ERIC T. BRADLOW, YE HU, and TECK-HUA HO*

In this note, the authors propose several extensions of the model of consumer learning in conjoint analysis that Bradlow, Hu, and Ho (2004) develop. They present a clarification of the original model, propose an integration of several new imputation rules, add new measurement metrics for pattern matching, and draw a roadmap for further real-world tests. The authors also discuss general modeling challenges when researchers want to mathematically define and integrate behavioral regularities into traditional quantitative domains. They conclude by suggesting several critical success factors for modeling behavioral regularities in marketing.

Modeling Behavioral Regularities of Consumer Learning in Conjoint Analysis

We welcome the constructive comments on our article (Bradlow, Hu, and Ho 2004; hereinafter BHH) by Alba and Cooke (2004), Rao (2004), and Rubin (2004). Because a major goal of our original article is to enrich conjoint analysis with a stronger behavioral foundation, we are pleased to hear from our colleagues in marketing, all of whom have both behavioral modeling and quantitative interests, and from Rubin, who first introduced the formal nomenclature of missing data methods to the statistics literature (Rubin 1976). We believe that such dialogue enables us to harness the strengths of various research paradigms and to make marketing theories more precise and predictive of actual consumer behavior.

We organize our responses to the three comments into four subsections. The first section includes general responses that touch on the issues of research language and mathematical formalism, and the last three sections are specific responses to the comments in terms of clarification, additional data analyses, and model extensions.

MODEL SIMPLICITY, RESEARCH LANGUAGE, AND MATHEMATICAL SPECIFICATION

By definition, a model is an approximate description of the real world. The degree of abstraction depends on the modeler's taste for simplicity and his or her research goals. We illustrate this point with the abstract task of drawing a map of the world. A map with every detail of the world is no longer a useful map, because it must be as big as the real world. A good map keeps only the important information from reality (e.g., direction, landmarks) and ignores trivial reality (e.g., a parking meter's location). The same analogy applies to modeling. A good model abstracts only what is significant and disregards "unnecessary" details of reality. In this regard, Little (1970) and Leeflang and colleagues (2000) provide excellent perspectives on the pros and cons of modeling. The a priori beliefs about the degree of significance of specific details and the choice of which details to include depend on the goal of the research. A major, if not singular, goal of our original article is to nest current extant models of imputation and test their relative predictive power. This goal considerably constrains our design and development of the proposed model. Consequently, BHH do not capture every single detail of an imputation process and should not have, given the research goal. That is, as the discussants pointed out, BHH do not incorporate other empirical regularities that have been shown to exist.

However, showing the existence of an important empirical regularity is a necessary but not sufficient condition for incorporating it into a formal model. To incorporate an empirical regularity into a formal model, the regularity must be specified in mathematical language (see, e.g., Camerer and Ho 1999; Camerer, Ho, and Chong 2004). This formal specification of an empirical regularity is in no way trivial. For example, it is a challenge to capture explicitly the cognitive effort required in a conjoint analysis task as a decision variable that respondents may choose to minimize; however, it would be indicative of the realism of the model if response times were to violate a principle on which the model is based. For example, although on the one hand profiles with fewer attributes may be easier to rate and process, as our data suggest, on the other hand, as Alba and Cooke

*Eric T. Bradlow is Associate Professor of Marketing and Statistics and Academic Director of the Wharton Small Business Development Center, The Wharton School, University of Pennsylvania (e-mail: ebradlow@wharton.upenn.edu). Ye Hu is a visiting assistant professor, Krannert School of Management, Purdue University (e-mail: hu18@purdue.edu). Teck-Hua Ho is William Halford Jr. Family Professor of Marketing, Haas School of Business, University of California, Berkeley (e-mail: hoteck@haas.berkeley.edu).

Journal of Marketing Research
Vol. XLI (November 2004), 392–396

(2004) point out, if imputation is effortful, response times may increase with missing attributes. This trade-off is unclear and is an interesting issue for additional testing; it might be teased out by a properly designed experiment built for that purpose. Similarly, we perceive no easy way to specify unprompted inferences mathematically. Therefore, the points raised by Alba and Cooke (2004) and Rao (2004) about the BHH model are essential if the model is to be wisely extended.

CLARIFICATIONS OF THE ORIGINAL MODEL

We clarify and examine some of the points raised by Alba and Cooke (2004), Rao (2004), and Rubin (2004). First, Alba and Cooke (2004) point out that if the respondents were well-informed, they would have figured out that the experimental design was an orthogonal one, and thus learning would not have been their explicit objective. Consequently, they would not exhibit imputation inference behavior. This conjecture is reasonable and may hold in certain contexts; however, it is testable by means of our model and data. If this suggestion were true, the results would have favored an "ignoring missing attribute(s)" model. We do not find support for this conjecture in either our in-sample or out-of-sample estimation results (see Table 5 in BHH). In general, our results suggest that subjects do not ignore missing attribute levels and that the BHH model provides a way for subjects to use available information to infer these missing levels.

Second, Rao (2004) notes that at least one of the earlier profiles must be complete (and not partial) for the imputation process to occur. In contrast to this concern, we note that there are two information sources for imputing missing attributes: the prior and the previously shown profiles. Thus, (1) even if none of the previously shown profiles is complete, "proper" imputation can occur on the basis of the prior counts, and (2) without the prior (albeit this is not our model), it would be most accurate to note that each missing (i.e., to be imputed) attribute must have appeared at least once before. With regard to the second point and in practice, because the missing attribute(s) in a partial-profile conjoint analysis usually rotate, the allowable imputation process would start fairly early. In addition, in our Study 1, we expose subjects to complete profiles in the learning phase, so the imputation process always can immediately occur.

Rao (2004) also calls for further study on the managerial importance of the BHH model and its prospects of application in different real-world scenarios. Because of the wide application of conjoint analysis in marketing research, there are plenty of examples for which the BHH model is relevant. For example, Ford Motor Company recently adopted "Vehicle Advisor," a procedure akin to adaptive conjoint analysis, to help consumers make vehicle choices (Figure 1). After consumers choose their preferred basic functionality, software provides them with a list of pairwise vehicles to compare. The side-by-side comparison (Figure 2) uses only a small subset of vehicle features. The BHH model applies to this example, especially when consumers make multiple comparisons. We believe that testing the BHH model with similar real-world examples is an important step toward its application in industry practice.

Third, Rubin (2004) points out that a complex model, such as the BHH model, is a prime candidate for posterior predictive checks (Rubin 1984; Rubin and Stern 1994) that can simulate data from the posterior predictive distribution. We wholeheartedly agree with this suggestion and note that

Figure 1
FORD MOTOR COMPANY VEHICLE ADVISOR: APPLICATION OF BHH

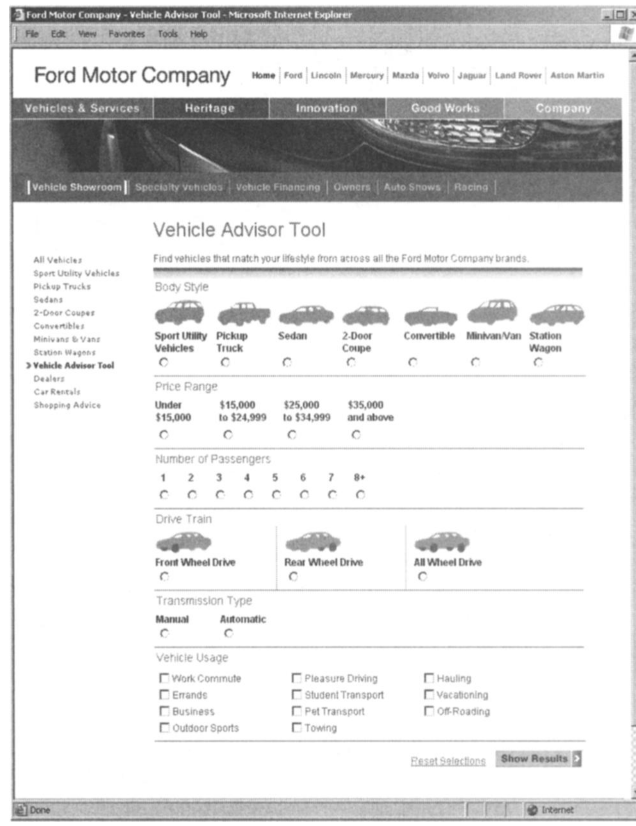
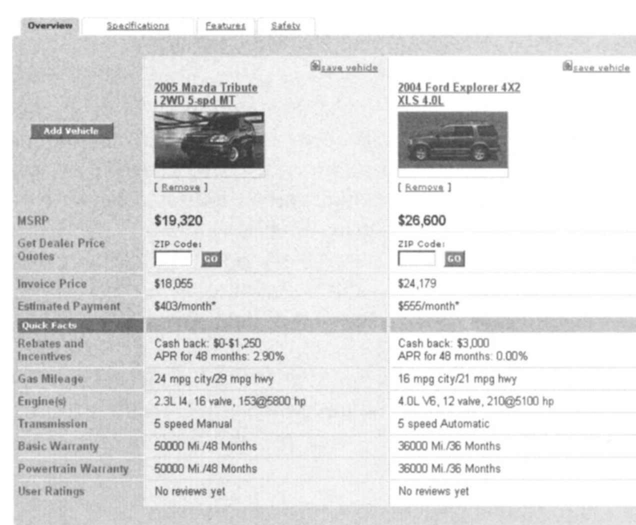


Figure 2
SIDE-BY-SIDE COMPARISON BY FORD VEHICLE ADVISOR



our out-of-sample fit assessment of the model was obtained by drawing holdout conjoint ratings from the model's predictive distribution directly within our Markov chain Monte Carlo sampler. However, we concede Rubin's (2004) point that the more common form of posterior predictive checks in the statistics literature assesses the features of the model (out-of-sample prediction is one on them), which would be a nice dimension to consider in further research.

ADDITIONAL DATA ANALYSES

Alba and Cook (2004) suggest that we provide additional support for our model by showing that the other decay parameters (λ_{i1}–λ_{i5}) do not correlate with the manipulated prior condition between price and maximum resolution. In Table 1, we show the results from such a correlation analysis. We do not find this relationship with the other decay parameters (Table 1), in support of our model.

MODEL EXTENSIONS

Both Alba and Cooke (2004) and Rao (2004) suggest that the imputation process could have sources of information other than historical attribute levels. For example, Rao (2004) suggests a different imputation process in which people impute missing values on the basis of their importance or the partworths themselves. This suggestion is intriguing and would be fairly easy to implement if we simply modeled the following:

$$\text{(1)} \qquad \text{logit}\{\Pr[x'_{ij}(t) = 1]\} = \alpha_i + \kappa\beta_{ij} + \varepsilon_{ij},$$

where α_i represents the baseline propensity of a person to impute a higher level, β_{ij} is the partworth of respondent i to attribute j, κ is a slope parameter, and ε_{ij} is a stochastic error term. This implementation is also related to Alba and Cooke's (2004) comment about evaluative consistency, in which profiles that perform well on attributes with higher relative importance tend to have the missing attributes imputed higher. This consistency could be accomplished within our framework by modeling the following:

$$\text{(2)} \qquad \text{logit}\{\Pr[x'_{ij}(t) = 1]\} = \alpha_i + \kappa\beta_{ij}x_{ij} + \varepsilon_{ij},$$

where x_{ij} is the vector of observed levels in the current profile. Whether the subjects use attributes with high importance more is an empirical question. Although such extensions enrich the behavioral versatility of the BHH model, they also may lead to potential estimation issues because of higher-order interaction terms involving the βs. However, because the imputation parameters (ωs) in BHH are not restricted, they should be able to capture, at least partially, effects such as which attribute gets more weight in imputation.

Some additional issues raised about our model on various dimensions include the choice of independent values for the prior experience counts $N_{ij}(0|l_j)$, the ability to handle more than two attribute levels, symmetry in the imputation model, and the ability to impute values outside the observed set. We address each of these points in turn.

First, although the prior experience counts $N_{ij}(0|l_j)$ are specified separately for each attribute level, this does not imply that the prior correlation structure is ignored. If two or more attribute levels are highly (weakly) correlated, both their prior experience counts will be simultaneously high (low). Therefore, if both are missing, both would be more likely based on their a priori values to be imputed as the higher level. Note that the model we use,

$$\text{(3)} \qquad N_{ij}(0|l_j) \sim \text{Poisson}[\exp(\varsigma_i + \omega_j)],$$

allows for the possibility that the ω_js are correlated, even though they are drawn i.i.d. from a common prior.

Second, the handling of more than two attribute levels is an important issue to bring our model closer to practical usage. There is no restriction in our framework that two attribute levels are necessary, and our model for the imputation of a given level could be extended so that the probability that a given level is imputed is given by the following:

$$\text{(4)} \quad \Pr[x'_{ij}(t) = l_j] = \begin{cases} 1 & \text{if } r_{ij}(t) = 1 \text{ and } x'_{ij}(t) = l_j \\ 0 & \text{if } r_{ij}(t) = 1 \text{ and } x'_{ij}(t) \neq l_j \\ \dfrac{N_{ij}(t|l_j)}{\sum_{l_j} N_{ij}(t|l_j)} & \text{if } r_{ij}(t) = 0, \end{cases}$$

which is a direct extension of Equation 4 in BHH. Recall that BHH use a binary Hamming matching metric, where $I[x_{ij}(t'), x_{ij}(t)]$ equals 1 for a match and equals 0 for a nonmatch. For nominal variables (e.g., gender, color), binary matching may be the best way to measure how similar two attribute levels are; they are either the same or different with no measurement in between. For continuous variables (e.g., price in dollars), if the linear assumption holds, instead of binary matching, a Euclidean distance $d[x_{ij}(t'), x_{ij}(t)]$, normalized with respect to the range of attribute j, could be easily defined between two attribute levels (e.g., \$4 is more similar to \$5 than to \$3). Specifically, we define the following:

$$\text{(5)} \qquad d[x_{ij}(t), x_{ij}(t')] = \frac{|x_{ij}(t') - x_{ij}(t)|}{\underset{\forall k_j, k_j^0;\, k_j \neq k_j^0}{\text{MAX}} \left| l_{jk_j} - l_{jk_j^0} \right|},$$

where k_j represents level k of attribute j. A more common case in conjoint analysis is to treat continuous variables as ordinals with multiple levels (e.g., price is treated as discrete rather than continuous). In this case, we could choose one of the three similarity measurements—a Euclidean distance, an equally spaced distance between

Table 1

CORRELATION BETWEEN λ AND MANIPULATED PRICE–RESOLUTION COVARIANCE

		λ_{i1}	λ_{i2}	λ_{i3}	λ_{i4}	λ_{i5}
One missing	Correlation	–.054	–.057	.049	–.169	.059
	(*p*-value)	(.719)	(.704)	(.743)	(.256)	(.693)
Two missing	Correlation	.003	.061	.050	–.095	–.019
	(*p*-value)	(.987)	(.697)	(.752)	(.546)	(.903)

levels, or a binary Hamming level matching—and empirically validate which is more likely to hold. At the same time, Alba and Cooke's (2004) concern that it may be easier for subjects to impute when all attributes are binary is legitimate. We believe it is an interesting and open empirical question. However, we could argue the opposite, because when more attribute levels are provided in a conjoint analysis, it is easier for subjects to impute a non-base level for the missing presence-manipulated attribute, which would lead to results different from the no-imputation model assumption in which the base level is assumed for a missing attribute.

Third, with regard to symmetry and the example brought up by Alba and Cooke (2004), in the particular case given in BHH's Figure 1, Panel B, $N_{i2}(3|1) = N_{i3}(3|1)$ if Attributes 2 and 3 had been reversed at Time 3. However, we note that $N_{i2}(3|0) \neq N_{i3}(3|0)$ in these cases, and thus the probabilities would not be the same. In the case given in BHH, the probability of imputing a 1 is $(\lambda_2 + \lambda_3)/(\lambda_2 + \lambda_3 + \lambda_2^2)$, whereas in this new case, it would be 1 because there would be no matches if a 0 was imputed. Nevertheless, even though the probabilities are different, it would be an interesting extension, as Alba and Cooke (2004) suggest, to create an extended model with $\lambda_{i(j)j'}$, where (j) denotes the focal item being imputed and j′ is the information source. This formulation originally appeared in the BHH model; however, we simplified it for parsimony. If a model with focal weights were applied, an increased parameter space would lead to the need for more careful thought about the design of the conjoint study (e.g., number of profiles).

Fourth, we consider the issue about imputing outside the observed set of values. We agree that our discrete contingency table formulation has, at its core, a kernel at which a discrete number of levels is imputed. In addition, in the current formulation, it will be within the observed set. An extension of our model in which the set of possible levels is specific to trial t and/or prior information observed, as long as that set remains discrete, is possible within the Bayesian paradigm and could be inferred along with the missing attribute values. This extension would represent a significant change to the model, which we would wholly endorse.

CONCLUSION

We take this response as an opportunity to further highlight our attempt to integrate behavioral research findings into a traditional quantitative domain (i.e., partial-profile conjoint analysis) in marketing. When modeling consumer imputation, we consider two key aspects: the source of information and the process. The most likely sources of information for imputation in a partial-profile conjoint analysis include consumer prior knowledge and the product profiles provided in the task. In general, the former is not observable in a conjoint task but can be modeled with latent variables, as in BHH. The latter is what we used directly in BHH. Behavioral research (as suggested by Alba and Cooke 2004; Rao 2004) has documented sufficient findings related to how consumers might make inferences in a partial-profile conjoint task. Our research goals, the taste for simplicity, and the lack of any other existing formal specification prompted us to focus on capturing only a limited set of relevant regularities and nesting commonly used extant models. Furthermore, research in the fields of psychology and educational testing (Bradlow and Thomas 1998; see Rubin 2004) has considered a conceptually similar problem of inferring levels (or abilities in the educational testing case) on the basis of missing information or responses. We believe that the following elements are crucial to continue to bridge behavioral and quantitative research, as we have attempted:

1. *Mathematical formalism.* To incorporate behavioral regularities into standard models, these regularities must be expressed in mathematical terms: "What is its representation?" Ideally, this question should be answered by a joint effort of both behavioral researchers and modelers, which is why the subsequent point is important.
2. *Interdisciplinary collaboration.* As modelers, we could have benefited greatly from working with a colleague with greater behavioral training. Such collaboration may have made the proposed model more realistic without increasing its complexity. Previous research of a similar nature (e.g., Bodapati and Drolet 2003; Hardie, Johnson, and Fader 1993; Kahn and Raju 1991; Kivetz, Netzer, and Srinivasan 2004) has benefited from such interdisciplinary collaboration and skills. As Rubin (2004) points out, statisticians provide many tools, which in this case include the concepts of latent variable modeling and classic fractional factorial design, to bond behavioral theories with mathematical modeling. We regard collaborations not only between behavioral researchers and modelers within the marketing domain itself but also across different fields (e.g., economics, operations, psychology, sociology, statistics) as a way to undertake challenging and important research in marketing in the future.

REFERENCES

Alba, Joseph W. and Alan D.J. Cooke (2004), "When Absence Begets Inference in Conjoint Analysis," *Journal of Marketing Research,* 41 (November), 382–87.

Bodapati, Anand V. and Aimee Drolet (2003), "A Hybrid Choice Model that Uses Absolute and Ordinal Value Information," working paper, Anderson School, University of California, Los Angeles.

Bradlow, Eric T., Ye Hu, and Teck-Hua Ho (2004), "A Learning-Based Model for Imputing Missing Levels in Partial Conjoint Profiles," *Journal of Marketing Research,* 41 (November), 369–81.

——— and N. Thomas (1998), "Item Response Theory Models Applied to Data Allowing Examinee Choice," *Journal of Educational and Behavioral Statistics,* 23 (3), 236–43.

Camerer, Colin F. and Teck-Hua Ho (1999), "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica,* 67 (4), 837–74.

———, ———, and Juin-Kuan Chong (2004), "A Cognitive Hierarchy Model of Games," *Quarterly Journal of Economics,* 119 (3), 861–98.

Hardie, Bruce G.S., Eric J. Johnson, and Peter S. Fader (1993), "Modeling Loss Aversion and Reference Dependence Effects on Brand Choice," *Marketing Science,* 12 (4), 378–94.

Kahn, Barbara E. and Jagmohan S. Raju (1991), "Effects of Price Promotions on Variety-Seeking and Reinforcement Behavior," *Marketing Science,* 10 (4), 316–37.

Kivetz, Ran, Oded Netzer, and V. Srinivasan (2004), "Alternative Models for Capturing the Compromise Effect," *Journal of Marketing Research,* 41 (August), 237–57.

Leeflang, Peter S.H., Dick R. Wittink, Michel Wedel, and Philippe A. Naert, eds. (2000), *Building Models for Marketing Decisions.* New York: Kluwer Academic Publishers.

Little, John D.C. (1970), "Models and Managers: The Concept of a Decision Calculus," *Management Science,* 16 (8), B466–85.

Rao, Vithala R. (2004), "Comments on Conjoint Analysis with Partial Profiles," *Journal of Marketing Research,* 41 (November), 388–89.

Rubin, Donald B. (1976), "Inference and Missing Data," *Biometrika,* 63 (3), 581–92.

——— (1984), "Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician," *Annals of Statistics*, 12, 1151–72.

——— (2004), "Design and Modeling in Conjoint Analysis with Partial Profiles," *Journal of Marketing Research,* 41 (November), 390–91.

——— and H. Stern (1994), "Testing in Latent Class Models Using a Posterior Predictive Check Distribution," in *Latent Variables Analysis: Applications for Developmental Research*, A. Von Eye and C.C. Clogg, eds. London: Sage Publications, 420–38.