

Modeling the Effects of Including/Excluding Attributes in Choice Experiments on Systematic and Random Components

Towhidul Islam
Associate Professor
Department of Marketing and Consumer Studies
University of Guelph, Canada
Email: islam@uoguelph.ca

Jordan J. Louviere*
Professor of Marketing
Co-Director, Centre for the Study of Choice (CenSoC)
Faculty of Business, University of Technology, Sydney
Email: jordan.louviere@uts.edu.au

and

Paul F. Burke
Lecturer
School of Marketing
Faculty of Business, University of Technology, Sydney
Email: paul.burke@uts.edu.au

THIS PAPER IS TO BE CITED AS:

Islam, Towhidul, Jordan J. Louviere, and Paul F. Burke. "Modeling the effects of including/excluding attributes in choice experiments on systematic and random components." *International Journal of Research in Marketing* 24, no. 4 (2007): 289-300.
DOI: <http://dx.doi.org/10.1016/j.ijresmar.2007.04.002>

The authors gratefully acknowledge funding provided by the Social Sciences and Humanities Research Council (SSHRC) of Canada Grant # 046288 and the University of Guelph Internal SSHRC Research Grant # 045415. We also acknowledge helpful comments from Kenneth Train on an earlier draft of this paper.

Modeling the Effects of Including/Excluding Attributes in Choice Experiments on Systematic and Random Components

Abstract

This paper examines the impact of attribute presence/absence in choice experiments using covariance heterogeneity models and random coefficient models. Results show that attribute presence/absence impacts both mean utility (systematic components) and choice variability (random components). Biased mean effects can occur by not accounting for choice variability. Further, even if one accounts for choice variability, attribute effects can differ because of attribute presence/absence. Managers who use choice experiments to study product changes or new variants should be cautious about excluding potentially essential attributes. Although including more relevant attributes increases choice variability, it also reduces bias.

Keywords: Discrete Choice Experiments, Choice Models, Missing Attribute Information

1. Introduction

Several decades of research have focused on what consumers do if attribute level information is missing (e.g., Huber & McCann, 1982; Johnson & Levin, 1985; Jaccard & Wood, 1988). For example, a consumer may evaluate two digital cameras but only know about one in terms of how many photos it can store. The missing information paradigm considers such problems, particularly issues related to what level (value), if any, consumers infer about missing attribute levels. Prior work has examined cases involving “missing” information based on the observation that consumers have limited memory capacity (e.g., Bettman, Luce, & Payne, 1998; Miller, 1956). A commonly cited example is one in which consumers cannot recall an attribute level for an option seen in a store earlier but know the level of another option in a current store (Lynch, Marmorstein, & Weigold, 1988).

In other cases, some attribute information may be missing for all options, which may be unrelated to consumers’ cognitive capabilities (e.g., Kardes, Posavac, & Cronley, 2004) but are related to managers’ decisions. For example, retailers may limit information to a certain set of features for several reasons: because of limited display space or information space on shelves, online displays or in catalogues; to help consumers deal with complexity; reducing catalogue or online display costs; or to “put their best face forward.” Similarly, regulations may not require labels to list all key information, or there may be insufficient package space to provide it.

Researchers also decide to limit feature information in order to control the length and complexity of experiments or surveys. For example, each feature added to a choice experiment increases its size and complexity (Louviere & Woodworth, 1983; Louviere, Hensher, & Swait, 2000; Swait and Adamowicz 2001); hence, researchers must decide which attributes to include or exclude.

Issues of whether or how preferences/choices change due to differences in attribute-information provision is both academically and practically relevant because we do not know

how including/excluding such information will impact preferences/choices, except in limited cases (e.g., Yates et al., 1978; Meyer, 1981; Sanbonamatsu et al., 1992; Kivetz & Simonson, 2000). For example, consumers may respond to subsets of attributes in different ways, such as devaluing options that omit information about certain attributes. Alternatively, they may perceive greater/lesser risk or uncertainty about options, leading to more choice variability.

A key contribution of this paper is to design and implement discrete choice experiments (DCEs) that systematically and independently manipulate the inclusion of attributes and examine the consequences of doing so under controlled conditions. A further contribution is to show that inclusion/exclusion of attributes in DCEs not only impacts mean utility (“systematic utility component”) but also impacts choice variability (“random utility component”). The latter contribution addresses Louviere’s (2001) call to recognise that actions by managers and experimenters can impact moments of distributions other than the mean. A final contribution is a comparison of covariance heterogeneity models (CHMs), with random coefficient models, in terms of prediction accuracy and their ability to capture attribute inclusion/exclusion effects.

2. Conceptual Framework

Kardes, Posavac, and Cronley (2004) comprehensively reviewed missing information and consumer inferences, and we use their ideas to develop a conceptual framework to anticipate potential impacts of varying the inclusion of attributes in DCEs. They describe several ways that consumers may respond to missing information, embedded in a broader context of inference-making. They note that consumers will make inferences about missing attribute information only if consumers notice that information is missing; if consumers do notice, their sensitivity to missing attribute information should increase, and the probability of choosing options with missing information should decrease. Thus, if relevant information is missing in a DCE, choice variability should decrease because consumers should choose more consistently

and should be more likely to choose options about which they have more information. Kardes et al. (2004) also noted that, when consumers compare options, this increases sensitivity to missing information relative to when consumers examine options separately. DCEs involve comparisons, so we expect to find these effects in DCEs. We also expect consumers to focus more on missing information relating to relevant attributes.

Sanbonmatsu et al. (2003) noted that if consumers notice that potentially relevant attributes are missing, they should rely less on presented attributes. That is, if consumers do not notice or care about missing information, available attribute information should be weighted more than missing information. Sanbonmatsu et al. (2003) show this effect for open-ended (thought listing) and closed-ended (constant sum scale) measures of attribute importance. In many cases, however, consumers seem to ignore missing information and give more weight to presented attributes regardless of the amount of information presented. Yet, if consumers notice that important attributes are missing, evaluation uncertainty should increase, which should increase choice variability. A related issue involves "non-diagnostics" (Troutman & Shanteau, 1977) or "dilution" (Zukier, 1982) effects, whereby individuals process information irrelevant to the task objectives. For example, Carpenter, Glazer, and Nakamoto (1994) noted that consumers may use any attribute that is present to compare options, even those that are ambiguous or irrelevant in evaluating options. Sanbonmatsu et al. (2003) showed further that consumers try to use available information, inferring that any such information is somehow relevant.

The more attributes used to describe options in a DCE, the greater the task complexity, if all else is equal (e.g., Lussier & Olshavsky 1979; Miller, 1956). Kardes, Posavac, and Cronley (2004) suggest that, because additional attributes increase task complexity, more attributes may lead to more spontaneous inferences, greater use of heuristic decision rules, and more inter-attribute trade-offs — all of which should increase (within subject) choice

variability (e.g., Payne, Bettman, & Johnson, 1993). Also, as task complexity increases, (between subject) choice variability should increase because consumers differ in how they make inferences about missing attributes and their use of decision heuristics.

Further, adding attributes to DCEs should lead to more ‘preference uncertainty’, which refers to how certain consumers are about either which option they prefer or to what degree. Preference uncertainty is associated with attribute conflicts (Fischer, Luce, & Jia, 2000). Attribute conflicts arise if options have both negative and positive attribute levels, such as a car with good acceleration, braking and handling, but poor comfort and safety. Roberts and Urban (1988) also noted that consumers’ uncertainty about brands increases as they evaluate more information about each brand. This also is consistent with work by Peterson and Pitz (1986; 1988), who found that consumer uncertainty increases with more attributes. Thus, these sources suggest that consumers’ abilities to evaluate options should improve with fewer attributes.

Haaijer, Kamakura, and Wedel (2000) studied response latencies, choice uncertainty, and error in DCEs and found that faster choices were more systematic than slower ones. This finding was supported by Rose and Black (2006) and Haaijer et al. (2000), who showed that response times impact preference heterogeneity and choice consistency in DCEs. Thus, we expect that the subjects in DCEs with fewer attributes will make faster choices; the cited studies related to response latency suggest that we also should expect more systematic choices (i.e., less response variability).

We follow Louviere (2001) in suggesting that choice variability can be independently identified and modeled separately from mean utility. This focus on both preference and choice variability is similar to the JUMP model of Chandrashekar et al. (2000), who suggested that various factors may separately impact both judgment magnitude (i.e., systematic utility components) and judgment uncertainty (i.e., choice variability). However, our research differs

from this model in several key ways, such as focusing on choices (not ratings) and using experimental (not observational) data. We also note that the JUMP model is similar to choice models that parameterize preference heterogeneity using covariate main effects and interactions.

Swait and Louviere (1993); Louviere, Hensher, and Swait (2000); and Louviere (2001) noted that all empirical choice models confound estimated parameters with choice variability, which poses statistical inference issues unless one controls for and/or accounts for differences in choice variability. We are unaware of prior work in the missing-information paradigm that controlled for choice variability differences. In this study, we show that missing attribute information has little effect on choices, but if one does not control for choice variability differences, one is likely to conclude that there are effects on choices.

To show such effects, we use covariance heterogeneity models (CHMs) similar to those discussed by Louviere, Hensher, and Swait (2000, p. 195) to model unobserved heterogeneity sources using covariates. CHM use is growing in applied economics (e.g., DeShazo & Fermo 2002; Swait & Adamowicz, 2001b), and CHMs also have been used in marketing (e.g., Delleart, Brazell, & Louviere, 1999). CHMs allow one to decompose choice experiment effects on both systematic and random components. This allows one to capture the unobserved variability associated with a) each attribute level, and b) the inclusion/exclusion of attributes across experimental conditions. Much prior work in marketing and applied economics has used random coefficient models to capture preference heterogeneity, so we compare CHMs with these models. We now describe and discuss the CHMs that we used, followed by a discussion of random coefficient models.

3. Modeling the Impact of Including/Excluding Attributes

The conceptual basis for the models and DCEs used to produce data for our model estimation is random utility theory (RUT). RUT posits that preferences are latent and unobservable, such that the latent utility that consumer ‘n’ associates with option ‘i’, U_{in} , can be expressed as an additive function of a systematic component, V_{in} , and a random component, ε_{in} (Manski, 1977; McFadden, 1974; Thurstone, 1927). That is,

$$U_{in} = V_{in} + \varepsilon_{in}. \quad (1)$$

Systematic utility (V_{in}) is assumed to be a function of predictors represented by the parameters of a generalized regression function (Ben-Akiva & Lerman, 1985). We define the variable x_{inkp} , which takes the value 1 if option ‘i’ evaluated by consumer ‘n’ has level ‘p’ on attribute ‘k’, the value 0 if another level is evaluated, or the value -1 if this level is unknown. Hence, we capture included levels of each attribute with dummy codes and its exclusion with effects-coded variables. We define a set of parameters (β) to capture the effect that each of these coded attribute levels have on mean utility. That is,

$$V_{in} = \sum_{k=1}^K \sum_{p=1}^{k_p} \beta_{kp} x_{inkp} = \beta X_i, \quad (2)$$

where K is the total number of attributes, and k_p is the number of levels for attribute ‘k’. Thus, Equation (2) specifies each measurable component relative to attribute ‘k’ being missing. In most published applications of choice models, β_k captures preferences for an attribute level, ‘p’, relative to a base level of attribute k. We estimate β_{km} , the impact of each attribute on mean utility when attribute information is missing, via Equation (3):

$$\beta_{km} = - \sum_{p=1}^{k_p} \beta_{kp} \quad (3)$$

We also propose that included/excluded attribute information impacts random error component variability (i.e., response variability), σ_ε^2 . In any model of latent utility, the scale

(magnitude) of the parameter vector is inversely related to the standard deviation of the error distribution, σ_ε . For example, errors in multinomial logit (MNL) models are assumed to be independently and identically distributed (IID) extreme value type 1 variates, with zero location parameter and scale parameter, λ , where $\lambda = \text{SQRT}(\pi^2/6\sigma_\varepsilon^2)$. McFadden (1974) showed that the probability that option 'i' is chosen from a set of 'J' options (dropping 'n' for simplicity), can be written as Equation (4):

$$P_i = \frac{\exp(\lambda V_i)}{\sum_{j=1}^J \exp(\lambda V_j)} . \quad (4)$$

In any given data set, λ is not identified and so must be made constant (e.g., “1”). Ben-Akiva and Lerman (1985) showed that the estimated parameter vector, $\hat{\beta}$, is $\lambda\beta$ and not true β ; hence, estimates of β are confounded with λ in all choice models (McFadden 1974; Swait & Louviere, 1993; Train 2003). Thus, by ignoring λ , researchers can make incorrect inferences about consumers’ preferences.

Because scale is confounded with model parameter estimates, prior empirical results suggesting that consumers devalue options with missing information by devaluing missing attributes may be due to not controlling for choice variability differences. That is, because λ is inversely proportional to the error variance, factors that increase error variance decrease estimated parameter magnitudes ($\hat{\beta}$), which may lead to misinterpreting such effects as devaluation. Thus, comparisons among different conditions that include and exclude attribute levels are confounded by variability differences and should be accounted for in any analysis.

Using a CHM approach, we address the identification problem by specifying the variance of the random component to be a function of attributes varied in DCEs. Various names have been applied to CHMs in prior work, creating some confusion in nomenclature. For example, Swait and Adamowicz (2001a) called it a “Parameterized Heteroscedastic MNL

model” (PHMNL), using it to specify λ as a function of an entropy-based measure of task complexity. DeShazo and Fermo (2002) called it a “Heteroscedastic MNL model” (HMNL), using it to specify scale as a function of covariates that measure the quantity of information in a choice set and the way in which information is configured. Delleart et al. (1999) also called it “HMNL”, using it to specify scale as a function of absolute price levels and price differences.

We expect that varying attribute level information about option 'i' — i.e., whether an attribute is missing — will impact choice variability. Following Louviere et al. (2000), Swait and Adamowicz (2001a), DeShazo and Fermo (2002), and Delleart et al. (1999), we represent these effects by a design matrix, X_i , with scale effects specified as $\lambda = \exp(\alpha X_i)$, where λ is a variance-scale ratio of various conditions relative to a reference condition. The resulting choice model is Equation (5):

$$P_i = \frac{\exp(\lambda V_i)}{\sum_{j=1}^J \exp(\lambda V_j)} = \frac{\exp(\exp(\alpha X_i)(\beta X_i))}{\sum_{j=1}^J \exp(\exp(\alpha X_j)(\beta X_j))}. \quad (5)$$

The significant elements of α and β imply noteworthy effects on mean utility and scale, respectively. Scale multiplies the systematic component elements, so significant scale effects magnify effects of corresponding attributes in the systematic component, implying interactions between scale effects and systematic component effects. Hence, CHM allows both non-linear and non-additive effects on preferences.

CHMs can be viewed as a type of random coefficients model because scale is inversely proportional to error variance, and a CHM with all attribute effects in the scale function captures variability in the estimates across a sample population. Thus, the CHM results we report later estimate the variability in each attribute level, and significant estimates can be graphed to visualize relationships between attribute levels and unobserved variability.

Finally, we compare CHM with the mixed logit model (MIXL) of Revelt and Train (1998). MIXL can be written as:

$$P_i = \exp[(\beta_k + \varpi_k)X_{ki}] / \sum_{j \in C} \exp[(\beta_k + \varpi_k)X_{kj}], \quad (6)$$

where β_k is a vector of random effects, including alternative-specific intercepts with associated disturbance terms, ϖ_k , that correspond to the design matrix of covariates X_{ki} and X_{kj} . The disturbance terms allow each of the $k = 1, \dots, K$ parameter vector elements to be distributed across the sample population. We specify the K -element vector in the same way as the CHMs. All attributes are effects coded, but each attribute has an additional “level” to estimate the effect of it as “missing”. We follow Revelt and Train (1998) in assuming that random effects are normally distributed with mean β_k and variance ϖ_k . Because MIXL is not a closed form, we estimate it with simulated maximum likelihood (SML) (Revelt & Train, 1998), modifying Train’s GAUSS code (<http://elsa.berkeley.edu/~train/software.html>) to do this.

We think that this is the first CHM and MIXL comparison with comparable statistical effects. Specifically, both CHM and MIXL estimate conditional means associated with each attribute level. However, CHM estimates the scale effects associated with each attribute level, while MIXL estimates standard deviations. CHM assumes the errors are non-independent extreme value type-1 random variates, while MIXL assumes that errors are IID extreme value type-1 random variates, with attribute level effects distributed as standard normal random variates (with estimated standard deviations). Thus, we expect both models will provide similar statistical descriptions of the effects; both capture unobserved variability but do so in different ways.

4. Impact of Missing Information on Choices

Our CHM specification allows us to anticipate signs and parameter-estimate magnitudes based on prior work. For example, work in economics and psychology suggests that consumers

evaluate attribute level information about products, combining the level values to holistically evaluate each option (e.g., Fishbein & Azjen, 1975; Louviere, 1988). In the missing information paradigm, the experimental participants *know* the attribute exists, but there is no information for one or a few options. In our case, “missing” attributes are *not* shown, so our experimental participants do not know that a missing attribute describes choice options (unless they make assumptions or inferences).

Prior work consistently shows that missing information impacts systematic utility components, but as we noted earlier, we are unaware of attempts to account for differences in choice variability. For example, Meyer (1981) showed that if attribute information is missing, consumers tended to assign a score equal to their adaptation level. Huber and McCann (1982) showed that consumers used beliefs about price and quality correlations to infer missing price or quality levels. Broniarczyk and Alba (1994) noted that consumers’ existing (i.e., prior) beliefs can influence their inferences. Thus, common themes in this paradigm are a) the role of inferences or beliefs that go beyond the information provided (e.g., Bruner, 1957), and b) devaluation of choice options (e.g., Sanbonamatsu et al., 1992; Johnson & Levin, 1985; Meyer, 1981; Simmons & Lynch, 1991). In the case of devaluation, discounting of options that have missing information is a typical hypothesis. For example, Meyer (1981) noted that missing attribute levels are assigned “below neutral” utility value and that this devaluation impacts the overall evaluation. Johnson and Levin (1985) found that evaluations decreased with more missing information. Simmons and Lynch (1991) suggested that devaluation is due to negative cues provided by missing information.

Our CHM specifications estimate β_{km} to capture effects on mean utility when a level of attribute 'k' is missing. As noted, prior work suggests that inferred levels of missing attributes will be negative. Our CHM specifications include estimates that capture effects on the scale component for missing attribute levels (α_{km}): positively signed α_{km} imply that, if a level of

attribute 'k' is missing, response variability decreases. Signs and magnitudes of α_{km} matter when interpreting behavioral outcomes because such effects suggest that option devaluation due to missing attribute information may be due to scale differences. Specifically, observationally equivalent results can occur with slightly negative and insignificant β_{km} combined with positive α_{km} (i.e., less choice variability). As previously noted, we expect response variability to decrease if an attribute is excluded in a DCE condition; therefore, we expect the α_{km} estimates to be significant and positive.

5. Research Approach

We designed an experiment to obtain choice data to estimate the model effects described earlier and control for as many other effects as possible. Specifically, if each attribute can be included or excluded, attribute information has two states. Thus, for K attributes, there are 2^K total ways to vary information (include/exclude) for any set of K attributes.

We studied two product categories (delivered pizzas and packaged juices) and varied the inclusion of seven attributes in both categories. Each category had core attributes that were always included (brand name and price for pizzas; brand name, price, and type of juice for juices). We used an orthogonal fractional factorial design to create 32 “master conditions,” in which each attribute occurs equally often and each pair of attributes co-occur equally often (Hahn & Shapiro 1966, Experimental Plan Code 6c). Known as a “Resolution 5” design, it allows all main effects and two-way interactions to be estimated independently of one another but not independently of higher-order interactions, which must be assumed insignificant. The master design can be split into two orthogonal main effects designs and used to conduct rigorous split-half comparisons, as we later discuss.

We selected pizza and fruit juice to study because these are familiar to our student respondents, who can afford to purchase them. These also represent different product types

with different price points, allowing cross-category comparisons. Attributes and levels are in column 1 in Appendix A and B. Pizza attributes and levels are based on local pizza supplier menus and ingredient lists; juice attributes were based on package information from supermarkets. Attributes and levels were refined using focus groups with participants from the study population. In every condition, brand and price are always present and vary across four levels.

In each of the 32 master conditions we used an orthogonal main effects plan (OMEPE) to create product profile attribute/level descriptions for attributes assigned to that condition. For example, condition one, pizzas, has five 4-level attributes and five 2-level attributes ($2^5 \times 4^5$). Condition two, juices, has four 4-level attributes and four 2-level attributes ($2^4 \times 4^4$). We created different designs for each of the 32 conditions but based all designs on a common OMEPE from the 8×4^8 factorial. That is, we assigned all 4-level attributes to a fixed set of columns in the 8×4^8 OMEPE, and we assigned all 2-level attributes to other columns by collapsing each 4-level column into two two-level columns (2×2 sub-designs). This ensures that each attribute always varies in the same way with other attributes regardless of the conditions where it occurs. The full experiment consists of 32×32 treatments (product profiles or scenarios) = 1024×2 categories = 2048 total scenarios. It is worth noting that one could also design the experiments as choices among pairs, triples, etc. Optimal design theory for these problems recently has become available (e.g., Street, Burgess and Louviere, 2005), making it feasible to control for differences in design efficiency.

We randomly assigned 18 subjects to each of the 32 master conditions to complete the paper and pencil survey. Each subject evaluated the 32 product profiles in the master condition to which they were assigned. Participants were recruited from undergraduate marketing classes, receiving class credit consistent with university ethics guidelines.

Subjects evaluated each pizza or juice description one-at-a-time and decided whether to buy each of the 32. We used this binary response task because a) such tasks are used in applied economics and marketing (Louviere et al., 2000), and b) they resemble traditional conjoint tasks. On average, subjects took approximately 20 minutes to complete the tasks.

6. Results

6.1 Simulation Comparing Predictive Performance of MIXL and CHM

To obtain additional insights about in- and out-of-sample performance of CHM and MIXL, we designed a simulation study by varying parameter values that reflect different degrees of sensitivity (2), preference heterogeneity (2) and choice consistency (2), as described in Appendix C. We also used four sample sizes (10, 20, 30, and 50); therefore, the simulation was a $2^3 \times 4$ factorial. In the interests of brevity, we do not present these results (these are available on request). The findings can be summarized as follows:

- If there is preference heterogeneity in any attribute, MIXL does well on in-sample predictions, regardless of sample size or scale;
- If there is no preference heterogeneity in attributes, neither MIXL nor CHM is consistently superior for in-sample predictions; and
- CHM is superior for out-of-sample predictions, especially at higher scale values.

Generally, MIXL outperformed CHM for in-sample fits, but CHM outperformed MIXL in out-of-sample fits. That is, MIXL generally performed better at predicting the choices of the hypothetical consumers from which it was estimated (in-sample). However, CHM generally performed better in predicting out-of-sample choices or generalizing predictions.

The reason for this seemingly contradictory result is that MIXL does not capture the data generation process (DGP) when there is no preference heterogeneity; perhaps more disappointingly, MIXL seems to identify preference heterogeneity when there is none. CHM

captures preference heterogeneity via the estimates associated with relative scale and seems to capture the mean DGP better than MIXL. Overall, the simulation results support the empirical results reported here that show that MIXL is better in-sample, and CHM is better out-of-sample.

Following Andrews, Ainslie, and Currim (2002), we analysed our simulation results with ANOVA and multiple analysis of covariance (MANOVA). Dependent variables are in- and out-of-sample model fits measured by R-sq (MXL_in sample R-Sq; MXL_out sample R-sq; CHM_in sample R-sq; CHM_out sample R-sq); independent variables are a) heterogeneity in an important attribute (low, high); b) heterogeneity in a non-important attribute (low, high); and c) scale (low, high). We find significant main effects for heterogeneity in important attributes and scale (details available from authors).

Turning our attention back to our experiments, we summarize the proportions of subjects who “would buy” for present/absent attributes (Table 1) to set the scene for the results that follow. For Pizza, variances differ on six of seven attributes, while means differ for two of seven attributes (at 95% C.I.). For Juice, variances differ for four of seven attributes, while means differ for two of seven attributes. In general, omitting information (simplifying sampling choice) has a positive impact on choice. These results suggest there are larger variance differences in mean responses when attributes are present versus missing.

Insert Table 1 Here

6.2. Overall model comparisons

The estimated model parameters of the CHM and MIXL results are in Appendix A (Pizza) and B (Juice). Not surprisingly, CHM and MIXL outperform MNL for both in-sample and out of sample tests; thus, for reasons of brevity we omit MNL results (available on request). MNL is a restriction on the more general CHM; CHM relaxes the IID error assumption by allowing error variances to co-vary systematically with particular attributes. MIXL relaxes the common

parameter vector assumption of MNL by allowing a distribution of preference parameters within the sample.

Because subjects completed all choice scenarios, we can calculate the proportion of “yes” responses for each profile (32) in each master condition. Thus, we can calculate traditional R-squares for each model in each condition by comparing model-predicted choice probabilities with observed yes proportions. We evaluate in-sample fits using likelihood ratio tests (LRTs) and use R-square values for out-of-sample comparisons. For brevity, we simply note that both MIXL and CHM consistently outperform MNL on these tests.

MIXL is consistently superior to CHM for in-sample fits for both categories. Noting that CHM and MIXL have the same number of parameters, the LRT results are: a) Pizza CHM Log L = - 11013.1; Pizza MIXL Log L = -9907.9; b) Juice CHM Log L = - 11501.6; and Juice MIXL Log L = -10735.9.

We used split-half cross-validations to compare out-of-sample fits based on R-square values for predicted choice proportions for all 32 Master Conditions x 32 scenarios = 1024 observations. We calculated the proportion of would buy responses for each of the 1024 cases and used the models to predict the proportion of “yes” responses. We then split the 32 conditions into two orthogonal halves. This allows rigorous split-sample comparisons of out-of-sample fits instead of merely predicting based on some arbitrary set of “holdout choices” (Table 2).

----- **Insert Table 2 here** -----

In two out of four comparisons, CHM outperforms MIXL in terms of out-of-sample fit. We use consumer choices to estimate parameters for each person via Bayes rule and use these estimates to calculate in-sample fits (i.e., R-square). For out-of-sample fits, we randomly selected 16 profiles from the 32 available profiles for each person to calibrate individual level

estimates using Bayes rule, and then we used these estimates to predict the remaining 16 profiles.

The most intuitive explanation for our results is that MIXL is better at making in-sample predictions because it accounts for unobserved variability as preference heterogeneity, expressing each sampled individual as a deviation from the overall sample. Hence, MIXL captures between-subject differences, while CHM captures all sources of variability, including both within- and between-subject variability. Generalizing MIXL to out-of-sample choices is problematic if there is new and potentially different between-subject variability. Louviere (2006) and Louviere and Meyer (2007) suggest that many potential sources of variability that may be constant in an estimation dataset are not constant in a prediction sample. Hence, MIXL will do poorly at out-of-sample predictions if the out-of-sample responses differ in some material way with regard to within- and/or between-subject heterogeneity. In turn, this implies that MIXL may fail to capture some relevant sources of variability in some datasets, posing issues for generalizing MIXL to other data sources (samples). Thus, models that can capture components of error variability may provide more generalizable predictions than random coefficient models.

6.3. Impact of missing attributes on systematic and random components

The results in Tables 3a and 3b focus on missing information effects. MIXL pizza results indicate several significant effects for included/excluded attribute information; for example, if information about extras such as free delivery or salads is missing, the probability of saying “yes” decreases, which is consistent with prior work that has shown missing information implied devaluation. However, only one effect is significant (salad, $p=0.01$). MIXL pizza results also imply that the probability of saying “yes” is more likely if the crust is missing than present, contradicting prior missing information results, although Sanbonmatsu et al. (1992) and Kardes et al. (2004) note that option devaluation occurs only if consumers

recognize that information is missing. Overall, the results for the Pizza and Juice datasets are similar.

----- **Insert Table 3a and Table 3b here** -----

Yet once variability differences are captured in the CHM, the pizza results change. That is, there are no significant missing attribute effects in the systematic utility component, but all effects are significant and positive in the scale component. This indicates that the propensity to say “yes” did not change if attributes are missing; while respondents neither devalued nor inferred more positive values, within- and/or between-person choice variability decreased significantly.

MIXL juice results imply that including attribute information significantly increases the propensity to say “yes” for pack-size, added sugar, and for all types of packaging material studied. Propensity to say “yes” decreases if calcium is missing. MIXL juice results indicate that the estimated impacts on the random component are similar to the pizza results.

CHM juice results imply that less variability is associated with missing attributes for all except one attribute (at the 90% C.I.). In contrast to the CHM pizza results, the effects of type of packaging information, pack size, added sugar, and percentage of real juice on the systematic component is significant after accounting for random component variability. Hence, missing information about pack size, added sugar, and type of packaging increases purchase likelihoods, but missing information about the percentage of real juice decreases these (Table 3b).

It is worth noting that the systematic utility component results may depend on attribute levels experienced by subjects in real life in each category and/or what they believe are the real levels when each attribute is included. That is, if subjects are familiar with better/worse attribute levels than experimenters present, on average this should decrease/increase propensity to say “yes” in response to variation in levels. For example, our results suggest that the sample

is more likely to say “yes” if delivery time information is missing. Without other information, respondents may believe that (on average) real delivery times are faster than the average of the levels that we varied when delivery time was included (10, 20, 30, and 40 min.). If true, it could lead to higher estimates about the likelihood of purchase, which suggests that future research should consider how expectations relate to attribute levels varied in DCEs (and other experiments).

Similarly, many juice products have relatively low real juice percentages, so it may be that the sample inferred that (on average) the percent of juice was lower if this information was missing. In turn, this implies that if information about some attributes is missing, consumers may be more likely to say “yes” or “no” simply because they like/dislike the levels they see when information is included. In any case, the cross-category results for missing attribute effects on systematic utility components warrants further investigation, and a logical hypothesis to test is that the effect can be manipulated by an experimenter’s choice of levels and/or that the choice of levels interacts with respondent expectations.

6.4. Impact of levels of included attributes on random components

Both pizza and juice price results show an inverse-U relationship between price levels and choice variability, such that choice variability is highest for the highest price levels, as shown in Figures 1a (pizza) and 1b (juices). Louviere (2001) noted that variability may be higher for extreme levels if a) consumers respond less consistently to higher prices because they don’t know if extreme price levels are riskier (they may be higher or lower than expected); and/or b) price sensitivity differences across consumers are larger at extreme price levels.

----- Insert Figure 1a and Figure 1b here -----

The results show that choice variability decreases if attribute information is missing and is systematically related to the levels of several attributes. For example, choice variability exhibits an inverse-U shaped relationship with number of toppings and delivery times, which

may be due to different sensitivity to the levels, leading to non-linear conditional utilities. If the conditional response functions are non-linear and one aggregates across consumers, errors in model estimates are likely to be larger for extreme than middle levels. That is, if utility decreases at a decreasing rate for some, but decreases at an increasing rate for others, response curves will differ more at extreme than middle levels.

The random component results also show differences in choice variability associated with qualitative attribute levels. For example, juice packaging results (Appendix A, CHM results) show that the sample preferred glass ($t=12.35$) and were indifferent to aluminum ($t=-1.30$), but neither packaging type affected the random component ($t=0.46$ and 0.21 for glass and aluminum, respectively). They disliked plastic bags ($t=-15.28$) and preferred tetra packs ($t=3.10$); their choices reflected consistency in these preferences, as captured by a significant and negative scale effect ($t=-2.78$ for plastic; $t=-2.35$ for tetra).

6.5. Comparison of Mean Parameter Estimates of CHM and MIXL

Both models yield the same mean utility estimates for the attribute levels up to scale, again indicating that their primary difference lies in assumptions about unobserved variability. Specifically, the estimated parameters are the same up to a scaling factor that reflects differences in unobserved variability in the data sources from which the models were estimated (Swait & Louviere, 1993). This suggests convergent validity, with correlations between CHM and MIXL mean estimates, respectively, of 0.976 for Pizza and 0.953 for Juice.

Recall that parameter vectors estimated from choice data, $\hat{\beta}$, are confounded with scale. In theory, CHM estimates true β . We can adjust the data with scales recovered from CHM, which should allow us to recover true β for MNL. Therefore, we adjusted for CHM scale and re-estimated MNL with an adjusted design matrix, λX , instead of X . We obtained a revised mean ratio of CHM/(MNL_adjusted with scale), which was not statistically different from 1 for either category ($p=0.64$ for Pizza; $p=0.31$ for Juice; details available on request).

7. Discussion and Conclusions

We described and discussed research designed to give better insights into the likely effects of including/excluding attributes in DCEs. We used MIXLs and CHMs to separate the effects associated with 32 master conditions used to vary attribute inclusion/exclusion, while nesting separate DCEs under each master-condition that only varied “included” attributes. This approach allowed us to identify and specify both systematic and random components as a function of included attribute levels. Our results suggest that excluded/missing attribute information effects depend on the attribute levels and/or category that researchers choose. That is, our CHM pizza results suggest that missing attribute information did not impact the systematic component but increased the scale component, implying that choice consistency increased if attribute information was missing. For juice, missing attribute information affected both components. Additives and packaging attributes increased propensity to say “yes” if missing; similar to the pizza results, missing attributes increased choice consistency. This missing information increased choice consistency in both categories, suggesting that consumers may respond as if “what I don't know probably won't hurt me,” or ignore information not presented, which may lead to more consistent choices.

Behaviorally, our results suggest that decisions to exclude attributes in DCEs will affect choice consistency, but the impacts on propensity to buy products are unclear. Decreasing the amount of attribute information presented leads to more consistent choices, but choosing whether to buy with less information depends on the attribute levels of product profiles relative to consumers’ real-life experiences with the attributes. From a statistical modeling perspective, our comparison of CHMs and MIXL using simulated data supported our empirical findings. CHMs fared worse on in-sample fits but fared better on out-of-sample, split-half predictions, suggesting that CHMs warrant more attention by researchers. Our CHM models allowed us to separate systematic and random component effects associated with attributes and capture non-

linear and non-additive impacts of attributes on utility and choice. CHMs also are closed-form, which may be advantageous in some cases (e.g., quicker estimation). Because of the way in which we specified MIXL and CHM, we can visualize relationships not only between attribute levels and systematic effects but also between attribute levels and scale-component effects. We also found several numerical attributes, such as price, number of toppings, and percent real fruit juice, that were associated with U-shaped error variance functions, implying more variability at extreme levels than middle levels. The latter is consistent with Louviere's (2001) notion that extreme levels may impact within- and between-subject variability, although our modeling approach does not allow these two effects to be separated.

There are several implications of our work for academic and commercial DCE research. Failure to account for unobserved variability in comparing or pooling DCE results is likely to lead to biased conclusions about mean attribute effects. Thus, one should exercise caution in excluding attributes in DCEs and compare effects across conditions only after taking variability differences into account. One also should be cautious about assuming constant random errors because our results and a growing body of empirical work suggest that they are neither constant nor non-systematic.

Our results also suggest that unobserved choice variability is likely to be an important behavioral phenomenon in its own right, as noted by Louviere (2001), Louviere et al. (2002), and Louviere and Meyer (2007). However, we know little about unobserved variability. Despite much research on variability between individuals (preference heterogeneity) in the last decade, this is only one type of unobserved variability. Thus, the field would benefit from more research into sources of unobserved variability in real and experimental markets.

Overall, our results suggest that the decisions researchers make about which attributes (and associated attribute levels) to include in DCEs impact both mean utility and choice variability. This probably also applies to real markets, such that managers' marketing actions

impact both components. For example, if price levels impact both mean utility and choice variability, at least part of the price impact may be due to scale differences, not real utility differences. Louviere and Meyer (2007) showed that scale impacts on price effects can be as much as 35% to 40% of the estimated price effect. In some cases, manipulating prices may lead to more or less consumer uncertainty or heterogeneity, while in other cases, it may lead to changes in both. To accurately assess pricing and other policies, managers need to know which is which. CHMs can separate these effects and relate them to observable covariates like age, education and geographical location, providing insights into impacts on both components.

References

- Andrews, R.L., Ainslie, A. & Currim, I.S. (2002). An Empirical Comparison of Logit Choice Models with Discrete versus Continuous Representations of Heterogeneity. *Journal of Marketing Research*, 39(4), 479-487.
- Ben-Akiva, M. & Lerman, S. R. (1985). *Discrete-Choice Analysis*. Cambridge: MIT Press.
- Bettman, J.R., Luce, M.R., & Payne, J.W. (1998). Constructive Consumer Choice Processes. *Journal of Consumer Research*, 25(December), 187-217.
- Broniarczyk, S.M., & Alba, J.W. (1994). The Role of consumers' Intuitions in Inference Making. *Journal of Consumer Research*, 21(December), 393-407.
- Bruner, J.S. (1957). Going Beyond the Information Given. In H.E. Gruber, K.R. Hammond, & R. Jessor (Eds.), *Contemporary Approaches to Cognition* (pp. 41-69). Cambridge: MA, Harvard University Press.
- Carpenter, G.S., Glazer, R., & Nakamoto, K. (1994). Meaningful Brands from Meaningless Differentiation - the Dependence on Irrelevant Attributes. *Journal of Marketing Research*, 31, 339-350.
- Chandrashekar, M., McNeilly, K., Russ, F.A., & Marinova, D. (2000). From Uncertain Intentions to Actual Behavior: A Threshold Model of Whether and When Salespeople Quit. *Journal of Marketing Research*, 37(4), 463-79.
- Dellaert, B., Brazell, J., & Louviere, J.J. (1999). The Effect of Attribute Variation on Consumer Choice Consistency. *Marketing Letters*, 10(2), 139-147.
- DeShazo, J.R., & Fermo, G. (2002). Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management*, 44(1), 123-143.
- Fischer, G.W., Luce, M.F., & Jia, J. (2000). Attribute Conflict and Preference Uncertainty: Effects on Judgement Time and Error. *Management Science*, 46, 88-103.

- Fishbein, M., & Ajzen, I. (1975). *Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research*. MA: Addison-Wesley.
- Hahn, G.J., & Shapiro, S.S. (1966). A Catalog and Computer Program for the Design and Analysis of Symmetric and Asymmetric Fractional Factorial Experiments. *G.E. Technical Report No. 66-0-165*, General Electric Research and Development Center, Schenectady, New York (May).
- Haaijer, R., Kamakura, W., & Wedel, M. (2000). Response Latencies in the Analysis of Conjoint Choice Experiments. *Journal of Marketing Research*, 37, 376-382.
- Huber, J., & McCann, J. (1982). The Impact of Inferential Beliefs on Product Evaluations. *Journal of Marketing Research*, 19(August), 324-333.
- Jaccard, J., & Wood, G. (1988). The Effects of Incomplete Information on the Formation of Attitudes Toward Behavioral Alternatives. *Journal of Personality and Social Psychology*, 54(4), 580-591.
- Johnson, R.D., & Levin, I.P. (1985). More than Meets the Eye: The Effect of Missing Information on Purchase Evaluations. *Journal of Consumer Research*, 12, 169-177.
- Kardes, F.R., Posavac, S.S., & Cronley, M.L. (2004). Consumer Inference: A Review of Processes, Bases and Judgment Context. *Journal of Consumer Psychology*, 14, 230-256.
- Kivetz, R., & Simonson, I. (2000). The Effect of Incomplete Information on Consumer Choice. *Journal of Marketing Research*, 37, 427-448.
- Louviere, J.J. (1988). *Analyzing Decision Making: Metric Conjoint Analysis*, Sage University Papers Series Number 67. Newbury Park, CA: Sage Publications, Inc.
- Louviere, J.J. (2001). What if Consumer Experiments Impact Variances as Well as Means? Response Variability as a Behavioral Phenomenon. *Journal of Consumer Research*, 28, 506-511.

- Louviere, J.J. (2006). What You Don't Know Might Hurt You: Some Unresolved Issues In The Design and Analysis of Discrete Choice Experiments. *Environmental and Resource Economics*, 34(1), 173-188.
- Louviere, J.J., & Woodworth, G.W. (1983). Design and Analysis of Simulated Consumer Choice or Allocation Experiments. *Journal of Marketing Research*, 20, 350-367.
- Louviere, J.J., & Meyer, R.J. (2007). Formal Choice Models of Informal Choices: What Choice Modeling Research Can (and Can't) learn from Behavioral Theory. *Review of Marketing Research*, 4, (in press).
- Louviere, J.J., Meyer, R.J., Bunch, D.S., Carson, R., Dellaert, B., Hanemann, W.M., Hensher, D., & Irwin, J. (1999). Combining Sources of Preference Data for Modelling Complex Decision Processes. *Marketing Letters*, 10(3), 187-204.
- Louviere, J.J., Carson, R., Ainslie, A., Cameron, T.A., DeShazo, J.R., Hensher, D., Kohn, R., Marley, T., & Street, D. (2002). Dissecting the Random Component of Utility. *Marketing Letters*, 13(3), 177-193.
- Louviere, J.J., Hensher, D., & Swait, J. (2000). *Stated Choice Methods: Analysis and Applications*. Cambridge, UK: Cambridge University Press.
- Lussier, D.A., & Olshavsky, R.W. (1979). Task Complexity and Contingent Processing in Brand Choice. *Journal of Consumer Research*, 6(September), 154-165.
- Lynch, J.G., Marmorstein, H., & Weigold, M.F. (1988). Choices from Sets Including Remembered Brands: Use of Recalled Attributes and Prior Overall Evaluations. *Journal of Consumer Research*, 15(2), 169-85.
- Manski, C. (1977). The Structure of Random Utility Models. *Theory and Decision*, 8, 229-254.
- Marley, A.A.J., & Louviere, J.J. (2005). Some Probabilistic Models of Best, Worst, and Best-Worst Choices. *Journal of Mathematical Psychology*, 49, 464-480.

- McFadden, D. (1981). Econometric Models of Probabilistic Choice. In C. Manski, & D. McFadden (Eds.), *Structural Analysis of Discrete with Econometric Applications* (pp. 198-272). Cambridge, MA: MIT Press.
- McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in Econometrics* (pp. 105-142). New York: Academic Press.
- Meyer, R.J. (1981). A Model of Multiattribute Judgements Under Attribute Uncertainty and Informational Constraint. *Journal of Consumer Research*, 18, 428-41.
- Miller, G.A. (1956). The Magical Number Seven, Plus Or Minus Two: Some Limits On Our Capacity For Processing Information. *Psychological Review*, 63, 81-97.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The Adaptive Decision Maker*. Cambridge, UK: Cambridge University Press.
- Peterson, D., & Pitz, D.F. (1986). The Effects of Amount of Information on Prediction of Uncertain Quantities. *Acta Psychologica*, 61, 229-241.
- Peterson, D., & Pitz, D.F. (1988). Confidence, Uncertainty, and the Use of Information. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14(1), 229-241.
- Revelt, D., & Train K. (1998). Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics*, 53, 647-657.
- Roberts, J.H., & Urban, G.L. (1988). Modeling Multiattribute Utility, Risk, and Belief Dynamics for New Consumer Durable Brand Choice. *Management Science*, 34, 167-185.
- Rose, J.M., & Black, I.R. (2006). Means Matter, but Variances Matter Too: Decomposing Response Latency Influences on Variance Heterogeneity in Stated Preference Experiments. *Marketing Letters*, 17, 295-310.

- Sanbonmatsu, D.M., Kardes, F.R., & Herr, P.M. (1992). The Role of Prior Knowledge and Missing Information in Multiattribute Evaluation. *Organizational Behavior & Human Decision Processes*, 51(1), 76-91.
- Sanbonmatsu, D.M., Kardes, F.R., Houghton, D.C., Ho, E.A., & Posavac, S.S. (2003). Overestimating the Importance of the Given Information in Multiattribute Consumer Judgment. *Journal of Consumer Psychology*, 13, 289-300.
- Simmons, C.J., & Lynch, J.G. (1991). Inference Effects without Inference Making? Effects of Missing Information on Discounting and Use of Presented Information. *Journal of Consumer Research*, 17(March), 477-491.
- Street, D., Burgess, L. & Louviere, J.J. (2005). Quick and Easy Choice Sets: Constructing Optimal and Nearly Optimal Stated Choice Experiments. *International Journal of Research in Marketing*, 22, 459-470.
- Swait, J., & Adamowicz, W. (2001a). Incorporating the Effect of Choice Environment and Complexity into Random Utility Models. *Organizational Behavior and Human Decision Processes*, 86(2), 141-167.
- Swait, J., & Adamowicz, W. (2001b). The Influence of Task Complexity on Consumer Choice: A Latent Class Model of Decision Strategy Switching. *Journal of Consumer Research*, 28(1), 135-148.
- Swait, J., & Louviere, J.J. (1993). The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. *Journal of Marketing Research*, 30(August), 305-314.
- Thurstone, L. (1927). A Law of Comparative Judgement. *Psychological Review*, 34, 273-286.
- Train, K. (2003). *Discrete Choice Methods with Simulation*. Cambridge, UK: Cambridge University Press.

Troutman, M.C., & Shanteau, J. (1977). Inferences Based on Nondiagnostic Information.

Organizational Behavior and Human Performance, 19, 43-55.

Yates, J.F., Jagacinski, C.M., & Faber, M.D. (1978). Evaluation of Partially Described

Multiattribute Options. *Organizational Behavior and Human Performance*, 21(2), 240-251.

Zukier, H. (1982). The Dilution Effect: The Role of the Correlation and the Dispersion of

Predictor Variables in the Use of Nondiagnostic Information. *Journal of Personality and Social Psychology*, 43(6), 1163-1174.

Appendix A: Detailed Model Estimation of Pizza Results

Parameters	CHM Model						Mixed Logit					
	Systematic Component			Scale Component			Systematic Component			Variance Component		
	Est. (β)	t-stat	prob.	Est. (α)	t-stat	prob.	Est. (β)	t-stat	prob.	Est. (ϖ)	t-stat	prob.
Pizza Hut	0.0729	3.3	0.00	-0.0068	-0.2	0.87	0.1998	4.9	0	2.0137	40.5	0
Dominos	-0.0130	-0.5	0.59	-0.0186	-0.5	0.64	-0.0539	-1.4	0.16	0.5214	10.9	0
Pizza Pizza	0.1852	8.8	0.00	0.0188	0.4	0.71	0.3452	9.3	0	0.4352	8.3	0
Ginos	-0.2451	-12.2	0.00	0.0066	0.2	0.84	-0.4911	-10.5	0	1.0571	21.6	0
Price = \$10	0.5271	26.1	0.00	0.0030	0.1	0.94	1.1143	27.9	0	1.4738	27.4	0
Price = \$12	0.2793	14.1	0.00	0.0647	2.0	0.05	0.4972	14.4	0	0.0875	1.8	0.07
Price = \$14	-0.1155	-6.1	0.00	0.0416	1.0	0.31	-0.2608	-7.7	0	0.2356	3.9	0
Price = \$16	-0.6909	-31.9	0.00	-0.1093	-2.9	0.00	-1.3507	-26.2	0	1.1507	21.9	0
Regular Crust	0.1225	2.8	0.01	-0.0057	-0.1	0.89	0.1811	3.5	0	0.0232	0.4	0.66
Thin Crust	-0.0939	-2.6	0.01	-0.0796	-1.2	0.22	-0.1373	-2.4	0.02	0.8573	14.9	0
Thick Crust	-0.0353	-1.0	0.31	-0.1055	-1.8	0.07	-0.1251	-2.4	0.02	0.2256	3.7	0
Pan Crust	-0.0026	-0.1	0.96	-0.0265	-0.4	0.66	-0.0706	-1.4	0.17	0.0109	0.2	0.86
Crust Missing	0.0093	0.2	0.82	0.2173	3.8	0.00	0.1519	2.9	0	0.6440	11.1	0
1 Topping	-0.8064	-23.0	0.00	-0.1455	-2.8	0.01	-1.3256	-19.5	0	1.2090	17.7	0
2-3 Toppings	0.1248	4.0	0.00	0.0063	0.2	0.88	0.2210	4.1	0	0.4721	6.5	0
4-5 Toppings	0.2484	8.0	0.00	0.0433	0.9	0.38	0.4222	8.0	0	0.1815	3.0	0
6 or more	0.4333	12.2	0.00	-0.1066	-2.4	0.02	0.5819	9.9	0	0.5960	9.9	0
Toppings Missing	-0.0001	0.0	1.00	0.2025	4.3	0.00	0.1005	1.7	0.08	0.9036	13.9	0
10 min Del. Time	0.5432	14.9	0.00	-0.1111	-2.4	0.02	0.8469	14.6	0	0.5783	10.4	0
20 min Del. Time	0.2530	7.7	0.00	-0.0394	-0.8	0.44	0.4201	8.2	0	0.0459	0.7	0.46
30 min Del. Time	-0.0838	-2.7	0.01	0.0294	0.5	0.60	-0.1359	-2.7	0.01	0.2329	4.0	0
40 min Del. Time	-0.7195	-21.1	0.00	-0.0893	-1.8	0.08	-1.1757	-19.2	0	0.8484	14.7	0
Del. Time Missing	0.0071	0.2	0.83	0.2104	4.1	0.00	0.0446	0.8	0.42	0.0087	0.2	0.88
Free Del. = No	-0.1556	-7.9	0.00	-0.0001	0.0	1.00	-0.2695	-6.8	0	0.0481	0.9	0.36
Free Del. = Yes	0.1480	7.4	0.00	-0.0606	-2.1	0.04	0.2795	7.0	0	0.2552	4.3	0
Free Del. Missing	0.0076	0.4	0.70	0.0607	2.1	0.04	-0.0100	-0.3	0.80	0.3033	5.4	0
Bread = No	-0.0996	-5.1	0.00	-0.0489	-1.6	0.12	-0.2077	-5.5	0	0.0310	0.5	0.59
Bread = Yes	0.1041	4.9	0.00	-0.0977	-3.1	0.00	0.1343	3.5	0	0.2053	3.5	0
Bread Missing	-0.0045	-0.2	0.83	0.1466	4.7	0.00	0.0734	1.9	0.06	0.1743	3.0	0
Wings = No	-0.1275	-6.8	0.00	0.0463	1.5	0.13	-0.2557	-6.0	0	0.5401	10.6	0
Wings = Yes	0.1406	6.5	0.00	-0.1361	-4.4	0.00	0.1959	4.6	0	0.6048	11.5	0
Wings Missing	-0.0131	-0.7	0.52	0.0898	2.9	0.00	0.0598	1.4	0.16	1.1449	22.0	0
Salad = No	-0.0627	-3.1	0.00	-0.0300	-1.0	0.34	-0.0469	-1.2	0.24	0.0893	1.7	0.09
Salad = Yes	0.0486	2.4	0.02	-0.0493	-1.6	0.11	0.1473	3.7	0	0.0626	1.5	0.15
Salad Missing	0.0141	0.7	0.48	0.0793	2.6	0.01	-0.1004	-2.5	0.01	0.1519	3.1	0
Intercept	0.1996	17.3	0.00	-0.0692	-2.6	0.01	0.3533	5.9	0	0.1559	3.0	0

Appendix B: Detailed Model Estimation of Juice Results

Parameters	CHM Model						Mixed Logit					
	Systematic Component			Scale Component			Systematic Component			Variance Component		
	Est. (β)	t-stat	prob.	Est. (α)	t-stat	prob.	Est. (β)	t-stat	prob.	Est. (ϖ)	t-stat	prob.
Minute Maid	0.0404	2.8	0.01	0.0109	-1.5	0.13	0.1036	3.1	0.00	0.7790	15.5	0.00
Tropicana	0.1223	8.2	0.00	-0.0580	-1.5	0.13	0.2811	8.3	0.00	0.3061	7.1	0.00
McCain	-0.0510	-3.6	0.00	0.0136	0.5	0.65	-0.1060	-3.3	0.00	0.0717	1.1	0.29
Del Monte	-0.1110	-7.9	0.00	0.0335	0.8	0.40	-0.2790	-8.0	0.00	0.4008	9.9	0.00
Price = \$0.30	0.2487	17.2	0.00	0.0293	0.8	0.46	0.6175	17.6	0.00	0.8030	17.5	0.00
Price = \$0.40	0.1204	8.7	0.00	0.0443	1.0	0.30	0.2694	8.6	0.00	0.0170	0.4	0.66
Price = \$0.50	-0.0520	-3.7	0.00	0.0213	0.6	0.58	-0.1630	-5.2	0.00	0.0020	0.0	0.97
Price = \$0.60	-0.3170	-20.5	0.00	-0.0950	-2.6	0.01	-0.7240	-17.0	0.00	0.8221	19.2	0.00
Juice Type -- Apple	-0.0040	0.4	0.66	0.0609	2.8	0.01	-0.0290	-1.3	0.21	0.3963	13.0	0.00
Juice Type -- Orange	0.0036	0.4	0.66	-0.0610	-2.8	0.01	0.0286	1.3	0.21	0.3963	13.0	0.00
Prep (concentrate)	-0.1030	-7.7	0.00	0.0031	0.1	0.93	-0.2340	-5.8	0.00	0.4110	8.7	0.00
Prep (fresh)	0.1041	7.7	0.00	-0.0630	-1.7	0.08	0.2057	5.3	0.00	0.2600	4.4	0.00
Prep Missing	-0.0010	-0.1	0.92	0.0603	1.7	0.08	0.0281	0.7	0.48	0.6710	12.7	0.00
10% real juice	-0.7520	-27.7	0.00	0.0518	1.1	0.26	-1.3760	-18.0	0.00	1.4297	18.7	0.00
40% real juice	-0.3390	-13.9	0.00	0.1069	1.5	0.13	-0.5300	-9.7	0.00	0.5400	7.1	0.00
70% real juice	0.3508	12.0	0.00	-0.1030	-1.8	0.07	0.7805	14.6	0.00	0.3132	5.7	0.00
100% real juice	0.8367	21.6	0.00	-0.4480	-9.2	0.00	1.1723	21.0	0.00	0.5109	7.2	0.00
% real juice Missing	-0.0970	-3.2	0.00	0.3925	7.1	0.00	-0.0470	-0.8	0.43	2.7940	40.2	0.00
Number of Packs (3)	-0.1680	-6.5	0.00	-0.2090	-3.0	0.00	-0.3340	-7.0	0.00	0.0205	0.4	0.72
Number of Packs (6)	-0.0340	-1.5	0.14	-0.0560	-0.8	0.42	-0.1040	-2.2	0.03	0.0150	0.2	0.83
Number of Packs (9)	0.0206	0.9	0.36	-0.0340	-0.4	0.73	0.0055	0.1	0.91	0.0375	0.5	0.60
Number of Packs (12)	0.1111	5.2	0.00	0.0960	1.6	0.11	0.2650	5.5	0.00	0.0350	0.4	0.73
No. of Packs Missing	0.0706	3.0	0.00	0.2020	2.7	0.01	0.1675	3.5	0.00	0.0080	0.1	0.91
Vit. C = Not Added	-0.0770	-5.6	0.00	-0.0710	-1.5	0.13	-0.1530	-4.0	0.00	0.2445	4.6	0.00
Vit. C = Added	0.0953	7.2	0.00	0.0054	0.1	0.92	0.2089	5.3	0.00	0.0940	1.5	0.13
Vit. C Missing	-0.0180	-1.3	0.18	0.0659	1.3	0.20	-0.0560	-1.4	0.15	0.1500	2.6	0.01
Sugar = Unsweetened	-0.0120	-0.8	0.42	-0.0590	-1.6	0.12	-0.0120	-0.3	0.78	0.6608	12.5	0.00
Sugar = Sweetened	-0.0420	-2.8	0.01	-0.0980	-2.6	0.01	-0.1180	-2.9	0.00	0.5322	7.9	0.00
Sugar Missing	0.0538	3.6	0.00	0.1567	4.2	0.00	0.1301	3.1	0.00	1.1930	19.8	0.00
Calcium = Not Added	-0.0470	-3.5	0.00	-0.0490	-1.3	0.20	-0.0310	-0.8	0.43	0.2074	3.7	0.00
Calcium = Added	0.0345	2.7	0.01	0.0494	1.3	0.19	0.1104	2.8	0.00	0.1849	2.9	0.00
Calcium Missing	0.0128	1.0	0.34	-0.0010	0.0	0.99	-0.0790	-2.0	0.04	0.3920	6.5	0.00
Packaging (Glass)	0.2860	12.4	0.00	0.0291	0.5	0.65	0.5913	11.5	0.00	0.3080	4.2	0.00
Packag. (Aluminum)	-0.0290	-1.3	0.20	0.0124	0.2	0.84	-0.1360	-2.5	0.01	0.5842	10.6	0.00
Packaging (Tetra)	0.0791	3.1	0.00	-0.1470	-2.4	0.02	0.1532	2.9	0.00	0.5233	8.7	0.00
Packag. (Plastic Bag)	-0.4070	-15.3	0.00	-0.1550	-2.8	0.01	-0.8600	-14.0	0.00	0.9452	16.9	0.00
Packaging Missing	0.0714	2.9	0.00	0.2607	4.3	0.00	0.2518	4.6	0.00	1.745	28.4	0.00
Intercept	0.0061	0.7	0.47	0.1151	4.8	0.00	-0.1320	-2.3	0.02	0.1815	3.2	0.00

Appendix C: Simulation Design

We simulated choices based on experiments with four attributes. Attributes A and B were always present, but C and D were varied. Attribute C was assumed more important than D. This results in four conditions (cond 1: 4 profiles with A and B varied; cond 2: 8 profiles with A, B, and D varied; cond 3: 8 profiles with A, B, and C varied; cond 4: 16 profiles with A, B, C and D varied). The mean coefficients for the data generation process (DGP) are $\beta_A = 0.15$, $\beta_B = -0.70$, $\beta_C = 1.0$, and $\beta_D = 0.10$. Our choice of important and non-important parameter values are based on theoretical work on best-worst scaling by Marley and Louviere (2005). For example, if there are 4 attributes in a particular set and a consumer reports their most and least preferred options, the relative importance of the 4 attributes are 8, 4, 2 and 1 from most to least preferred attributes. The importance weight of least important attribute is $1/8 (= 0.12)$ relative to the most important attribute. Our choice of $\beta_C = 1.0$ and $\beta_D = 0.10$ is based on that theoretical reasoning. The preference heterogeneity is assumed “zero”, “low” (standard deviation of preference heterogeneity, $\varpi = \beta$, i.e., coefficient of variation, $CV = 1$), and “high” ($\varpi = 2\beta$, i.e., $CV = 2$). Our choice of $CV = 0, 1$, and 2 is based on empirical findings of the most frequent ratios from 9 published articles from different disciplines (e.g., Revelt and Train 1998; a complete list is available from authors). Condition 4 is used as the reference condition to calculate relative scales. The data generation process is $U_{in} = \beta X_i + e_{in}$. Here e_{in} is distributed as extreme value. Following Train (2003, p 210), $e = -\ln(-\ln \mu)$, where μ is a draw from uniform distribution. The variance of the error term is $\pi^2/6\lambda^2 = \pi^2/6$ (assumed scale parameter $\lambda = 1$). Two scale values were used ($\lambda = 1$ or 3). Andrews, Ainslie, and Currim (2002) simulation study, which investigated merits of discrete and continuous representation of preference heterogeneity, used variances 1:4 (equivalent to scale 1 and 2). Our empirical scale estimates from current data varied from 1 to 3 (31 scale estimates for Pizza and 31 for Juice; 32 conditions for each), and we observed that scales varied from 1 to 3 in other research projects. Following estimation of mixed logit model, each individual’s posterior prediction was made using Bayes rule.

Figure 1a (Pizza): Random Component Estimates vs. Attribute Levels

Legend: □ = Mixed Logit ● = CHM

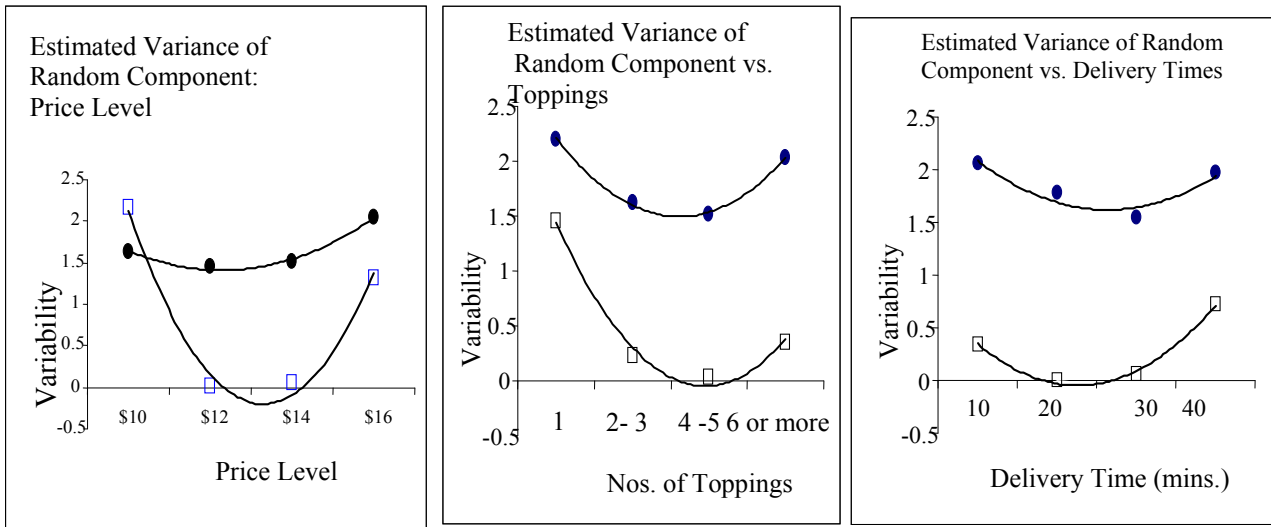


Figure 1b (Juice): Random Component Estimates vs. Attribute Levels

Legend: □ = Mixed Logit ● = CHM

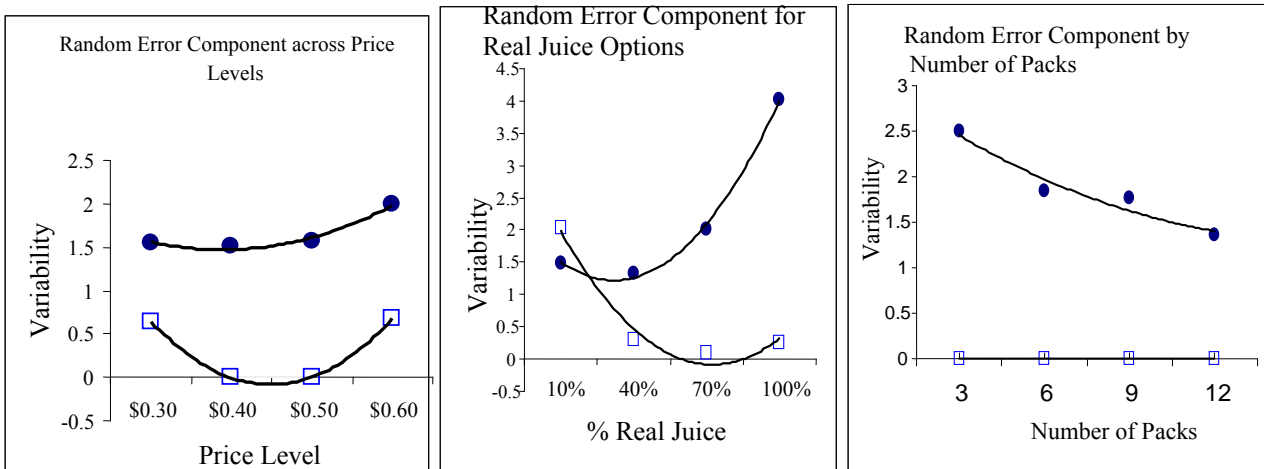


Table 1: Proportion who would buy the product under Presence/Absence of Attributes

Pizza Attributes	Mean proportion that “would buy”		Test of Equality	
	Missing	Present	Variances (Sig.) [^]	Means (Sig.) [†]
Number of Toppings	0.577	0.563	0	0.057
Delivery Time	0.587	0.553	0	0
Free Delivery	0.563	0.576	0	0.074
Garlic Bread/Sticks	0.563	0.577	0	0.057
Chicken Wings	0.571	0.569	0.592	0.789
Salad	0.562	0.578	0	0.026
Crust Type	0.577	0.563	0	0.053

Juice Attributes	Mean proportion that “would buy”		Test of Equality	
	Missing	Present	Variances (Sig.) [^]	Means (Sig.) [†]
Concentrate/Fresh	0.507	0.508	0.814	0.906
% of real juice	0.509	0.506	0.388	0.659
Number of Packs	0.521	0.494	0	0
Vitamin C	0.506	0.509	0.486	0.724
Sugar	0.518	0.497	0	0.005
Calcium	0.514	0.501	0.010	0.099
Packaging Materials	0.520	0.495	0.001	0.001

[^] - Levene’s Test for equality of variances (two tailed significance reported). Null hypothesis is equality of variances.

[†] - Independent Sample t-test for equality of means (two tailed significance reported)

Table 2: Percent Correct In and Out of Sample Predictions

		In Sample			Out of Sample		
		MNL	CHM	MIXL	MNL	CHM	MIXL
Model estimated with first set of 16 conditions.	Pizza	0.780	0.807	0.862	0.672	0.731	0.694
	Juice	0.719	0.789	0.825	0.588	0.600	0.667
Model estimated with second set of 16 conditions.	Pizza	0.753	0.803	0.855	0.706	0.736	0.721
	Juice	0.694	0.744	0.801	0.620	0.647	0.660

Table 3a: Pizza Model Parameters

Parameters	CHM Model					Mixed Logit					
	Systematic Component		Scale Component ¹			Systematic Component		Variance Component ²			
	Est. (β)	Prob.	Est. (α)	Prob.		Est. (β)	Prob.	Est. (ϖ)	Prob.		
Crust Missing	0.0093	0.82	0.2173	0.00	**	0.1519	0.00	**	0.644	0.00	**
Toppings Miss.	-0.0001	0.99	0.2025	0.00	**	0.1005	0.09		0.9036	0.00	**
Del. Time Miss.	0.0071	0.83	0.2104	0.00	**	0.0446	0.42		0.0087	0.88	
Free Del. Miss.	0.0076	0.70	0.0607	0.04	*	-0.0100	0.80		0.3033	0.00	**
Bread Miss.	-0.0045	0.83	0.1466	0.00	**	0.0734	0.06		0.1743	0.00	**
Wings Miss.	-0.0131	0.52	0.0898	0.00	**	0.0598	0.16		1.1449	0.00	**
Salad Miss.	0.0141	0.48	0.0793	0.01	*	-0.1004	0.01	*	0.1519	0.00	**

* significant at 0.05 level; ** significant at 0.01 level

¹ Positive α estimates indicate decrease in variability. Scale component $\lambda = \exp(\alpha) = \text{SQRT}(\pi^2/6\sigma_e^2)$.

Table 3b: Juice Model Parameters

Parameters	CHM Model					Mixed Logit						
	Systematic Component		Scale Component			Systematic Component		Variance Component				
	Est. (β)	Prob.	Est. (α)	Prob.		Est. (β)	Prob.	Est. (ϖ)	Prob.			
Prep Miss.	-0.0014	0.92	0.0603	0.08		0.0261	0.48		0.6710	0.00	**	
Realjuice Miss.	-0.0965	0.00	**	0.3925	0.00	**	-0.0471	0.43		2.7938	0.00	**
N_Pack Miss.	0.0706	0.00	**	0.2020	0.01	**	0.1675	0.00	**	0.0083	0.91	
Vitamin C Miss.	-0.0180	0.18		0.0659	0.20		-0.0559	0.15		0.1503	0.01	**
Sugar Miss.	0.0538	0.00	**	0.1567	0.00	**	0.1301	0.00	**	1.1930	0.00	**
Calcium Miss.	0.0128	0.34		-0.0007	0.99		-0.0791	0.04	*	0.3923	0.00	**
Pkg Mat. Miss.	0.0714	0.00	**	0.2607	0.00	**	0.2518	0.00	**	1.7450	0.00	**

* significant at 0.05 level; ** significant at 0.01 level