-level models, the main recommendation is to look for more flexibility in the approaches being used.

REFERENCES

- Adamowicz, W., Boxall, P., Williams, M., Louviere, J., 1998. Stated preference approaches for measuring passive use values: Choice experiments and contingent valuation. *American Journal of Agricultural Economics* 80 (1), 64–75.
- Beggs, S., Cardell, S., Hausman, J., 1981. Assessing the potential demand for electric cars. *Journal of Econometrics* 17 (1), 1–19.
- Berndt, E., Hall, B., Hall, R., Hausman, J., 1974. Estimation and inference in non-linear structural models. *Annals of Economic and Social Measurement* 3/4, 653–665.
- Bhat, C.R., Castelar, S., 2002. A unified mixed logit framework for modeling revealed and stated preferences: formulation and application to congestion pricing analysis in the San Francisco Bay area. *Transportation Research Part B* 36 (7), 593–616.
- Bhat, C.R., Sardesai, R., 2006. The impact of stop-making and travel time reliability on commute mode choice. *Transportation Research Part B* 40 (9), 709–730.
- Bierlaire, M., 2003. BIOGEME: a free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transport Research Conference, Monte Verità, Ascona.*
- Bonnans, J.-F., Gilbert, J., Lemaréchal, C., Sagastizábal, C., 2006. *Numerical optimization: Theoretical and practical aspects*, Universitext (second revised ed. of translation of 1997 French ed.) edition. Springer-Verlag, Berlin.
- Bradley, M.A., Daly, A., 1994. Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation* 21 (2), 167–184.
- Brownstone, D., Hansen, M., Madanat, S., 2010. Review of “Bay Area/California High-Speed Rail Ridership and Revenue Forecasting Study”. UC Berkeley Research Report UCB-ITS-RR-2010-1. Institute of Transportation Studies, University of California, Berkeley, CA.
- Chapman, R., 1984. An approach to estimating logit models of a single decision maker’s choice behavior. *Advances in Consumer Research* 11, 656–661.
- Chapman, R., Staelin, R., 1982. Exploiting rank ordered choice set data within the stochastic utility model. *Journal of Marketing Research*, 288–301.
- Cherchi, E., Cirillo, C., Ortúzar, J. de D., 2009. A mixed logit mode choice model for panel data: accounting for different correlation over time periods. Paper presented at the International Choice Modelling Conference, Harrogate.

- Cirillo, C., Lindveld, K., Daly, A., 2000. Eliminating bias due to the repeated measurements problem in SP data. In: Ortúzar, J. de D. (Ed.), *Stated Preference Modelling Techniques: PTRC Perspectives 4*. PTRC Education and Research Services Ltd, London.
- Daly, A., Hess, S., 2011. Simple methods for panel data analysis. paper presented at the 90th Annual Meeting of the Transportation Research Board, Washington, D.C.
- Ebling, C., Frischknecht, B., Louviere, J., 2010. Keep it simple: Easy ways to estimate choice models for single consumers. In: Ballantine, P., Finsterwalder, J. (Eds.), *Proceedings of the Australian and New Zealand Marketing Academy Conference 2010*. University of Canterbury, Christchurch, New Zealand, pp. 1–7.
- Frischknecht, B., Eckert, C., Louviere, J., 2011a. Estimating choice models (or equivalent) for each person in a sample. Paper presented at the Second International Choice Modeling Conference, Leeds.
- Frischknecht, B., Eckert, C., Louviere, J., 2011b. Simple Ways to Estimate Choice Models for Single Individuals. *Censoc Working Paper*, <http://www.censoc.uts.edu.au/researchoutput/wp11-006.pdf>.
- Hanley, N., Wright, R., Koop, G., 2002. Modelling recreation demand using choice experiments: climbing in Scotland. *Environmental and Resource Economics* 22, 449–466.
- Hess, S., Rose, J.M., 2009. Allowing for intra-respondent variations in coefficients estimated on repeated choice data. *Transportation Research Part B* 43 (6), 708–719.
- Hess, S., Train, K., 2011. Recovery of inter- and intra-personal heterogeneity using mixed logit models. *Transportation Research Part B* 45 (7), 973–990.
- Hess, S., Rose, J.M., Hensher, D.A., 2008. Asymmetrical preference formation in willingness to pay estimates in discrete choice models. *Transportation Research Part E* 44 (5), 847–863.
- Hess, S., Hensher, D., Daly, A., 2012. Not bored yet: revisiting respondent fatigue in stated choice surveys. *Transportation Research Part A* 46 (3), 626–644.
- Louviere, J.J., Woodworth, G., 1983. Design and analysis of simulated consumer choice and allocation experiments: A method based on aggregate data. *Journal of Marketing Research* 20, 350–367.
- Louviere, J., Street, D., Burgess, L., Wasi, N., Islam, T., Marley, A., 2008. Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information. *Journal of Choice Modelling* 1 (1), 128–163.
- Luce, D., Suppes, P., 1965. Preference, utility, and subjective probability. In: Bush, R., Galanter, E. (Eds.), *Handbook of Mathematical Psychology*. John Wiley and Sons Inc., New York, pp. 249–410.
- Marshall, D., Chan, S., Curry, J., 2011. A head-to-head comparison of the traditional (top-down) approach to choice modeling with a proposed bottom-up approach. 2010 Sawtooth Software Conference, Newport Beach, CA, 309–319.
- Paag, H., Daly, A., Rohr, C., 2001. Predicting the use of the Copenhagen Harbour Tunnel. In: Hensher, D.A. (Ed.), *Travel Behaviour Research: the Leading Edge*. Pergamon, Oxford, pp. 627–646.
- Plott, C., 1996. Rational individual behavior in markets and social choice processes: the discovered preference hypothesis. In: Arrow, K., Colombaro, C.

- E., Perleman, M., Schmidt, C. (Eds.), *Rational Foundations of Economic Behavior*. Macmillan, London.
- Risa Hole, A.R., 2004. Forecasting the demand for an employee park and ride service using commuters stated choices. *Transport Policy* 11 (4), 355–362.
- Yáñez, M.F., Cherchi, E., Heydecker, B., Ortúzar, J. de D., 2011. On the treatment of repeated observations in panel data: Efficiency of mixed logit parameter estimates. *Networks and Spatial Economics* 11 (3), 393–418.