

Modeling Behavioral Regularities of Consumer Learning in Conjoint Analysis

Eric T. Bradlow, Ye Hu, and Teck-Hua Ho *

November 15, 2007

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

Modeling Behavioral Regularities of Consumer Learning in Conjoint Analysis

ABSTRACT

In this note, we propose several extensions of the model of consumer learning in conjoint analysis developed by Bradlow, Hu, and Ho (2004). We present 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. We discuss general modeling challenges when one wants to mathematically define and integrate behavioral regularities into traditional quantitative domains. We conclude by suggesting several critical success factors for modeling behavioral regularities in marketing.

We welcome the constructive comments on our paper (Bradlow, Hu and Ho 2004; BHH hereafter) by Alba and Cooke (2004), Rao (2004), and Rubin (2004). Since a major goal of our paper 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 into the statistics literature (Rubin 1976). We believe such dialogue will allow us to harness the strengths of varied research paradigms and make marketing theories more precise and predictive of actual consumer behavior.

We would like to organize our responses to the three comments along four subsections. The first section includes general responses that touch on the issues of research language and mathematical formalism and the last three 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 a modeler's taste for simplicity and her research goals. Let us illustrate this point with the (abstract) task of drawing a map of the world. A map with every single detail of the world is no longer a useful map, since it has to be as big as the real world. A good map keeps only the important information from reality (such as direction, landmarks, etc.) but ignores trivial reality such as a parking meter's location. The same analogy applies to modeling. A good model abstracts only what's significant but disregards "unnecessary" details of reality. In this regard, Little (1970) and Leeflang et al. (2000) both provide excellent perspectives on the pros and cons of modeling. The (a priori) beliefs in the degree of significance of those 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 paper is to nest current extant models of imputation and test their relative predictive power. This has considerably constrained our design and development of the proposed model. Consequently, BHH does not capture every single detail of an

imputation process and in fact it should not have, given the research goal. That is, as pointed out by the discussants, there are other empirical regularities that have been shown to exist that BHH does not incorporate.

However, showing the existence of an important empirical regularity is only a necessary but not sufficient condition for incorporating it into a formal model. To incorporate an empirical regularity into a formal model, it is necessary that the regularity be specified in mathematical language (see Camerer and Ho (1999) and Camerer et. al (2004) for recent examples). This formal specification of an empirical regularity is in no way a trivial. For example, it is a challenge to explicitly capture 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, while on the one hand, profiles with less attributes may be easier to rate and process as our data suggests, on the other hand, as pointed out by Alba and Cooke (2004), if imputation is effortful response times may increase with missing attributes. This trade-off is unclear and is an interesting issue for future testing; one which could be teased out by a properly designed future experiment built for that purpose. Similarly, we see no easy way to specify “unprompted” inferences mathematically. It is why 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. We clarify these points and discuss some of the possible extensions next.

CLARIFICATIONS OF THE ORIGINAL MODEL

We would like to clarify and examine some of the points raised by Alba and Cooke (2004), Rao (2004), and Rubin (2004).

- Alba and Cooke (2004) point out that if the respondents were “well-informed”, they would have figured out that the experimental design would be an orthogonal one and 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 this is testable using our model and data. Indeed, if this were true, the results would have favored an “ignoring missing attribute(s)” model. We do not find support for this conjecture in both our in- and out-of-sample estimation results (see Table 5 in BHH). In general, our results suggest subjects do not ignore missing attribute levels and the BHH model provides one way in which subjects use available information to infer these missing levels.

- Rao (2004) points out that at least one of the earlier profiles need to be complete (and not partial) for the imputation process to occur. We note, in contrast (counter) to this concern, that there are two information sources for imputing missing attributes: the prior and the previously shown profiles. Thus, (a) even if none of the previously shown profiles are complete, “proper” imputation can occur based on the prior counts; and (b) without the prior (albeit this is not our model), what would be most accurate would be to say that each attribute that is missing (to be imputed) must have appeared before at least once. Note, with regards to (b), and in practice, since the missing attribute(s) in a partial profile conjoint analysis are usually rotating, the “allowable” imputation process would start fairly early. In addition, in our study 1, since subjects are exposed to complete profiles in the learning phase, it is always the case that the imputation process can immediately occur.

- Rao (2004) also calls for further study on the managerial importance of BHH 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 where the BHH model is relevant. For instance, Ford Motor Company recently adopted “Vehicle Advisor”, a procedure akin to adaptive conjoint analysis, to help consumers make vehicle choices (Figure 1). After preference of basic functionality is given, the software provides the consumers with a list of pairwise vehicles to compare. The side-by-side comparison (Figure 2) uses only a small subset of vehicle features. Obviously, BHH applies to this example, especially when multiple comparisons are made by consumers. We believe testing BHH with real-world examples like this is a very important step towards its applications in industry practice.

[Insert Figure 1 and Figure 2 Here.]

- Rubin (2004) points out that a complex model such as the one utilized here is a “prime” candidate for the use of posterior predictive checks (Rubin 1984; Rubin and Stern 1994) that allows one to simulate data from the posterior predictive distribution. We wholeheartedly agree with this suggestion and point out that our out-of-sample fit assessment of the model was indeed obtained by drawing hold-out conjoint ratings from the model’s predictive distribution directly within our MCMC sampler. However, we concede Rubin (2004)’s point that the more common form of posterior predictive checks in the statistics literature are to assess the features of the model (out of sample prediction is one on them), which would be a nice dimension to be taken into consideration for future research.

ADDITIONAL DATA ANALYSES

Alba and Cook (2004) suggest that we provide additional support for our model by showing that the other decay parameters ($\lambda_{i1}-\lambda_{i5}$) do not correlate with the manipulated prior condition between price and maximum resolution. Table 1 shows the results from such correlation analysis. Indeed, we do not see this relationship with the other decay parameters (Table 1). This provides additional support of our model.

[Insert Table 1 Here.]

MODEL EXTENSIONS

Both Alba and Cooke (2004) and Rao (2004) suggest that the imputation process could have more sources of information other than historical attribute levels. For instance, Rao (2004) suggests a differing imputation process in which people impute missing values based upon their importance, or on the part-worths themselves. This is quite intriguing, and would be fairly easy to implement as we could simply model

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

where α_i represents the baseline propensity of an individual to impute a higher level, β_{ij} is the part-worth of individual i to attribute j , κ is a slope parameter, and ε_{ij} is a stochastic error term. This is also very much related to the comment by Alba and Cooke (2004) on evaluative consistency where profiles that perform well on attributes with higher relative importance tend to have the missing attributes imputed higher. This could be accomplished within our framework by modeling

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

where x_{ij} is the vector of observed levels in the current profile. Whether the subjects do utilize attributes with high importance more is an empirical question. Although such extensions enrich the behavioral versatility of BHH, they may lead to potential estimation issues because of higher order interaction terms involving the β 's. On the other hand, since the imputation parameters (ω 's) in BHH are not restricted, they should be able to, at least partially, capture such effects as to which attribute gets more weight in imputation.

Some additional issues raised about our model on a number of dimensions: 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, deserve some attention. 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, it does *not* imply that the prior correlation structure is ignored. In fact, if two or more attribute levels are highly (weakly) correlated then both their prior experience counts will be simultaneously high (low), and hence if both are missing, then they would both be more likely based on the a priori values to be imputed as the higher level. This can be seen by noting that the model we utilize

$$N_{ij}(0 | l_j) \sim \text{Poisson}(\exp(\zeta_i + \omega_j)), \quad (3)$$

allows for the possibility that the ω_j 's are correlated even though they are drawn i.i.d. from a common prior.

The issue raised about handling more than two attribute levels is an important one to bring our model closer to practical usage. Certainly, 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 a given level is imputed is given by

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

a direct extension of Equation (4) in BHH. Recall that BHH use a binary Hamming matching metric where $I(x_{ij}(t'), x_{ij}(t)) = 1$ for a match and 0 for a non-match. For nominal variables (such as gender, color, etc.), binary matching may be the best way to measure how “similar” two attribute levels are: they are either the same or different; there’s no measurement in-between. For continuous variables (such as 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 \$3). Specifically, we define:

$$d(x_{ij}(t), x_{ij}(t')) = \frac{|x_{ij}(t') - x_{ij}(t)|}{\text{MAX}_{\forall k_j, k_j^0, k_j \neq k_j^0} |l_{jk_j} - l_{jk_j^0}|} \quad (4)$$

where k_j represents level k of attribute j . A more common case in conjoint analysis is to treat continuous variables as ordinal with multiple levels (e.g., price is more often treated as discrete instead of continuous). In this case, we could choose one of the three similarity measurements: a Euclidean distance,

an equally-spaced distance between levels, or a binary Hamming level-matching and empirically validate which is most likely to hold. At the same time, Alba and Cooke (2004)'s concern that it is possible that it's easier for the subjects to impute when all attributes are binary is legitimate. We think it is an interesting and open empirical question. One can, however, easily argue the opposite since when more attribute levels are provided in a conjoint analysis, it will make it easier for the subject to impute a non-base-level for the missing PM attribute, an action that would lead to results different from a no-imputation model assumption where the base-level is assumed for a missing attribute.

With regards to symmetry, and the example brought up by Alba and Cooke (2004), while it is true in the particular case given in Figure 1B that $N_{i2}(3|1) = N_{i3}(3|1)$ had attributes 2 and 3 been reversed at time 3, we note that $N_{i2}(3|0) \neq N_{i3}(3|0)$ under the two cases and thus the probabilities would not be the same. In the case given in the manuscript, the probability of imputing a 1 is $(\lambda_2 + \lambda_3) / (\lambda_2 + \lambda_3 + \lambda_2^2)$ where in this new case it would be 1 as there would be no matches if a 0 was imputed. Nevertheless, even though the probabilities are different, it would be an interesting extension as suggested by Alba and Cooke (2004) to have an extended model where we have $\lambda_{i(j)j'}$, where (j) denotes the focal item being imputed and j' is the information source. In fact, this was the original formulation of the BHH model; however, was simplified for parsimony. Certainly if a model with focal weights were applied, an increased parameter space would have to lead to more careful thought about the design of the conjoint study (e.g. number of profiles).

Lastly, we consider the issue about imputing outside of the observed set of values. We agree completely that our discrete contingency table formulation has, at its core, a kernel where a discrete number of levels are possibly imputed. Also, it is true that under 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 the trial t and/or the past information observed, as long as that set remained discrete, is possible within the Bayesian paradigm utilized here and could be inferred along with the missing attribute values as well. This would be a significant change in the model, one in 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. The former is in general not observable in a conjoint task, but could be modeled via latent variables as done in BHH. The latter is what we used directly in BHH. Behavioral research (as suggested by Alba and Cooke 2004 and 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 have prompted us to focus on capturing only a limited set of relevant regularities and nesting commonly used extant models. Furthermore, research done in the field of psychology, and in educational testing (Bradlow and Thomas 1998, as pointed out by Rubin 2004), have also considered a conceptually similar problem of inferring levels (or abilities in the educational testing case) based on missing information/responses. We believe the following elements will be crucial in the future so that the bridging of behavioral and quantitative research, as we have attempted, can continue to occur:

1. **Mathematical formalism.** To incorporate behavioral regularities into standard models, it is necessary that these regularities are able to 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 point 2, below, is important.
2. **Interdisciplinary collaboration.** BHH are modelers and could have benefited greatly from working with a colleague with greater behavioral training. Such kind of collaboration would have made the proposed model more realistic perhaps without increasing the model complexity. Previous research of similar nature (e.g., Kahn and Raju 1991; Hardie, Johnson, and Fader 1993; Bodapati and Drolet 2000; Kivetz, Netzer, and Srinivasan 2004) have surely benefited from such “interdisciplinary collaboration”

and skills. Also pointed out by Rubin (2004), statisticians provide us with numerous 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 see collaborations, not only between behavioral researchers and modelers within the marketing domain itself, but across different fields (e.g., economics, operations, psychology, sociology, and statistics), as a way to go to undertake challenging and important research in marketing in the future.

Table 1 Correlation Between λ and Manipulated Price-Resolution Covariance

		λ_{i1}	λ_{i2}	λ_{i3}	λ_{i4}	λ_{i5}
1 missing	Correlation	-0.054	-0.057	0.049	-0.169	0.059
	(p-value)	(0.719)	(0.704)	(0.743)	(0.256)	(0.693)
2 missing	Correlation	0.003	0.061	0.050	-0.095	-0.019
	(p-value)	(0.987)	(0.697)	(0.752)	(0.546)	(0.903)

Figure 1 Ford Motor Company Vehicle Advisor: Application of BHH

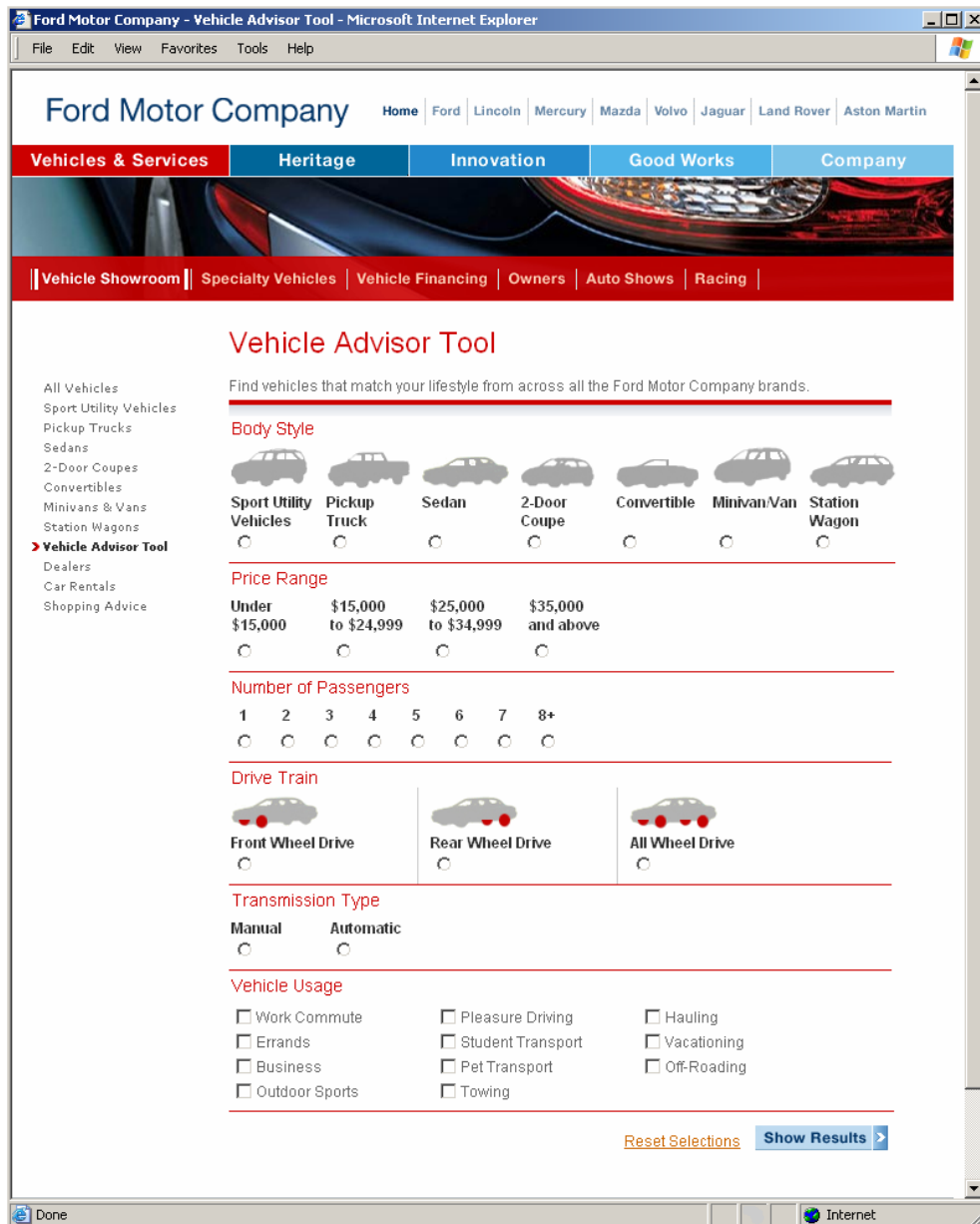




Figure 2 Side by side Comparison at Ford Vehicle Advisor

Overview		Specifications	Features	Safety
Add Vehicle	save vehicle 2005 Mazda Tribute i 2WD 5-spd MT  [Remove]	save vehicle 2004 Ford Explorer 4X2 XLS 4.0L  [Remove]		
	MSRP \$19,320	\$26,600		
Get Dealer Price Quotes	ZIP Code: <input type="text"/> GO	ZIP Code: <input type="text"/> GO		
Invoice Price	\$18,055	\$24,179		
Estimated Payment	\$403/month*	\$555/month*		
Quick Facts				
Rebates and Incentives	Cash back: \$0-\$1,250 APR for 48 months: 2.90%	Cash back: \$3,000 APR for 48 months: 0.00%		
Gas Mileage	24 mpg city/29 mpg hwy	16 mpg city/21 mpg hwy		
Engine(s)	2.3L I4, 16 valve, 153@5800 hp	4.0L V6, 12 valve, 210@5100 hp		
Transmission	5 speed Manual	5 speed Automatic		
Basic Warranty	50000 Mi./48 Months	36000 Mi./36 Months		
Powertrain Warranty	50000 Mi./48 Months	36000 Mi./36 Months		
User Ratings	No reviews yet	No reviews yet		
Overview		Specifications	Features	Safety

REFERENCES

- Alba, Joseph W. and Alan Cooke (2004), "When Absence Regrets Inference in Conjoint Analysis," *Journal of Marketing Research*, 41(4), XXX-XXX.
- 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, E.T. and N. Thomas (1998), "Item Response Theory Models Applied to Data Allowing Examinee Choice," *Journal of Educational and Behavioral Statistics*, 23(3), 236-243.
- Bradlow, Eric T., Ye Hu, and Teck-Hua Ho (2004), "A Learning-Based Model for Imputing Missing Attribute Levels in Partial Conjoint Profiles," *Journal of Marketing Research*, 41(4), XXX-XXX.
- Camerer, Colin F. and Teck-Hua Ho (1999), "Experience-Weighted Attraction Learning in Normal Form Games," *Econometrica*, 67(4), 837-874.
- Camerer, Colin F., Teck-Hua Ho, and Juin-Kuan Chong (2004), "A Cognitive Hierarchy Model of One-Shot Games," *Quarterly Journal of Economics* 119(3), forthcoming.
- 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), 378-394.
- Kahn, Barbara E. and Jagmohan S. Raju (1991), "Effects of Price Promotions on Variety-Seeking and Reinforcement Behavior," *Marketing Science*, 10(4), 316-337.
- Kevitz, Ran, Oded Netzer, and V. Srinivasan (2004), "Alternative Models for Capturing the Compromise Effect," *Journal of Marketing Research*, 41(August), forthcoming.

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-B485.

Rao, Vithala (2004), "Comments on Eric Bradlow, Ye Hu and Teck-Hua Ho, 'A Learning-Based Model'," *Journal of Marketing Research*, 41(4), XXX-XXX

Rubin, Donald B. (1976), "Inference and Missing Data.," *Biometrika*, 63(3), 581-592.

Rubin, Donald B. (1984), "Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician," *Annals of Statistics* 12, 1151-1172.

Rubin, Donald B. and H. Stern (1994), "Testing in Latent Class Models Using a Posterior Predictive Check Distribution," *Latent Variables Analysis: Applications for Developmental Research*. A. Von Eye and C.C. Clogg (eds.), Sage Publications, Ch. 17, 420-438.

Rubin, Donald B. (2004), "Comments on Eric Bradlow, Ye Hu and Teck-Hua Ho, 'A Learning-Based Model'," *Journal of Marketing Research*, 41(4), XXX-XXX.