Sawtooth Software

RESEARCH PAPER SERIES

Perspectives Based on 10 Years of HB in Marketing Research

Greg M. Allenby,
The Ohio State University
and
Peter E. Rossi,
University of Chicago

530 W. Fir St.
Sequim, WA 98382
(360) 681-2300
www.sawtoothsoftware.com

Perspectives Based on 10 Years of HB in Marketing Research

[1]Greg M. Allenby, Ohio State University
Peter E. Rossi, University of Chicago

Abstract

Bayes theorem, and the Bayesian perspective to analysis, has been around for hundreds of years. It has just recently experienced a tremendous increase in popularity. In this paper, we describe the Bayesian approach, the recent revolution in marketing research that has occurred, and the revolution that is about to occur.

1. Introduction

Hierarchical Bayes (HB) methods were first applied to marketing problems in the early 90s. Since, this time, over 50 HB papers have been published in marketing journals. The popularity of the HB approach stems from several unique advantages afforded this approach over conventional methods. The Bayesian approach offers more accurate solutions to a wider class of research problems than previously possible, facilitating the integration of data from multiple sources, dealing directly with the discreteness (i.e., lumpiness) of marketing data, and tracking the uncertainty present in marketing analysis, which is characterized by many "units" of analysis (e.g., respondents), but relatively sparse data per unit.

The purpose of this paper is to provide an introduction to, and perspective on, Bayesian methods in marketing research. This paper begins with a conceptual discussion of Bayesian analysis and the challenge of conducting Bayesian analysis on marketing data. Hierarchical Bayes (HB) models and Markov chain Monte Carlo (MCMC) estimators are introduced as a means of dealing with these challenges. The MCMC estimator replaces difficult analytic calculations with an iterative, simple, computational procedure. When coupled with HB models, the combination offers researchers the dual capability of analyzing larger problems with more accuracy. Some challenges of implementing and estimating HB models are then discussed, followed by a perspective on where marketing research analysis is headed with this new tool. Readers interested in a formal and more complete account of these methods are referred to Rossi and Allenby (2003).

2. Bayes Theorem and Marketing Data

Bayes theorem was originally published in 1764 as "An Essay toward Solving a Problem in the Doctrine of Chances" by the Royal Society of London, England. In his essay, Bayes

[1] The authors may be contacted at allenby.1@osu.edu; peter.rossi@gsb.uchicago.edu

proposed a formal rule for accounting for uncertainty. In its simplest form, if H denotes a hypothesis and D denotes data, the theorem states that

$$\Pr(H|D) = \Pr(D|H) \times \Pr(H) / \Pr(D) \tag{1}$$

Where Pr(.) denotes a probabilistic statement. That is, Pr(H) is a probabilistic statement about the hypothesis before seeing the data, and Pr(D|H) is the probability of the data given (i.e., conditional on) the hypothesis. Bayes rule can be derived from usual operations with probabilities:

$$\begin{aligned}\Pr(H,D) &= \Pr(H) \times \Pr(D|H) \\ &= \Pr(D) \times \Pr(H|D)\end{aligned} \tag{2}$$

which, after equating the two expressions on the right of the equal signs and dividing both sides by Pr(D), yields the expression in equation (1).

Bayes theorem is useful in problems involving what has historically been called "inverse probability." In these problems, an analyst is given the data and, from that information, attempts to infer the random process that generated them by using equation (1). For example, the data (D) could be the yes-no response to the question "do you camp?" and the hypothesis (H) might be a set of factors (e.g., product ownership or enjoying other outdoor activities) that are hypothesized to be camping-related. The probability of the data given the hypothesized model can take many forms, including regression models, hazard models, and discrete choice models. Bayes theorem is used to derive probability statements about the unobserved data generating process by multiplying the probability of the data given the model, Pr(D|H), by the prior probability of the hypothesized model, Pr(H), and dividing by Pr(D).

Bayes theorem offers a formula that accounts for uncertainty in moving from the hypothesized data-generating process, Pr(D|H), to inferences about the hypothesis given the data, Pr(H|D). To a Bayesian, there is little difference between a hypothesis (H) and any other unobserved (latent) aspect of a model, including model parameters (β). Bayes rule is applied to any and all unobservables, with the goal of making inferences based on the rules of probability. A critical difference between Bayesian and non-Bayesian analysis is that Bayesians condition on the observed data (D) while non-Bayesians condition on the hypothesis (H) and model parameters (β). Non-Bayesian analysis proceeds by conditioning on the hypothesized model – i.e., assuming that the hypothesis is known with certainty – and searching for the best fitting model that maximizes Pr(D|H). The hypothesis, in reality, is never known, and such an assumption destroys the ability of the analyst to account for the uncertainty present in an analysis.

Accounting for uncertainty is important in the analysis of marketing research data. Marketing data contains large amounts of uncertainty, typically comprising many heterogeneous "units" (e.g., households, respondents, panel members, activity occasions) with limited information on each unit. These units may differ in their preferences, sensitivities, beliefs and motivations. In a conjoint analysis, for example, it is rare to have more than 20 or so evaluations, or choices, of product descriptions per respondent. In the analysis of direct marketing data, it is rare to have more than a few dozen orders for a customer within a given product category.

Researchers often report that predictions in marketing research analyses are too aggressive and unrealistic. A reduction in price of 10 or 20%, for example, results in an increase in market share that is known to be too large. Overly optimistic predictions can easily result if estimates of price sensitivity are incorrectly assumed known with certainty. When uncertainty is accounted for, market share predictions become more realistic and tend to agree with current and past experience. The inability to accurately account for data uncertainty also creates problems when integrating data from multiple sources, when conducting an analysis in which the output from one procedure is used as input to another, and in conducting analysis of latent processes, such as the use of brand beliefs in forming consideration sets.

While Bayes theorem is a conceptually simple method of accounting for uncertainty, it has been difficult to implement in all but the simplest problems. It is typically the case that the data, D, are assumed to arise from hypothesized models where Pr(D|H) and Pr(H), when multiplied together, take a form that leads to difficulty in constructing inferences about model parameters and making predictions. This occurs, for example, in the analysis of choice data where there is assumed to exist a latent, utility maximizing process. Accordingly, until recently, researchers in marketing and other fields have tended not to use Bayesian methods, and have instead conducted analysis based entirely on Pr(D|H).

3. Hierarchical Bayes (HB) Models

Recent developments in statistical computing have made Bayesian analysis accessible to researchers in marketing and other fields. The innovation, known as Markov chain Monte Carlo (MCMC), has facilitated the estimation of complex models of behavior that can be infeasible to estimate with alternative methods. These models are written in a hierarchical form, and are often referred to as hierarchical Bayes models. Discrete choice models, for example, assume that revealed choices reflect an underlying process where consumers have preferences for alternatives and select the one that offers greatest utility. Utility is assumed related to specific attribute levels that are valued by the consumer, and consumers are assumed to be heterogeneous in their preference for the attributes. The model is written as a series of hierarchical algebraic statements, where model parameters in one level of the hierarchy are unpacked, or explained, in subsequent levels.

$$\Pr(y_{ih} = 1) = \Pr(V_{ih} + \varepsilon_{ih} > V_{jh} + \varepsilon_{jh} \quad \text{for all } j) \tag{3}$$

$$V_i = x_i'\beta_h \tag{4}$$

$$\beta_h \sim \text{Normal}(\bar{\beta}, \Sigma_\beta) \tag{5}$$

where "i" and "j" denote different choice alternatives, y_{ih} is the choice outcome for respondent h, V_{ih} is the utility of choice alternative i to respondent h, x_i denotes the attributes of the i^{th} alternative, β_h are the weights given to the attributes by respondent h, and equation (5) is a "random-effects" model that assumes that the respondent weights are normally distributed in the population.

The bottom of the hierarchy specified by equations 3) – 5) is the model for the observed choice data. Equation 3) specifies that alternative j is chosen if the latent or unobserved utility is the largest among all of the alternatives. Latent utility is not observed directly and is linked to characteristics of the choice alternative and a random error in equation 4. Each respondent's part-worths or attribute weights are linked by a common distribution in equation 5. Equation 5) allows for heterogeneity among the units of analysis by specifying a probabilistic model of how the units are related. The model of the data-generating process, $\Pr(D_h|\beta_h)$, is augmented with a second equation $\Pr(\beta_h| \bar{\beta}, \Sigma_\beta)$ where $\bar{\beta}$ and Σ_β are what are known as "hyper-parameters" of the model, i.e., parameters that describe variation in other parameters rather than variation in the data. At the top of the hierarchy are the common parameters. As we move down the hierarchy we get to more and more finely partitioned information. First are the part worths which vary from respondent to respondent. Finally, at the bottom of the hierarchy are the observed data which vary by respondent and by choice occasion.

In theory, Bayes rule can be applied to this model to obtain estimates of unit-level parameters given all the available data, $\Pr(\beta_k|D)$, by first obtaining the joint probability of all model parameters given the data:

$$\Pr(\{\beta_h\}, \bar{\beta}, \Sigma_\beta|D) = [\, \Pi_h \Pr(D_h|\beta_h) \times \Pr(\beta_h| \bar{\beta}, \Sigma_\beta)\,] \times \Pr(\bar{\beta}, \Sigma_\beta) / \Pr(D) \qquad (6)$$

and then integrating out the parameters not of interest:

$$\Pr(\beta_k \mid D) = \int \Pr(\{\beta_h\}, \bar{\beta}, \Sigma_\beta|D)\, d\beta_{-k}\, d\bar{\beta}\, d\Sigma_\beta \qquad (7)$$

where "-k" denotes "except k" and $D=\{D_h\}$ denotes all the data. Equations (6) and (7) provide an operational procedure for estimating a specific respondent's coefficients (β_k) given all the data in the study (D), instead of just her data (D_k). Bayes theorem therefore provides a method of "bridging" the analysis across respondents while providing an exact accounting of the all the uncertainty present.

Unfortunately, the integration specified in equation (7) is typically of high dimension and impossible to solve analytically. A conjoint analysis involving part-worths in the tens (e.g., 15) with respondents in the hundreds (e.g., 500) leads to an integration of dimension in the thousands. This partly explains why the conceptual appeal of Bayes theorem, and its ability to account for uncertainty, has had popularity problems – its implementation was difficult except in the simplest of problems. Moreover, in simple problems, one obtained essentially the same result as a conventional (classical) analysis unless the analyst was willing to make informative probabilistic statements about hypotheses and parameter values prior to seeing the data, Pr(H). Marketing researchers have historically felt that Bayes theorem was intellectually interesting but not worth the bother.

4. The MCMC Revolution

The Markov chain Monte Carlo (MCMC) revolution in statistical computing occurred in the 1980s with the publication of papers by Geman and Geman (1984), Tanner and Wong (1987) and Gelfand and Smith (1990), eventually reaching the field of marketing with papers by Allenby and Lenk (1994) and Rossi and McCulloch (1994). The essence of the approach involves replacing the analytical integration in equation (4) with a Monte Carlo simulator involving a Markov chain. The Markov chain is a mathematical device that locates the simulator in an optimal region of the parameter space so that the integration is carried out efficiently, yielding random draws of all the model parameters. It generates random draws of the joint distribution $\Pr(\{\beta_h\}, \bar{\beta}, \Sigma_\beta|D)$ and all marginal distributions (e.g., $\Pr(\beta_h|D)$) instead of attempting to derive the analytical formula of the distribution. Properties of the distribution are obtained by computing appropriate sample statistics of the random draws, such as the mean, variance, and probability (i.e., confidence) intervals.

A remarkable fact of these methods is that the Monte Carlo simulator can replace integrals of any dimension (e.g., 10,000), with the only limitation being that higher dimensional integrals take longer to evaluate than integrals of only a few dimensions. A critical part of analysis is setting up the Markov chain so that it can efficiently explore the parameter space. An effective method of doing this involves writing down a model in a hierarchy, similar to that done above in equations (3) – (5).

MCMC methods have also been developed to handle the discreteness (i.e., lumpiness) of marketing choice data, using the technique of data augmentation. If we think of the data as arising from a latent continuous variable, then is it a relatively simple matter to construct an MCMC algorithm to sample from the posterior. For example, we can think of ratings scale data as arising from a censored normal random variable that is observed to be in one of k-1 "bins" or intervals for a k element scale. The resulting computational flexibility, when coupled with the exact inference provided by Bayes theorem, has lead to widespread acceptance of Bayesian methods within the academic field of marketing and statistics.

Diffusion of the HB+MCMC innovation into the practitioner community was accelerated by the existence of key conferences and the individuals that attended them. The American Marketing Association's Advanced Research Techniques (ART) Forum, the Sawtooth Software Conference, and the Bayesian Applications and Methods in Marketing Conference (BAMMCONF) at Ohio State University all played important roles in training researchers and stimulate use of the methods. The conferences brought together leading academics and practitioners to discuss new developments in marketing research methods, and the individuals attending these conferences were quick to realize the practical significance of HB methods.

The predictive superiority of HB methods has been due to the freedom afforded by MCMC to specify more realistic models, and the ability to conduct disaggregate analysis. Consider, for example, the distribution of heterogeneity, $\Pr(\beta_h|\bar{\beta}, \Sigma_\beta)$ in a discrete choice conjoint model. While it has long been recognized that respondents differ in the value they attach to attributes and benefits, models of heterogeneity were previously limited to the use of demographic covariates to explain differences, or the use of finite mixture models. Neither model is realistic – demographic variables are too broad-scoped to be related to attributes in a specific product category, and the assumption that heterogeneity is well approximated by a small number of customer types is more a hope than a reality. Much of the predictive superiority of HB methods is due to avoiding the restrictive analytic assumptions that alternative methods impose.

The disaggregate analysis afforded by MCMC methods has revolutionized analysis in marketing. By being able to obtain individual-level estimates, analysis can avoid many of the procedures imposed by analysts to avoid computational complexities. Consider, for example, analysis associated with segmentation analysis, target selection and positioning. Prior to the ability to obtain individual–level parameter estimates, analysis typically proceeded in a series of steps, beginning with the formation of segments using some form of grouping tool (e.g., cluster analysis). Subsequent steps then involved describing the groups, including their level of satisfaction with existing offerings, and assessing management's ability to attract customers by reformulating and repositioning the offering.

The availability of individual-level estimates has streamlined this analysis with the construction of choice simulators that take the individual-level parameter estimates as input, and allow the analyst to explore alternative positioning scenarios to directly assess the expected increase in sales. It is no longer necessary to conduct piece-meal analysis that is patched together in a Rube Goldberg-like fashion. Hierarchical Bayes models, coupled with MCMC estimation, facilitates an integrated analysis that properly accounts for uncertainty using the laws of probability. While these methods have revolutionized the practice of marketing research over the last 10 years, they require some expertise to implement successfully.

5. Challenges in Implementing HB

In addition to its widespread use in conjoint analysis because of Sawtooth Software, Bayesian models are being used by companies such as DemandTec to estimate price sensitivity for over 20,000 individual sku's in retail outlets using weekly sales data. These estimates of price sensitivity are used to identify profit maximizing retail prices. A challenge in carrying out this analysis is to estimate consumer price sensitivity given the basic assumption that rising prices are associated with declining sales for any offering, and that an increase in competitor prices will lead to an increase in own sales. Obtaining estimates of price sensitivity with the right algebraic signs is sensitive to the level of precision, or uncertainty, of the price-sales relationship.

One of the major challenges of implementing HB models is to understand the effect of model assumptions at each level of the hierarchy. In a conventional analysis, parameter estimates from a unit are obtained from the unit's data, $\Pr(\beta_h|D_h)$. However, because of the scarcity of unit-level data in marketing, some form of data pooling is required to obtain stable parameter estimates. Rather than assume no heterogeneity ($\beta_h = \beta$ for all h) or that heterogeneity in response parameters follow a deterministic relationship to a common set of covariates ($\beta_h = z_h'\gamma$) such as demographics, HB models often assume that the unit-level parameters follow a random-effects model ($\Pr(\beta_h \mid \bar{\beta}, \Sigma_\beta)$). As noted above, this part of the model "bridges" analysis across respondents, allowing the estimation of unit-level estimates using all the data $\Pr(\beta_h|D)$, not just the unit's data $\Pr(\beta_h|D_h)$.

The influence of the random-effect specification can be large. The parameters for a particular unit of analysis (h) now appear in two places in the model: 1) in the description of the model for the unit, $\Pr(D_h|\beta_h)$ and 2) in the random-effects specification, $\Pr(\beta_h|\bar{\beta}, \Sigma_\beta)$, and estimates of unit h's parameters must therefore employ both equations. The random-effects specification adds much information to the analysis of β_h, shoring up the information deficit that exists at the unit level with information from the population. This difference between HB models and conventional analysis based solely on $\Pr(D_h|\beta_h)$ can be confusing to an analyst and lead to doubt in the decision to use these new methods.

The influence of the unit's data, D_h relative to the random-effects distribution on the estimate of β_h depends on the amount of noise, or error, in the unit's data $\Pr(D_h|\beta_h)$ relative to the extent of heterogeneity in $\Pr(\beta_h|\bar{\beta}, \Sigma_\beta)$. If the amount of noise is large or the extent of heterogeneity is small, then estimates of β_h will be similar across units (h=1,2,…). As the data become less noisy and/or as the distribution of heterogeneity becomes more dispersed, then estimates of β_h will more closely reflect the unit's data, D_h. The balance between these two forces is determined automatically by Bayes theorem. Give the model specification, no addition input from the analyst is required because Bayes theorem provides an exact accounting for uncertainty and the information contained in each source.

Finally, the MCMC estimator replaces difficult analytic calculations with simple calculations that are imbedded in an iterative process. The process involves generating draws from various distributions based on the model and data, and using these draws to explore the joint distribution of model parameters in equation (6). This can be a time consuming process for large models, and a drawback is that HB models that longer to estimate than simpler models, which attempt only to identify parameter values that fit the data best. However, as computational speed increases, this drawback becomes less important.

6. New Developments in Marketing Research

In addition to improvements in prediction, HB methods have been used to develop new marketing research methods and insights, including new models of consumer behavior, new models of heterogeneity, and new decision tools. Discrete choice models have been developed to include carry-over effects (Allenby and Lenk 1994), quantity (Arora, Allenby and Ginter 1999, Allenby, Shively, Yang and Garratt 2003), satiation (Kim, Allenby and Rossi 2002), screening rules (Gilbride and Allenby 2003) and simultaneous effects (Manchanda, Chintagunta and Rossi, 2003 and Yang, Chen and Allenby 2003). Models of satiation facilitate identifying product characteristics that are responsible for consumers tiring of an offering, and have implications for product and product line formation. Screening rules are to simplify consumer decision making, and point to the features that are needed for a brand to be considered. These features are of strategic importance to a firm because they define the relevant competition for an offering. Finally, simultaneous models deal with the fact that marketing mix variables are chosen strategically by managers with some partial knowledge of aspects of demand not observed by the market researcher. For example, the sensitivity of prospects to price changes is used by producers to design promotions and by the prospects themselves when making their purchase decisions. Prices are therefore set from within the system of study, and are not independently determined. Incorrectly assuming that variables are independent can lead to biased estimates of the effectiveness of marketing programs.

Experience with alternative forms of the distribution of heterogeneity reveals that assuming a multivariate normal distribution leads to large improvements in parameter estimates and predictive fit (see, for example, Allenby, Arora and Ginter 1998). More specifically, assuming a normal distribution typically leads to large improvements relative to assuming that the distribution of heterogeneity follows a finite mixture model. Moreover, additional benefit is gained from using truncated distributions that constrain parameter estimates to sensible regions of support. For example, negative price coefficients are needed to solve for profit maximizing prices that are realistic.

Progress has been made in understanding the nature of heterogeneity. Consumer preferences can be interdependent and related within social and informational networks (Yang and Allenby, 2003). Moreover, heterogeneity exists at a more micro-level than the individual respondent. People act and use offerings in individual instances of behavior. Across instances, the objective environment may change with implications for consumer motivations and brand preferences (Yang, Allenby and Fennell 2002). Motivating conditions, expressed as the concerns and interests that lead to action, have been found to be predictive of relative brand preference, and are a promising basis variable for market segmentation (Allenby, et.al. 2002).

The new decision tools offered by HB methods exploit availability of the random draws from the MCMC chain. As mentioned above, these draws are used to simulate scenarios related to management's actions and to explore non-standard aspects of an analysis. Allenby and Ginter (1995) discuss the importance of exploring extremes of the distribution of heterogeneity to identify likely brand switchers. Other uses include targeted coupon delivery (Rossi, McCulloch and Allenby 1996) and constructing market share simulators discussed above.

7. A Perspective on Where We're Going

Freedom from computational constraints allows researchers and practitioners to work more realistically on marketing problems. Human behavior is complex, and, unfortunately, many of the models in use have not been. Consider, for example, the dummy-variable regression model used in nearly all realms of marketing research. This model has been used extensively in advising management what to offer consumers, at what price and through what channels. It is flexible and predicts well. But does it contain the right variables, and does it provide a good representation of the process that generated the data? Let's consider the case of survey response data.

Survey respondents are often confronted with descriptions of product offerings that they encode with regard to their meaning. In a conjoint analysis, respondents likely assess the product description for correspondence with the problem that it potentially solves, construct preferences, and provide responses. The part-worth estimates that conjoint analysis makes available reveal the important attribute-levels that provide benefit, but they cannot reveal the conditions that give rise to this demand in the first place. Such information is useful in guiding product formulation and gaining the attention of consumers in broadcast media. For example, simply knowing that high horsepower is a desirable property of automobiles does not reveal that consumers may be concerned about acceleration into high-speed traffic on the highway, stop and go driving in the city, or the ability to haul heavy loads in hilly terrain. These conditions exist up-stream from (i.e., prior to) benefits that are available from product attributes. The study of such upstream drivers of brand preference will likely see increased attention as researchers expand the size of their models with HB methods.

The dummy variable regression model used in the analysis of marketing research data is too flexible and lacks the structure present in human behavior, both when representing real world conditions and when describing actual marketplace decisions. For example, the level-effect phenomena described by Wittink et al. (1992) can be interpreted as evidence of model mis-specification in applying a linear model to represent a process of encoding, interpreting and responding to stimuli. More generally, we understand a small part of how people fill out questionnaires, form and use brand beliefs, employ screening rules when making actual choices, and why individuals display high commitment to some brands but not others. None of these processes is well represented by a dummy variable regression model, and all are fruitful areas of future research.

Hierarchical Bayes methods provide the freedom to study what should be studied in marketing, including the drivers of consumer behavior. It facilitates the study of problems characterized by a large number of variables related to each other in a non-linear manner, allowing us accurately to account for model uncertainty, and to employ an "inverse probability" approach to infer the process that generated the data. HB will be the methodological cornerstone for further development of the science of marketing, helping us to move beyond simple connections between a small set of variables.

References

Allenby, Greg M. and Peter J. Lenk (1994) "Modeling Household Purchase Behavior with Logistic Normal Regression," *Journal of the American Statistical Association*, 89, 1218-1231.

Allenby, Greg M. and James L. Ginter (1995) "Using Extremes to Design Products and Segment Markets," *Journal of Marketing Research*, 32, 392-403.

Allenby, Greg M., Neeraj Arora and James L. Ginter (1998) "On The Heterogeneity of Demand," *Journal of Marketing Research*, 35, 384-389.

Allenby, Greg, Geraldine Fennell, Albert Bemmaor, Vijay Bhargava, Francois Christen, Jackie Dawley, Peter Dickson, Yancy Edwards, Mark Garratt, Jim Ginter, Alan Sawyer, Rick Staelin, and Sha Yang (2002) "Market Segmentation Research: Beyond Within and Across Group Differences," *Marketing Letters*, 13, 3, 233-244.

Allenby, Greg M., Thomas S. Shively, Sha Yang and Mark J. Garratt (2003) "A Choice Model for Packaged Goods: Dealing with Discrete Quantities and Quantity Discounts," *Marketing Science*, forthcoming.

Arora, N. and Greg M. Allenby and James L. Ginter (1998), "A Hierarchical Bayes Model of Primary and Secondary Demand," *Marketing Science*, 17, 29-44.

Bayes, T. (1763) "An Essay Towards Solving a Problem in the Doctrine of Chances," *Philo. Trans. R. Soc London*, 53, 370-418. Reprinted in *Biometrika,* 1958, 45, 293-315.

Gelfand A.E. and A.F.M. Smith (1990) "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85, 398-409.

Geman, S. and D. Geman (1984) "Stochastic Relaxation, Gibbs Distribution and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 6, 721-741.

Gilbride, Timothy J. and Greg M. Allenby, "A Choice Model with Conjunctive, Disjunctive, and Compensatory Screening Rules," working paper, Ohio State University.

Kim, Jaehwan, Greg M. Allenby, and Peter E. Rossi (2002) "Modeling Consumer Demand for Variety," *Marketing Science*, 21, 3, 229-250.

Manchanda, P., P. K. Chintagunta and P. E. Rossi (2003), "Response Modeling with Non-random Marketing Mix Variables," working paper, Graduate School of Business, University of Chicago.

McCulloch, R. and P.E. Rossi (1984) "An Exact Likelihood Approach to Analysis of the MNP Model," *Journal of Econometrics*, 64, 207-240.

Rossi, Peter E., Robert E. McCulloch and Greg M. Allenby (1996) "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15, 321-340.

Rossi, Peter E. and Greg M. Allenby (2003) "Bayesian Methods and Marketing," working paper, Ohio State University.

Wittink, Dick, Joel Huber, Peter Zandan and Rich Johnson (1992) "The Number of Levels Effects in Conjoint: Where Does it Come From and Can It Be Eliminated?" *Sawtooth Software Research Paper Series*.

Yang, Sha, Greg M. Allenby and Geraldine Fennell (2002) "Modeling Variation in Brand Preference: The Roles of Objective Environment and Motivating Conditions," *Marketing Science*, 21, 1, 14-31.

Yang, Sha and Greg M. Allenby (2003) "Modeling Interdependent Consumer Preferences," *Journal of Marketing Research*, forthcoming.

Yang, Sha, Yuxin Chen and Greg M. Allenby (2003) "Bayesian Analysis of Simultaneous Demand and Supply," working paper, Ohio State University.

Bibliography of Bayesian Studies in Marketing

Ainslie, A. and Peter Rossi (1998), “Similarities in Choice Behavior across Product Categories,” *Marketing Science*, 17, 91-106.

Allenby, Greg, Neeraj Arora, Chris Diener, Jaehwan Kim, Mike Lotti and Paul Markowitz (2002) "Distinguishing Likelihoods, Loss Functions and Heterogeneity in the Evaluation of Marketing Models," *Canadian Journal of Marketing Research*, 20.1, 44-59.

Allenby, Greg M., Robert P. Leone, and Lichung Jen (1999), “A Dynamic Model of Purchase Timing with Application to Direct Marketing,” *Journal of the American Statistical Association*, 94, 365-374.

Allenby, Greg M., Neeraj Arora, and James L. Ginter (1998), “On the Heterogeneity of Demand,” *Journal of Marketing Research*, 35, 384-389.

Allenby, Greg M., Lichung Jen and Robert P. Leone (1996), “Economic Trends and Being Trendy: The Influence of Consumer Confidence on Retail Fashion Sales,” *Journal of Business & Economic Statistics,* 14, 103-111.

Allenby, Greg M. and Peter J. Lenk (1995), “Reassessing Brand Loyalty, Price Sensitivity, and Merchandising Effects on Consumer Brand Choice,” *Journal of Business & Economic Statistics*, 13, 281-289.

Allenby, Greg M. and James L. Ginter (1995), “Using Extremes to Design Products and Segment Markets,” *Journal of Marketing Research*, 32, 392-403.

Allenby, Greg M., Neeraj Arora, and James L. Ginter (1995), “Incorporating Prior Knowledge into the Analysis of Conjoint Studies,” *Journal of Marketing Research*, 32, 152-162.

Allenby, Greg M. and Peter J. Lenk (1994), “Modeling Household Purchase Behavior with Logistic Normal Regression,” *Journal of American Statistical Association*, 89, 1218-1231.

Allenby, Greg M. (1990) "Hypothesis Testing with Scanner Data: The Advantage of Bayesian Methods," *Journal of Marketing Research*, 27, 379-389.

Allenby, Greg M. (1990), “Cross-Validation, the Bayes Theorem, and Small-Sample Bias,” *Journal of Business & Economic Statistics,* 8, 171-178.

Andrews, Rick , Asim Ansari, and Imran Currim (2002) “Hierarchical Bayes versus finite mixture conjoint analysis models: A comparison of fit, prediction, and partworth recovery,” *Journal of Marketing Research*, 87-98

Ansari, A., Skander Essegaier, and Rajeev Kohli (2000), “Internet Recommendation Systems,” *Journal of Marketing Research,* 37, 363-375.

Ansari, A., Kamel Jedidi and Sharan Jagpal (2000), “A Hierarchical Bayesian Methodology for Treating Heterogeneity in Structural Equation Models,” *Marketing Science,* 19, 328-347.

Arora, N. and Greg M. Allenby (1999) “Measuring the Influence of Individual Preference Structures in Group Decision Making,” *Journal of Marketing Research,* 36, 476-487.

Arora, N. and Greg M. Allenby and James L. Ginter (1998), “A Hierarchical Bayes Model of Primary and Secondary Demand,” *Marketing Science*, 17, 29-44.

Blattberg, Robert C. and Edward I. George (1991), “Shrinkage Estimation of Price and Promotional Elasticities: Seemingly Unrelated Equations,” *Journal of the American Statistical Association,* 86, 304-315.

Boatwright, Peter, Robert McCulloch and Peter E. Rossi (1999),"Account-Level Modeling for Trade Promotion: An Application of a Constrained Parameter Hierarchical Model", *Journal of the American Statistical Association*, 94, 1063-1073.

Bradlow, Eric T. and David Schmittlein (1999), “The Little Engines That Could: Modeling the Performance of World Wide Web Search Engines,” *Marketing Science* 19, 43-62.

Bradlow, Eric T. and Peter S. Fader (2001), “A Bayesian Lifetime Model for the “Hot 100” *Billboard* Songs,” *Journal of the American Statistical Association,* 96, 368-381.

Bradlow, Eric T. and Vithala R. Rao (2000), “A Hierarchical Bayes Model for Assortment Choice,” *Journal of Marketing Research*, 37, 259-268.

Chaing, Jeongwen, Siddartha Chib and Chakrvarthi Narasimhan (1999), “Markov Chain Monte Carol and Models of Consideration Set and Parameter Heterogeneity,” *Journal of Econometrics* 89, 223-248.

Chang, K., S. Siddarth and Charles B. Weinberg (1999), “The Impact of Heterogeneity in Purchase Timing and Price Responsiveness on Estimates of Sticker Shock Effects,” *Marketing Science*, 18, 178-192.

DeSarbo, Wayne, Youngchan Kim and Duncan Fong (1999), “A Bayesian Multidimensional Scaling Procedure for the Spatial Analysis of Revealed Choice Data,” *Journal of Econometrics* 89, 79-108.

Edwards, Yancy and Greg M. Allenby (2002) "Multivariate Analysis of Multiple Response Data," *Journal of Marketing Research*, forthcoming.

Huber, J. and Kenneth Train (2001), “On the Similarity of Classical and Bayesian Estimates of Individual Mean Partworths,” *Marketing Letters*, 12, 259-269.

Jen, Lichung, Chien-Heng Chou and Greg M. Allenby (2003) "A Bayesian Approach to Modeling Purchase Frequency," *Marketing Letters*, forthcoming.

Kalyanam, K. and Thomas S. Shiveley (1998), "Estimating Irregular Pricing Effects: A Stochastic Spline Regression Approach," *Journal of Marketing Research,* 35, 16-29.

Kalyanam, K. (1996), "Pricing Decision under Demand Uncertainty: A Bayesian Mixture Model Approach," *Marketing Science,* 15, 207-221.

Kamakura, Wagner A. and Michel Wedel (1997), "Statistical Data Fusion for Cross-Tabulation," *Journal of Marketing Research,* 34, 485-498.

Kim, Jaehwan, Greg M. Allenby and Peter E. Rossi (2002), "Modeling Consumer Demand for Variety," *Marketing Science*, forthcoming.

Leichty, John, Venkatram Ramaswamy, and Steven H. Cohen (2001), "Choice Menus for Mass Customization," *Journal of Marketing Research* 38, 183-196.

Lenk, Peter and Ambar Rao (1990), "New Models from Old: Forecasting Product Adoption by Hierarchical Bayes Procedures," *Marketing Science* 9, 42-53.

Lenk, Peter J., Wayne S. DeSarbo, Paul E. Green and Martin R. Young, (1996), "Hierarchical Bayes Conjoint Analysis: Recovery of Partworth Heterogeneity from Reduced Experimental Designs," *Marketing Science,* 15, 173-191.

Manchanda, P., Asim Ansari and Sunil Gupta (1999), "The "Shopping Basket": A Model for Multicategory Purchase Incidence Decisions," *Marketing Science,* 18, 95-114.

Marshall, P. and Eric T. Bradlow (2002), "A Unified Approach to Conjoint Analysis Models," *Journal of the American Statistical Association*, forthcoming.

McCulloch, Robert E. and Peter E. Rossi (1994) "An Exact Likelihood Analysis of the Multinomial Probit Model," *Journal of Econometrics,* 64, 217-228

Montgomery, Alan L. (1997), "Creating Micro-Marketing Pricing Strategies Using Supermarket Scanner Data," *Marketing Science* 16, 315-337.

Montgomery, Alan L. and Eric T. Bradlow (1999), "Why Analyst Overconfidence About the Functional Form of Demand Models Can Lead to Overpricing," *Marketing Science,* 18, 569-583.

Montgomery, Alan L. and Peter E. Rossi (1999), "Estimating Price Elasticities with Theory-Based Priors," *Journal of Marketing Research,* 36, 413-423.

Neelamegham, R. and Pradeep Chintagunta (1999), "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets," *Marketing Science*, 18, 115-136.

Otter, Thomas, Sylvia Frühwirth-Schnatter and Regina Tüchler (2002) "Unobserved Preference Changes in Conjoint Analysis," Vienna University of Economics and Business Administration, working paper.

Putler, Daniel S. Kirthi Kalyanam and James S. Hodges (1996), "A Bayesian Approach for Estimating Target Market Potential with Limited Geodemographic Information," *Journal of Marketing Research*, 33, 134-149.

Rossi, Peter E., Zvi Gilula, Greg M. Allenby (2001), "Overcoming Scale Usage Heterogeneity: A Bayesian Hierarchical Approach," *Journal of the American Statistical Association*, 96, 20-31.

Rossi, Peter E., Robert E. McColloch and Greg M. Allenby (1996), "The Value of Purchase History Data in Target Marketing," *Marketing Science*, 15, 321-340.

Rossi, Peter E. and Greg M. Allenby (1993), "A Bayesian Approach to Estimating Household Parameters," *Journal of the Marketing Research*, 30, 171-182.

Sandor, Zsolt and Michel Wedel (2001), "Designing Conjoint Choice Experiments Using Managers' Prior Beliefs," *Journal of Marketing Research* 28, 430-444.

Seetharaman, P. B., Andrew Ainslie, and Pradeep Chintagunta (1999), "Investigating Household State Dependence Effects Across Categories," *Journal of Marketing Research* 36, 488-500.

Shively, Thomas A., Greg M. Allenby and Robert Kohn (2000), "A Nonparametric Approach to Identifying Latent Relationships in Hierarchical Models," *Marketing Science*, 19, 149-162.

Steenburgh, Thomas J., Andrew Ainslie, and Peder H. Engebretson (2002), "Massively Categorical Variables: Revealing the Information in Zipcodes," *Marketing Science*, forthcoming.

Talukdar, Debobrata, K. Sudhir, and Andrew Ainslie (2002), "Investing New Production Diffusion Across Products and Countries," *Marketing Science* 21, 97-116.

Ter Hofstede, Frenkel, Michel Wedel and Jan-Benedict E.M. Steenkamp (2002), "Identifying Spatial Segments in International Markets," Marketing Science, 21, 160-177.

Ter Hofstede, Frenkel, Youingchan Kim and Michel Wedel (2002), "Bayesian Prediction in Hybrid Conjoint Analysis," Journal of Marketing Research, 34, 253-261.

Wedel, M. and Rik Pieters (2000), "Eye Fixations on Advertisements and Memory for Brands: A Model and Findings," *Marketing Science*, 19, 297-312.

Yang, Sha and Greg M. Allenby (2000), "A Model for Observation, Structural, and Household Heterogeneity in Panel Data," *Marketing Letters*, 11, 137-149.

Yang, Sha, Greg M. Allenby, and Geraldine Fennell (2002a), "Modeling Variation in Brand Preference: The Roles of Objective Environment and Motivating Conditions," *Marketing Science*, 21, 14-31.

Yang, Sha and Greg M. Allenby (2002b), "Modeling Interdependent Consumer Preferences," *Journal of Marketing Research*, forthcoming.