Decomposing Unobserved Choice Variability In the Presence of Consumers' Taste Heterogeneity

Wuyang Hu
Ph.D. Candidate
Department of Rural Economy, GSB 515, University of Alberta
Edmonton, Alberta, T6G 2H1, Canada
Tel: (780)492-1518 Fax: (780)492-0268

Wiktor L. Adamowicz
Professor
Department of Rural Economy, GSB 515, University of Alberta
Edmonton, Alberta, T6G 2H1, Canada
Tel: (780)492-4603 Fax: (780)492-0268

Michele M. Veeman
Professor
Department of Rural Economy, GSB 515 University of Alberta
Edmonton, Alberta, T6G 2H1, Canada
Tel: (780)492-0270 Fax: (780)492-0268 E-mail: michele.veeman@ualberta.ca

Selected Paper prepared for presentation at the American Agricultural Economics Association Annual Meeting, Denver, Colorado, August 1-4, 2004

May 12, 2004

Decomposing Unobserved Choice Variability In the Presence of Consumers' Taste Heterogeneity

Wuyang Hu, Wiktor L. Adamowicz, and Michele M. Veeman[*,**]

Key Words: Context Effects, Heterogeneity, Random Parameters, Reference Point Effects, Stated Preference

Abstract

Heterogeneous tastes across consumers can be captured by random coefficients in a mixed logit (ML) model. However, other types of factors that may not directly affect taste could cause choices to vary, such as choice context, choice task complexity, and demographic characters. This paper jointly considers taste heterogeneity around reference-dependent attributes and other choice variability through inclusion of a scale function, based on data from a stated preference experiment for bread. Results demonstrate that modeling other sources of choice variability in addition to taste heterogeneity increases the model fit, although the improvement is not dramatic.

* Authors are, respectively, Ph.D. candidate and professors at the Department of Rural Economy, University of Alberta.

** Funding support from Genome Canada, Genome Prairie, and the Alberta Agricultural Research Institute is acknowledged.

Decomposing Unobserved Choice Variability In the Presence of Consumers' Taste Heterogeneity

Introduction

In a mixed logit (ML) model (also called a random parameter logit model), heterogeneities in consumers' tastes can be explicitly modeled through the distribution of coefficients associated with variables of interest. Valuable insights can also be obtained by decomposing the taste coefficient into several additive covariates to analyze the source of heterogeneity. Any significant taste heterogeneity that cannot be explained by covariates included in the model is usually classified as unobserved (Hensher and Greene 2003). It is not difficult to see that holding other factors constant, the more knowledge that can be collected about the heterogeneity, the better an economic model may explain and predict behaviour. This raises the question of how one can, to the fullest extent, use relevant information contained in observed consumers' choices to get a better understanding of unobserved heterogeneity.

By modeling taste heterogeneity through random coefficients in a ML specification, a researcher implicitly makes a behavioural assumption that all differences in consumers' choices are reflected by their taste variations[1]. This is likely to be an over-statement. Louviere et al. (2002) argued that unobserved heterogeneity, such as that described in a ML model, is just one of the many types of factors that cause choices to vary. These researchers used the term "variability" to account for reasons for choices to vary other than taste heterogeneity. We follow this terminology in the analysis. The variability in choices (within one individual or across individuals) may come from such factors as task complexity, the response mode, survey locations, time pressures or other aspects of the decision process (Louviere 2001, McFadden 2001). Literature in behavioural economics and psychology has made advances in recognising factors that form consumers' decisions (Payne et al. 1992 and Rabin 1998). McFadden (1998) provided a synthesis of these factors, collecting these into four overlapping categories: context effects, reference point effects, availability effects, and superstition effects. This paper constructs reference point effect measures of consumers' perceptions on price and a quality attribute for a food product. Using context effects (including the complexity effect) and demographic factors as a representation of these issues, the paper further provides a method to account for other variability in choice in addition to consumers' taste heterogeneity associated with reference point effect.

The importance of unobserved variability to the estimation of economic models has been investigated during the past decade. Variations in taste parameters cannot fully incorporate overall variability in choices (Louviere 2001). Researchers have formulated a systematic approach in order to model unobserved variability by taking advantage of the random (to the analyst) disturbance term in individuals' utility specifications (e.g., Swait and Louviere 1993; Swait and Adamowicz 2001a and b). From an assessment of relevant literature, Louviere (1996) came to the conclusion that once unobserved variability is explicitly controlled, many utility parameter differences across different studies may become negligible, leading a large proportion of the heterogeneity in taste (as seen in utility parameters) to be accounted for. In order to use unambiguous language, in the following discussion we follow the general terminology used in the literature by referring to the error term of the random utility model (RUM) as the random component and the random term in a random parameter as the stochastic component. We model overall choice variability through the random component, while explicitly accommodating taste heterogeneity through the stochastic component of random coefficients.

Modeling Choice Variability and Taste Heterogeneity

The random component in a RUM is the representation of all factors that affect individuals' choices that are known to those individuals but are unobservable from the analyst's perspective (McFadden 1974). Louviere et al. (2002) noted that since the random component coalesces all pertinent sources of unobserved variability that contribute to the differences in choices, one can achieve a valuable understanding of choice variability by explicitly modeling the covariance structure of the random component in a RUM. Hensher et al. (1999) appraised efforts along this line of research. Some researchers modeled the statistical impact of the unobserved (sometimes uncontrollable) variability in choices made under different scenarios. Examples include Adamowicz et al. (1994), Louviere et al. (1993), and Brownstone et al. (2000). These studies treat the random component purely as a scale factor to normalize estimates obtained from different data sources in order to support cross-evaluation of the validity of various results.

Employing the same modeling structure, another set of researchers consider the scale parameter derived from analysis of the random component as a behavioural vehicle by adding decision making factors to the scale parameter. Swait and Louviere (1993) analyzed disparities between parameter estimation from different choice tasks in investigating apparent differences in consumers' cognitive process. Other researchers have specified more direct models to elicit factors that contribute to the variability in choices. Hensher et al. (1999) used consumers' response to the average value of different alternatives to explain choice variability. Hensher et al. (2001) investigated the impact of the number of choice sets as a measure of the effect of complexity in consumer choice. Bradley and Daly (1994) explicitly measured the fatigue level in choice situations as an explanatory variable for choice variability. Swait and Adamowicz (2001a and b) generalized the fatigue effect and incorporated choice environment and a measure of complexity as covariates for the scale differences. Using a set of revealed preference (RP) data, Swait and Stacey (1996) modeled the impact of inter-purchase time and state-dependent on consumers' choice behaviour. Louviere and Hensher (2001) concluded that factors like consumers' demographic characteristics, choice environment and context, geographical and spatial allocations, and time factor can all be potential elements accounting for the variability of consumers' choices.

The researchers noted above have explored theoretical and empirical methods to explicitly model choice variability. They treat the random component as the overall cause of the variability in choices by modeling only the scale parameter. However, one can consider decomposing choice variability by assuming that the stochastic component explains the unobserved heterogeneity while the random component integrates the rest of the choice variability. Swait and Adamowicz (2001b) pointed out that these two effects often come hand in hand in a choice model and suggested that might be appropriate to model and interpret some forms of choice variability as taste heterogeneity and vice versa. Insofar as unobserved heterogeneity and variability are coupled, joint estimation was expected to be more efficient than independent estimation.

Swait and Bernardino (2000) outlined a potential approach for accomplishing this goal. Through a nested logit (NL) model, these authors accommodate taste heterogeneity across different alternatives in different nests while controlling the scale factors (inclusive value in a NL model) among nests. They concluded that if the differences across nests are not appropriately treated, it is likely that taste differences would seem to dominate. However, there are some limitations associated with using the NL model. First, as is well known, the underlying behavioural implication indicated by a NL model may not be consistent with utility maximization when the estimated scale parameters (inclusive values) fall beyond the range of [0,1] (McFadden 1978). An example is given by Swait et al. (2003). Swait and Bernardino (2000) also noted that more complicated specification of the NL, for example a random parameter version, may confound the interpretation of the nesting structure. Second, the NL model requires that the random component has a generalized extreme value distribution, which is restrictive, can only partially relax the IIA assumption and has difficulties in handling panel data (Train 2003 p111).

Louviere (2001) and Louviere et al. (2002) commented that since the impacts of the stochastic and random components are usually confounded, it generally requires special treatment to separate these. With the development of the ML model, taste heterogeneity can be uniquely modeled. Moreover, the ML model is flexible enough to allow any type of choice covariance structure (McFadden and Train 2000)[2] and provides a promising way to separate taste heterogeneity and choice variability. Brownstone et al. (2000) estimated a ML model with explicit consideration of the scale parameter. These authors did not address the scale factor as a manifestation of the unobserved choice variability.

Rather, they estimated the scale factor as a purely statistical nuance to enable merging of data sets from a revealed preference (RP) and stated preference (SP) survey. Breffle and Morey (2000) classified respondents into eight groups prior to estimation and compare the differences in the implied scales. However, this pre-estimation cluster analysis is not efficient for determining the appropriate magnitude of scales.

This analysis differs from the previous studies. A ML model is adopted to account for taste heterogeneity across the sample and choice variability is jointly considered by estimating a scale function. The scale function has a clear behavioural interpretation in that it is a function of choice context, choice set complexity and respondents' demographic characteristics. These factors are treated as endogenous to choice variability and are estimated jointly with the other parameters in the ML model.

Econometric Models

In a typical choice experiment, respondents are asked to state their preferences (usually indicated as the most preferred alternative) in each choice occasion they are assigned to (Swait and Adamowicz 2001a). This structure gives a string of stated choices for each individual and therefore constitutes a set of panel observations. According to random utility theory, the indirect utility of individual i choosing alternative j can be specified as:

$$U_{ijt} = \beta_i X_{ijt} + \varepsilon_{ijt}, \qquad (1)$$

where t indexes choice occasions and β_i is a vector of coefficient representing taste. β_i is allowed to be different for each individual respondent to incorporate heterogeneity associated with taste. ε_{ijt} is the random component, which can be viewed as a union of all other effects that cause choice variability (Louviere 2001). If the analyst can assume that a cumulative distribution function, in particular a Gumbel distribution, for the random component exists to a finite parameter vector, the probability of individual i choosing alternative j at the t-th choice occasion can be written as:

$$\overline{P}_{ijt} = \int \frac{\exp\left(\lambda_{ijt}\beta_i X_{ijt}\right)}{\sum_k \exp\left(\lambda_{ikt}\beta_i X_{ikt}\right)} f\left(\beta\right) d\beta = \int P_{ijt} f\left(\beta\right) d\beta \qquad (2)$$

There are several note-worthy points in this specification. First, ε_{ijt} is specified as independent across individuals. Second, $\overline{P}_{ijt}$ represents the choice probability under the mixed logit model. As the second equality shows, $\overline{P}_{ijt}$ is the conventional conditional logit probability P_{ijt} integrated over the density of the random parameters. Third, $f\left(\beta\right)$ is the probability density function for random coefficients. To keep our notation clean, β_i is used in equation (2), however not all coefficients in equation (1) need to be specified as random coefficients. $f\left(\beta\right)$ gives the density of those that are random. Also, random coefficients can be assumed to distribute independently or may have a joint multivariate distribution and in this latter case, $f\left(\beta\right)$ can be generalized to represent the joint density function of the random coefficients. Fourth, λ_{ijt} is the scale parameter that accounts for the overall unobserved variability of choices and is the inverse of the standard deviation of the model. A general expression to represent the scale parameter is $\lambda_{ijt}\left(Z_{ijt}\right)$, which indicates that this is a function rather than a single parameter, and may vary across alternatives, survey respondents, or choice situations.

For β_i, in addition to the mean and standard deviation of the stochastic component, one can specify covariates to model the shift of the mean in response to various explanatory variables. In general, one can define:

$$\beta_i = b_0 + b_q Y_{iq} + b_e e_i \, , \qquad (3)$$

where Y_{iq} is a vector of individual-specific variables and $e_i \sim N(0,1)$. Similarly, we desire to explicitly model the source of unobserved variability in choices. Therefore $\lambda_{ijt}(Z_{ijt})$ can be defined as follows:

$$\lambda_{it} = \exp(\gamma_w Z_{itw}) = \frac{1}{\sigma_{it}}, \tag{4}$$

where Z_{itw} is a vector of variables representing the differences across choice sets and γ_w indicates the corresponding scale function parameters. Note that the modeling of the scale parameter is simplified by letting Z_{itw} vary only across choice situations and individuals, but not over alternatives. Variables Z_{itw} enter the scale in their exponential form to guarantee non-negative estimates of model variance, as λ_{it} is the inverse of the standard deviation. Equation (3) and (4) can be substituted back into (2) to complete the probability expression.

Obviously, the integral in equation (2) does not have a closed form but can be evaluated by simulation. Conditional on the d-th random draw of β_{id}, the simulated probability can be written as:

$$\widetilde{P}_{ijt} = \frac{1}{D}\sum_{d=1}^{D} \frac{\exp(\lambda_{it}\beta_{id} X_{ijt})}{\sum_k \exp(\lambda_{it}\beta_{id} X_{ikt})} \tag{5}$$

The corresponding simulated log-likelihood function is:

$$SLL = \sum_{i=1}^{N} \sum_{t}^{n_i} c_{ijt} \ln(\widetilde{P}_{ijt}), \tag{6}$$

where n_i denotes the number of choices individual i makes in the survey, and c_{ijt} =1 only when alternative j is chosen by individual i in the t-th occasion. Although *SLL* is a biased estimator of the true likelihood, it is efficient when the number of draws is large enough (Lee 1992; Hajivassiliou and Ruud 1994).

Data

The data employed for this analysis are obtained from a Canada-wide survey conducted in 2003. The survey is a stated preference choice experiment on pre-packaged sliced bread with possible genetically modified (GM) ingredients. Three types of GM labeling contexts applied in the survey: (a) mandatory labeling, where all products that contain GM ingredients must be labeled, requiring positive statements (e.g., "this product contains GM ingredients"); (b) voluntary labeling, where producers can choose whether to label their product; in this case producers only have incentives to label their products when these products do not contain GM ingredients, using negative statements (e.g., "this product does not contain GM ingredients") (Huffman et al. 2002); (c) any type of GM labeling, representing a situation when no specific labeling requirement is in practice. Each respondent was randomly assigned to a labeling scheme with eight replications of choice situations. Each choice situation contains three alternatives with the first two described by attributes while the third is a "buy none" option. The attributes and levels used in the design are presented in table 1.

Reference point effects derived from prospect theory (Kahneman and Tversky 1979) are captured by four dummy variables representing gains and losses associated with price and whether GM ingredients are present. Respondents' perceived price and belief on whether GM ingredients are present in their most-often purchased bread products are obtained before the choice exercise. These perceptions are their reference points. In the choice experiment, alternatives may have other attribute levels than the reference points, and these differences generate reference point effects. Specifically, we formulate measures of reference point effects following Hardie et al. (1993): For the price variable, denote the perceived price as P_r. If the price of the bread in an alternative is represented by P_a, then: if $P_a \leq P_r$, PGain (ie price gain) = 1 and PLoss (ie price loss) = 0; if $P_a > P_r$, PGain = 0 and PLoss = 1. For the third alternative in a choice set, PGain = PLoss = 0. Similarly, if GM ingredients are perceived to be present, $GM_r = 1$ otherwise $GM_r = 0$. If the actual GM contents of an alternative can be represented by GM_a, then $GM_a = 1$ means the product contains GM ingredients, otherwise $GM_a = 0$. We define NoGMGain = 1 and GMLoss = 0 when $GM_r - GM_a = 1$ and when $GM_a - GM_r = 1$, GMLoss = 1 and NoGMGain = 0. Again, for the third alternative in each choice occasion, NoGMGain = GMLoss = 0 for the GM attribute.

The price gain and loss variables are then interacted with the price variable. The GM gain dummy variable is interacted with NOGMO, and the GM loss dummy variable is interacted with GMO. For Y_{iq} in equation (3), we incorporate three variables: respondents' age, income, and GM knowledge level. GM knowledge level is a dummy variable which equals one if a respondent correctly answers each of five binary knowledge questions. Further variables considered to be important in explaining heterogeneity in Y_{iq} could be added. However, an excessively long list of covariates will unnecessarily complicate the estimation and lead to unstable results (Breffle and Morey 2000). After several trials, we finalized the specification with the current set of variables.

In terms of parameterizing vector Z_{itw} in equation (4), three types of variables are included. First, since we assess model choice variability that is rooted in the survey context and the survey includes three contexts defined by different types of GM labelling environments, variables capturing these labelling contexts are chosen for the scale function. Two dummy variables representing each of the mandatory and voluntary labelling contexts are selected. Second, following Hensher et al. (2001) the choice task number (1-8) is included in Z_{itw} to approximate task complexity (see Swait and Adamowicz (2001a and b) for a review of alternative measures of task complexity). A general hypothesis is that as the task overall becomes increasingly complex, as indicated by the task number moving from 1 to 8, consumers' preferences are likely to become less consistent (Swait and Admowicz 2001b). Third, since all relevant information/factors affecting choices must be processed by respondents before any actual choices are made (McFadden 2001), different individuals, characterized partly by their demographic characteristics, are likely to vary systematically in their different manners of processing information and making choices (de Palma et al. 1994; Hensher et al. 1999). Two demographic variables are therefore included: gender and college participation experience[3]. Descriptive statistics of relevant variables used in the analysis are summarized in table 2.

Estimation and Results

Questions that must be answered in estimation of a ML model are which coefficients in the utility function should be randomized and what type of distribution should be utilized to describe the stochastic component. McFadden and Train (2000) developed a test to help identify which variable should be associated with a random coefficient. However, the power of the test is low, and the critical value is difficult to retrieve. In regard to the type of distribution that should be used for random coefficients, the choice is likely to be a judgement that is dependent on the particular problem and the covariance structure that researchers want to establish for the overall random component of the model. In this study, we assume normally distributed random coefficients for the four reference point effect measures. The estimation results are presented in table 3. It is known that in order to identify the scale function, a base case scenario must be located (Swait and Louviere 1993) enabling other scale measures to be compared with this. The base case will have zero values for all related covariates, and as determined by equation (4), the scale parameter for the base case is one.

The model is significant and improves the fit slightly over a model version without the specification for the scale function (-3066.202 versus -3070.895 in *LL* function). All orthogonally designed bread attribute variables are highly significant. General implications of the estimates are: the higher the price, the less attractive a loaf of bread is to consumers; consumers prefer to buy bread rather than not, and in particular, they prefer nationally branded multigrain and whole wheat bread over white or partially whole wheat bread; the presence of GM ingredients is associated with large utility loss and the absence of GM ingredients incurs utility gain.

For random coefficients, knowledge about GM is not significant in any of the random coefficients. However, respondents' age has a positive impact on the negative price loss coefficient, indicating that older consumers are predicted to better cope with price losses than relatively younger consumers. Higher family incomes alleviate consumers' loss of utility associated with a price loss; higher family incomes also increase the utility for consumers when GM ingredients are not present in bread. All standard deviation estimates for the four random coefficients are significant, indicating that there is still a significant amount of heterogeneity that cannot be explained by the constant (b_0) and the three covariates used in the specification of the β_i estimates.

In the specification of the scale function, except for the dummy variable indicating a mandatory labelling scenario, all other parameters are at least marginally significant. Since the scale parameter is the inverse of the standard error of the model and the exponential function given in equation (4) is monotonically increasing in its argument, the larger is a parameter in the scale function, the smaller is the implied model standard deviation (variance). The coefficient of the voluntary labelling scenario dummy variable is positive in the scale function. This implies one of two things: first, compared with a situation in which no particular labelling policies are applied for GM bread products, the variance between consumers' choices is smaller in a voluntary labelling regime; or second, our model explains the unobserved variability in the voluntary labelling scenario better. In other words, we can be more confident in predicting consumers' behaviour in the voluntary labelling scenario than in the situations where there is no labelling or in the mandatory labelling scenario[4].

The preceding result can be justified. When no particular labelling rules apply, the market may contain different products with all possible labels (positive, negative, or a mixture of these two). In this situation, more products may appear to be different, and consumers may be confused by these labels. This can be interpreted as many different products increasing the complexity of choice tasks, resulting in less consistent choices (Mazzotta and Opaluch 1995). For the mandatory labelling scenario, the presence of GM ingredients can be a cause of uncertainties in terms of human health, the environment or other concerns. A consumer may not be likely to obtain sufficient information to resolve these uncertainties from a label that lists some ingredients with GM content (i.e., with stochastic qualities). Kinsey (1999) argues that from a consumers' perspective, a positive GM label statement in a mandatory labelling scenario may be viewed to work no better than no label at all. Swait and Adamowicz (2001b) pointed out that consumers' uncertainties can lead to inconsistent choices and to a larger variance in utility functions. In our case, holding other factors fixed, consumers' choices in a mandatory labelling environment are just as "noisy" as in the scenario of no labelling requirements. However,, in a voluntary labelling scenario, consumers are given definite information in terms of negative statement of GM ingredients (this product contains no GM ingredients). Thus consumers may be more certain about the quality of their chosen alternative which may indicate less volatility in terms of information presentation. These factors may significantly lower variation among choices in the voluntary labelling scenario, as indicated by the variable *Volun* in the scale function.

In this study, the variable "task" provides an approximation of the impact of survey complexity on consumer choices as the task number. Its effect is negative and significant in the scale function, indicating that as the choices proceeded (with the overall task becoming more complex), respondents started to make more inconsistent choices. In this situation a consumer may tend to pick an alternative simply to finish the choice task without effort to select the alternative that best represents their preferences. An extreme situation of the "pick any" effect could apply if tasks are so complex that consumers make random choices (Louviere 2001). Although the results from our model suggest the existence of the pick-any effect, it is possible that all behavioural processes discussed under the complexity effect may coexist[5]. Thus, a more precise interpretation is that the pick-any effect was found to be dominant. Distinguishing the mechanism that actually functioned behind the observed result is worthy of further research[6].

Gender was highly significant in shifting respondents' choices. If other factors are held constant, males tended to make more variable choices than females and the model predicts female consumers' choices better than that for males. Consumers' education level is marginally significant in explaining choice variability. Generally speaking, more consistent choices were made by consumers with post-secondary education.

Table 3 reports the estimates of the b and γ parameters in equation (3) and (4). These are not the actual random coefficients or the scale parameter. Simulations are used to obtain the mean and standard error estimates associated with the estimates of β and λ. To take the covariance between estimated parameters in the model into account, given the mean Y_{iq} variables, a vector of corresponding $b's$ is drawn from the multivariate normal distribution $MN(\theta, \Sigma_\theta)$, where θ is a vector of the means of the estimated parameters and Σ_θ is the correlation matrix between the parameters. Table 4 reports the simulated mean and standard deviations associated with the four random coefficients after 2000 replications. The simulated coefficient associated with the effect of price gain is not significant while the simulated coefficient for price loss is negative and significant. This verifies the asymmetric price reference point effect. A similar result was not observed for the GM attribute. The simulated coefficient for gain with the no-GM attribute (when a product was labelled as no-GM while respondents believed that the product contained GM ingredients before the choice) is not significant. However, the simulated coefficient for loss of the GM attribute (when a product was labelled as GM while respondents believed that the product did not contain GM ingredients) is significant but positive.

The reference point findings from this study do not necessarily disprove the existence of reference point effects for GM content or absence. Unlike the price factor where it can be anticipated that holding other factors constant, the higher is price, the less attractive is the product, whether GM ingredients are present or absent is a credence attribute with uncertain properties. Consumer responses to this attribute may not follow a standard pattern and the definition of gain or loss associated with GM ingredients may not be viewed similarly by all consumers. The study indicates that an average consumer sees the unexpected appearance of GM ingredients in bread products as desirable. This may arise if the presence of GM ingredients increases the variety of products that consumers can choose from, to the point that utility is increased. Diversity in attitudes to GM content warrants further assessment of possible GM-reference point effects.

For the scale parameter, a similar simulation approach can be conducted. Before the simulation, we classified the effects from variables in the Z_{itw} vector into six groups: a) mandatory labelling with low task effort (variable "task" reflects the number of product alternatives = 2.5, which is the average in the first four tasks); b) voluntary labelling with low task effort; c) no labelling requirement with low task effort; d) mandatory labelling with high task effort (here variable "task" = 5.5, which is the average for the last four tasks); e) voluntary labelling with high task effort; and f) no labelling requirement with high task effort. We define a representative consumer as a male consumer with some post-secondary education. Due to the exponential function used to define the scale parameter, a draw of 1 from the multivariate normal distribution can cause the scale parameter to be unreasonably large (e.g., in group f, $\exp(5.5) = 148$). Therefore, the parameter for the variable "task" is fixed at the mean and parameters for “mand”, “volun”, age, and education variables are drawn from the multivariate normal distribution. Table 5 reports these six scale parameters. Since the parameters for “task” are fixed, standard deviations for group c and f are not available.

All scale parameters are significant and three of the four testable scale parameters are significantly different from one. These scale parameter estimates can be compared with the base case (where scale is equal to one) or be interpreted relative to each other. Within either the low effort or the high effort scenario, voluntary labelling is associated with the lowest implied model standard deviation; i.e., the estimation in the voluntary labelling scenario is subject to the least choice variability. In the situation of no labelling requirements, variances among choices are noticeably larger than in the two specified labelling situations. Finally, within each labelling scenario (including no label), low task complexity (low effort) is associated with lower variances in choices than higher task complexity. These results are consistent with the earlier interpretation based on the signs of individual covariates in the scale function.

Discussion and Conclusions

Usually, in a mixed logit model, the variability in choices that cannot be captured by taste heterogeneity is treated as unobserved heterogeneity. This study demonstrates that in addition to taste heterogeneity, other factors may also cause variability in choices. We demonstrate the use of reference point effects and the heterogeneity in consumers’ evaluation that is associated with them. We also show that unobserved heterogeneity can be further explained by explicitly modeling other sources of variability in choices through the scale function. Our results are not completely consistent with the findings of some previous studies that analyzed the relationship between choice variability and specific taste heterogeneity and found that explicit consideration of the scale parameter greatly reduced or even eliminated the degree of heterogeneity (Kamakura et al. 1996; Swait and Bernardino 2000; Hensher et al. 1999; Louviere et al. 2002). However, our findings support the general conclusion that by simply estimating heterogeneity through coefficients or by focus only on variability through the random component, a researcher may miss the effects of some factors that can otherwise be discovered by jointly modeling both sources (Swait et al. 2002).

This study shows that unobserved heterogeneity can be separated from unobserved variability and that both may have significant impacts on choice predictions. However, as there are numerous factors that may affect consumers' choice behaviour and these often overlap with each other (McFadden 1999), it is difficult to distinguish and model all effects. In this analysis, several representative effects are investigated and are found to improve model fit.

Finally, it is noteworthy that the method proposed and applied in this study is not limited to a ML model. Other flexible models may be adopted depending on the research goal. These could include a mixed latent class model (if the purpose is to classify consumers rather than to know the preferences of the entire population or each specific individual); a mixed probit model (which may facilitate the estimation process) or a pure probit model (which may allow direct parameterisation of the model's covariance structure). These all provide grounds for future research effort.

References

Adamowicz, W., J. Louviere, and M. Williams (1994) "Combining Stated and Revealed Preference Methods for Valuing Environmental Amenities" *Journal of Environmental Economics and Management* 26, 271-292.

Ben-Akiva, M., D. McFadden, K. Train, J. Walker, C. Bhat, M. Bierlaire, D. Bolduc, A. Boersch-Supan, D. Brownstone, D. Bunch, A. Daly, A. De Palma, D. Gopinath, A. Karlstrom, M. A. Munizaga (2002) "Hybrid Choice Models: Progress and Challenges" *Marketing Letters*, 13(3): 163-175.

Bradley, M. and A. J. Daly (1994) "Use of Logit Scaling Approach to Test Rank-Order and Fatigue Effects in the Stated Preference Data" *Transportation* 21(2): 167-184.

Breffle, W. S., and E. R. Morey (2000) "Investigating Preference Heterogeneity in a Repeated Discrete-Choice Recreation Demand Model of Atlantic Salmon Fishing" *Marine Resource Economics*, 15: 1-20.

Brownstone, D., D. S. Bunch and K. Train (2000) "Joint Mixed Logit Models of Stated and Revealed Preferences for Alternatives-Fuel Vehicles" *Transportation Research B*, 34(5): 315-338.

Brownstone, D. and K. Train (1999) "Forecasting New Product Penetration with Flexible Substitution Patterns" *Journal of Econometrics* 89: 109-129.

De Palma, A., G. Myers, and Y. Papageorgiou (1994) "Rational Choice under an Imperfect Ability to Choose" *American Economic Review*, 84(3), 419-440.

Dhar, R. (1997) "Consumer Preference for a No-Choice Option" *Journal of Consumer Research*, 24: 215-231.

Foster, V. and S. Mourato (2002) "Testing for Consistency in Contingent Ranking Experiments" *Journal of Environmental Economics and Management*, 44: 309-328.

Hajivassiliou, V. and P. Ruud (1994) Classical Estimation Methods for LDV Models Using Simulation, Chapter 40, Engle, R. and D. McFadden (ed.) *Handbook of Econometrics, Vol. IV*, Elsevier, New York.

Hampton, J. (1998) "The Between-Subjects Experiment" Chapter 2, p15-38, in *Laboratory Psychology: A Beginner's Guide*, Nunn, J. (eds.), Hove, East Susses, UK: Psychology Press.

Hardie, B. G. S., E. J. Johnson and P. S. Fader (1993) "Modelling Loss Aversion and Reference Dependent Effects on Brand Choice" *Marketing Science* Vol.12(4): 378-394.

Hensher, D. and W. H. Greene (2003) "The Mixed Logit Model: The State of Practice" *Transportation Research B*, 36(2): 133-176.

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

Hensher, D., P. Stopher, and J. Louviere (2001) "An Exploratory Analysis of the Effects of Numbers of Choice Sets in Designs Choice Experiment: An Airline Choice Application" *Journal of Air Transport Management* 7: 373-379.

Huffman, W., M. Rousu, J. Shogren, A. Tegene (2002) "Should the United States Regulate Mandatory Labeling for Genetically Modified Foods?" Working Paper # 02013, Department of Economics, Iowa State University.

Kamakura, W., B. Kim, and J. Lee (1996) "Modeling Preference and Structural Heterogeneity in Consumer Choice" *Marketing Science*, 15(2): 152-172.

Kahneman, D. and A. Tversky (1979), "Prospect Theory: An Analysis of Decision under Risk" *Econometrica* Vol. 47, 263-291.

Kinsey, J. D. (1999) "Genetically Modified Food and Fibre: A Speedy Penetration or a False Start?" *Cereal Foods World* 44: 487-489.

Lee, L. (1992) "On Efficiency of Methods of Simulated Moments and Maximum Simulated Likelihood Estimation of Discrete Response Models" *Econometrica*, 8: 518-552.

Louviere, J. (1996) "Combining Revealed and Stated Preference Data: The Rescaling Revolution" Paper presented at the Association of Environmental and Resource Economists Conference, Lake Tahoe.

Louviere, J. (2001) "What if Consumer Experiments Impact Variances as well as Means? Response Variability as a Behavioural Phenomenon" *Journal of Consumer Research*, 28:3, 506-511.

Louviere, J., R. Carson, A. Ainslie, J. R. Deshazo, T. Cameron, D. Hensher, R. Kohn, and T. Marley (2002) "Dissecting the Random Component of Utility" *Marketing Letters* 13(3): 177-193.

Louviere, J., M. Fox, and W. Moore (1993) "Cross-Task Validity Comparisions of Stated Preference Choice Models" *Marketing Letters* 4(3): 205-213.

Louviere, J., and D. Hensher (2001) Combining Sources of Preference Data, p125-144, in D. Hensher (ed.) *Travel Behaviour Research: The Leading Edge*, Pergamon Press, Oxford.

Mazzotta, M. J. and J. J. Opaluch (1995) “Decision Making When Choices are Complex: A Test of Heiner’s Hypothesis” *Land Economics*, 71:4, 500-515.

McFadden, D. (1974) Conditional Logit Analysis of Qualitative Choice Behaviour, p105-142, in P. Zarembka (ed.) *Frontiers in Econometrics*, Academic Press, New York.

McFadden D. (1978) Modeling the Choice of Residential Location, p75-96, in A. Karlqvist, L. Lundqvist, F. Snickars and j. Weibull, (ed.) *Spatial Interaction Theory and Planning Models*, North-Holland, Amsterdam.

McFadden, D. (1998) “Rationality for Economists” *Journal of Risk and Uncertainty*, Vol.19 No.1-3, 73-105.

McFadden, D. (1999) Econometric 244 Class Notes, Department of Economics, University of California at Berkeley.

McFadden D. (2001) “Economic Choices” *The American Economic Review*, 91(3), 351-378.

McFadden D. and K. Train (2000) “Mixed MNL Models for Discrete Response” *Journal of Econometrics*, 15(5): 447-470.

Payne, J., J. Bettman and E. Johnson (1992) “Behavioural Decision Research: A Constructive Process Perspective” *Annual Review of Psychology*, Vol.43, 87-131.

Rabin, M. (1998) “Psychology and Economics” *Journal of Economic Literature*, Vol.36, 11-46.

Swait, J. and W. Adamowicz (2001a) “Choice Complexity and Decision Strategy Selection” *Journal of Consumer Research*, 28(1):135-148.

Swait, J. and W. Adamowicz, (2001b) “Incorporating the Effect of Choice Environment and Complexity into Random Utility Models” *Organizational Behavior and Human Decision Processes*, 86(2):141-167.

Swait, J., W. Adamowicz, M. Hanemann, A. Diederich, J. Krosnick, D. Layton, W. Provencher, D. Schkade, and R. Tourangeau (2002) “Context Dependence and Aggregation in Disaggregate Choice Analysis” *Marketing Letters*, 13(3): 195-205.

Swait, J., W. Adamowicz, and M. Van Bueren (2003) “Choice and Temporal Welfare Impacts: Dynamic GEV Discrete Choice Models” *Working Paper*, Department of Rural Economy, University of Alberta.

Swait, J. and A. Bernardino (2000) “Distinguishing Taste Variation from Error Structure in Discrete Choice Data” *Transportation Research B*, 34(1):1-15.

Swait, J. and J. Louviere (1993) "The Role of the Scale Parameter in the Estimation and Use of Multinomial Logit Model" *Journal of Marketing Research*, 30: 305-314.

Swait, J. and E. C. Stacey (1996) "Consumer Brand Assessment and Assessment Confidence in Models of Longitudinal Choice Behaviour" Presented at the 1996 INFORMS Marketing Science Conference, March 7-10, 1996, Gainesville, Florida.

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

Table 1. Attributes and Levels in Choice Experiment Design

	Level 1	Level 2	Level 3	Level 4
Brand Name	store brand	national brand	-	-
Type of Flour	white	partial (60%) whole wheat	100% whole wheat	multi-grain
Price (CND)	$0.99	$1.49	$2.49	$3.49
GM or not	GM ingredients present	GM ingredients absent	not specified	-

Table 2. Model Variable Descriptions

Variable Name	Variable Description
Price	A continuous variable representing actual price
Buyno	Alternative specific constant representing the utility associated with choosing to buy none of the bread
Storeb	=1 if the bread has a store brand, otherwise=-1
White	=1 if the bread is white bread; =-1 if it is multi-grain and =0 otherwise
Partial	=1 if the bread is partial whole wheat; =-1 if it is multigrain and =0 otherwise
Whole	=1 if the bread is whole wheat; =-1 if it is multigrain and =0 otherwise
GMO	=1 if the bread has GM ingredients; =-1 if not specified or labelled as containing no GM ingredients, and =0 if it is the third option - "none of the above"
NOGMO	=1 if the bread does not contain GM ingredients; =-1 if not specified or labelled as containing GM ingredients, and =0 if it is the third option - "none of the above"
PG	price gain dummy varialbe interacted with Price
PL	price loss dummy variable interacted with Price
GMG	GM gain dummy variable interacted with variable GMO
GML	GM loss dummy variable interacted with variable NOGMO
Age	A continuous variable representing respondents' age
Income	A continuous variable representing respondents' income
Know	A dummy variable representing whether a respondent has answered all five GM knowledge questions correctly
Mand	=1 if the context is a mandatory labelling and =0 otherwise
Volun	=1 if the context is a voluntary labelling and =0 otherwise
Task	A continuous variable representing the task number
Male	=1 if the respondent is a male and =0 otherwise
College	=1 if the respondent received some post-secondary education and =0 otherwise

Table 3. Estimation Results

	Without Scale Function		With Scale Function	
	Coefficient	Std. Error	Coefficient	Std. Error
Reference Point Effect Measures with Random Parameters				
Constant in PG	0.4423	0.3034	0.2948	0.2974
PG-Age	-0.8242	0.5514	-0.6548	0.5512
PG-Income	-0.1768	0.2737	-0.0915	0.2854
PG-Know	-0.0194	0.1565	-0.0483	0.1663
Std. Dev. PG	0.7170***	0.1086	0.7479***	0.1092
Constant in PL	-0.6675***	0.1481	-0.7014***	0.1415
PL-Age	0.5140**	0.2532	0.5629**	0.2585
PL-Income	0.2827**	0.1363	0.3338**	0.1371
PL-Know	0.0837	0.0755	-0.0954	0.0795
Std. Dev. PL	0.5294***	0.0438	0.5388***	0.0444
Constant in GMG	-0.6010	0.6877	-0.5505	0.6081
GMG-Age	-0.1795	1.4231	-0.1901	1.3437
GMG-Income	1.3248**	0.5143	1.2683***	0.4790
GMG-Know	0.0136	0.3494	0.0457	0.3498
Std. Dev. GMG	0.8150***	0.2210	0.7874***	0.2295
Constant in GML	0.6963	0.6177	0.6166	0.5540
GML-Age	-0.9763	1.1069	-0.9066	1.0819
GML-Income	0.6612	0.6246	0.7624	0.6451
GML-Know	-0.2548	0.3360	-0.2714	0.3535
Std. Dev. GML	1.3415***	0.1714	1.2935***	0.1736
Attribute Variables with Fixed Coefficient				
Price	-0.5449***	0.0659	-0.6114***	0.0841
Buyno	-2.0962***	0.1391	-2.1011***	0.2063
Storeb	-0.1262***	0.0302	-0.1167***	0.0325
White	-0.4051***	0.0355	-0.4289***	0.0445
Partial	-0.2525***	0.0487	-0.2632***	0.0535
Whole	0.2195***	0.0407	0.2053***	0.0428
GMO	-0.6144***	0.0636	-0.7505***	0.0672
NOGMO	0.1972***	0.0582	0.1553***	0.0527
MGMO	-0.3258**	0.1437	-	-
VNOGMO	-0.1410	0.1401	-	-
Scale Function Parameters				
Mand	-	-	0.0645	0.0602
Volun	-	-	0.0940**	0.0484
Task	-	-	-0.0265**	0.0126
Male	-	-	-0.0768***	0.0215
College	-	-	0.0416*	0.0250
Adj. R^2	*0.161*		*0.168*	
LL	*-3070.900*		*-3066.202*	

*, **, *** indicates significant at the 10%, 5%, and 1% significance level respectively.

Table 4. Simulated Random Coefficients

	Coefficient	Std. Dev.
Price Gain	0.0155	0.1086
Price Loss	-0.2688*	0.0577
No GM Gain	0.1044	0.1709
GM Loss	0.4826*	0.1718

* Significant at the 5% significance level

Table 5. Overall Simulated Relative Scale Parameters

Groups	Mean Scale Parameter	Std. Dev.	Implied Std. Dev. (1/λ)
Mandatory/Low effort	1.1040*	0.0785	0.9058
Voluntary/Low effort	1.1371*[#]	0.0688	0.8794
No Labelling/Low effort	0.9035	-	1.1069
Mandatory/High effort	1.1954*[#]	0.0850	0.8365
Voluntary/High effort	1.2312*[#]	0.0745	0.8122
No Labelling/High effort	0.8345	-	1.1983

* Significant at the 5% significance level.

[#] Significantly different from 1 at the 5% significance level.

[1] The ML model can also be specified as an error component model by collecting the stochastic portion of random coefficients into the error term of the indirect utility function (Brownstone and Train 1999). Sharing the same computational property as the random coefficients ML model, the error component model can be used to model the variance structure of choices. However, this approach is only adopted to explicitly model a specific heteroskedastic substitution pattern (Train 2003, p160).

[2] McFadden and Train (2000) demonstrated that this can be achieved by specifying appropriate distributions for the random coefficients in a model. However, they continued the argument with the note that such a generalization of the ML model is most likely only feasible in theory. In practice, researchers usually choose distributions that are relatively convenient to work with, which in turn prohibits to some extent the model's ability to capture *any* arbitrary type of covariance structure.

[3] Other demographic variables can also be used. In theory, variables used to explain heterogeneity in taste parameters can also be included. However, treating the same variable as both the taste and the scale (context) covariate will make it impossible to interpret its effect.

[4] The variances of choices under the no labeling requirement and mandatory labeling scenarios are not statistically different, indicated by the non-significant coefficient for the mandatory labeling dummy variable in the scale function.

[5] It is possible to argue that choices may be more variable at the beginning of the task but more consistent towards the end. Due to learning effect, as choices are made consecutively, consumers will start to obtain more experience in choices and learn from their own previous choices (Hampton 1998), leading to stable choices and less overall choice variability (variance). An opposing possibility is that when the tasks become complex, consumers may start to feel tired of or less interested and may simplify their choices by always selecting the alternative that is the easiest to evaluate ("buy none" option in this paper). This effect is generalized as the simplifying heuristic in the behavior literature (Dhar 1997; Foster and Mourato 2002). Dhar (1997) termed this as the "status quo" bias, which will lead to a smaller choice variance.

[6] The effect of a squared term of the complexity measurement was also investigated. However, the model failed to converge.