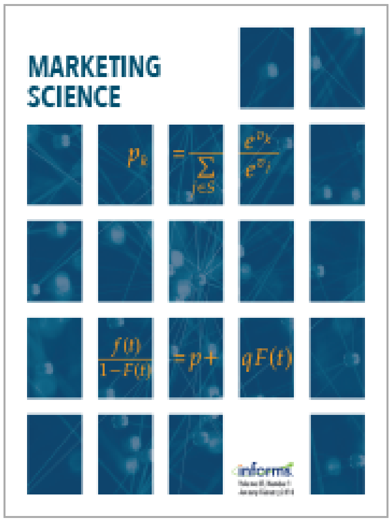

Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Bayesian Analysis of Hierarchical Effects

Sandeep R. Chandukala, Jeffrey P. Dotson, Jeff D. Brazell, Greg M. Allenby,

MARKETING SCIENCE
CELEBRATING 30 YEARS
Vol. 30, No. 1, January–February 2011, pp. 123–133
ISSN 0732-2399 | EISSN 1526-548X | 11 | 3001 | 0123

DOI 10.1287/mksc.1100.0602

Bayesian Analysis of Hierarchical Effects

Sandeep R. Chandukala
Kelley School of Business, Indiana University, Bloomington, Indiana 47405, sarchand@indiana.edu

Jeffrey P. Dotson
Owen Graduate School of Management, Vanderbilt University, Nashville, Tennessee 37203, jeff.dotson@owen.vanderbilt.edu

Jeff D. Brazell
The Modellers, LLC, Salt Lake City, Utah 84047, jeff.brazell@themodellers.com

Greg M. Allenby
Fisher College of Business, Ohio State University, Columbus, Ohio 43210, allenby.1@osu.edu

The idea of hierarchical, sequential, or intermediate effects has long been posited in textbooks and academic literature. Hierarchical effects occur when relationships among variables are mediated through other variables. Challenges in studying hierarchical effects in marketing include the large number of items present in most commercial studies and the presence of heterogeneous relationships among the variables. Existing approaches have dealt with the large number of variables by employing a factor structure representation of the data and have used standard mixture distributions for representing different response segments. In this paper, we propose a Bayesian model for the analysis of hierarchical data using the actual response items and incorporating heterogeneity that better reflects consumer stages in a decision process. Cross-sectional data from a national brand-tracking study are used to illustrate our model, where we find empirical support for a hierarchical relationship among media recall, brand beliefs, and intended actions. We find these effects to be insignificant when measured with standard models and aggregate analyses. The proposed model is useful for understanding the influence of variables that lead to intermediate as opposed to direct effects on brand choice.

Key words: hierarchical Bayes; mediation analysis; structural heterogeneity; variable selection
History: Received: April 19, 2010; accepted: August 6, 2010; Eric Bradlow served as the editor-in-chief and Robert Meyer served as associate editor for this article. Published online in *Articles in Advance* November 4, 2010.

1. Introduction

Hierarchical effects have been a fixture in marketing theory for decades, particularly in the context of explaining the effects of advertising. Hierarchical effects exist when relationships among variables are mediated by other variables. These occur when a decision process is seen as having distinct stages, such as beginning with the recognition of a need and terminating with choice and postpurchase evaluation (e.g., Vakratsas and Ambler 1996, Assael 1993). The AIDA framework discussed by E. St. Elmo Lewis in 1898 (Strong 1925) is one of the earliest and best-known examples, where advertising affects choice through a specific sequential process: from attention to interest to desire to action. Lavidge and Steiner (1961) extend this process to include knowledge, liking, preference, conviction, and ultimately purchase. Recent work has attempted to generalize these early advertising effects models by couching them in terms of cognition, affect, experience, and action (Vakratsas and Ambler 1999, Holbrook and Batra 1987). Although they differ in form, many models of consumer decision processes adhere to the general notion that consumers first learn, and then they know, feel, and do. That is, the effects of knowing on doing are assumed to be mediated, at least in part, by variables that reflect feelings and affect. A challenge in the analysis of hierarchical effects is the large number of variables present that reflect diverse aspects of behavior.

This paper proposes a Bayesian approach to studying hierarchical, or mediating, effects. Hierarchical effects are modeled by expressing the joint distribution of variables under study as a set of marginal and conditional distributions. Thus, for three sets of variables (A, B, C), we can write the joint distribution $[A, B, C]$ in terms of various factorizations. The model $[A, B, C] = [A \mid B][B \mid C][C]$ indicates that B mediates the influence of C on A, whereas the model $[A, B, C] = [A \mid B, C][B][C]$ indicates that both B and C directly influence A. The model $[A \mid B, C][B \mid C][C]$ indicates a direct effect of B and C on A and an effect of C on B. In these expressions, the vertical bar "|" denotes the conditioning argument, and the bracket notation implies a distribution.

Our approach to analyzing hierarchical effects in extended models of behavior is to employ three known aspects of modern Bayesian analysis. First, as stated above, we express hierarchical effects in terms of probability statements about the joint distribution of the data. In the model $[A, B, C] = [A \mid B][B \mid C][C]$, the variable A is conditionally independent of C given B. All information from C that affects A must flow through B. Conditional independence is a basic tenet of all hierarchical models. Our approach, however, differs from the typical hierarchical Bayes models in marketing, where interest focuses on explaining variation in effect sizes using random-effects models. Instead, we propose tools to investigate hierarchical effects among variables, not among coefficients.

Second, our model allows for heterogeneous segments that differ in terms of response coefficients and error variance. We employ the Bayesian method of data augmentation (Tanner and Wong 1987) to identify respondent segment memberships, greatly simplifying high-dimensional calculations associated with these models. Third, our model uses Bayesian variable selection (George and McCulloch 1993, Brown et al. 1998) to help reduce the high-dimensional mapping between sets of variables. We show below that both these modeling aspects are necessary for our data.

Our model generalizes a statistical procedure commonly known as "mediation analysis" (Baron and Kenny 1986), where a series of estimated regression relationships is used to assess the presence of a hierarchical relationship. Traditional mediation analysis involves using standard regression procedure and comparing effect sizes across different combinations of variables. More recently, Iacobucci et al. (2007) propose using structural equation models in a series of univariate regression models to handle multiple mediating variables. Our approach, which involves testing for conditional independence rather than the comparing effect sizes (Zhao et al. 2010), is more flexible and can be used to investigate multiple hierarchical effects and multiple levels of a hierarchy. Our model is related to two alternative approaches to studying hierarchical effects—structural equations models (SEMs; see Jedidi et al. 1997) and Gaussian graphical models (Whittaker 1990). SEMs employ the property of conditional independence on a factor representation of the data. In contrast, our model is based on the actual questionnaire items. We show below that much information is lost by trying to represent hierarchical mediating relationships with factors that are abstractions of the actual questionnaire items. We compare our best-fitting model that incorporates heterogeneity and variable selection with a finite mixture SEM and find that there is an inherent advantage of using our proposed conditional structure.

Gaussian graphical models represent questionnaire items with a multivariate normal distribution, using a parameterization based on the precision matrix instead of the covariance matrix. In graph theory models, conditional relationships among variables are modeled by zeroing out off-diagonal elements of the precision matrix, which is the same as setting coefficients to zero in a regression model; i.e., $\beta_{i,j} = -a_{i,j}/a_{i,i}$, where $A = \{a_{i,j}\}$ is the precision matrix. Our various factorizations correspond to zeroing out large blocks of off-diagonal elements of the precision matrix, and our use of Bayesian variable selection can be thought of as fine-tuning the conditional relationships by setting additional elements of A to zero. Also, unrestricted model search, using Gaussian graphical models, with a large number of variables (greater than 10 or 15) is also problematic, challenging, and computationally burdensome (Jones et al. 2005).

Our model extends existing graph theory models by allowing for heterogeneous response segments and incorporating variable selection. It also offers a more natural method of detecting the presence of conditional relationships using the prior associated with Bayesian variable selection methods where effect sizes smaller than a specified value are shrunk toward zero. In our model, priors are specified directly on the regression coefficients $\{\beta_{i,j}\}$ that correspond to the effect sizes. In Gaussian graphical models, priors are specified on elements of the precision matrix $\{a_{i,j}\}$ using a Wishart distribution (Letac and Massam 2007). Although there is a well-defined relationship between the regression coefficients and parameters of the precision matrix as shown above, the induced prior on effect sizes in a graph theory model does not lend itself to variable selection. We find that our use of Bayesian variable selection is very helpful in interpreting the hierarchical effects present in our data. We compare our best-fitting model with the current existing graphical model in the literature and demonstrate a better fit.

The remainder of this paper is organized as follows. Section 2 presents a Bayesian model for estimating heterogeneous hierarchical effects with a prior specification that facilitates variable selection. A simulation study is presented in the electronic companion (available as part of the online version that can be found at http://mktsci.pubs.informs.org) to demonstrate key features of the proposed method. Section 3 describes the empirical study used to investigate different hierarchical effects, and results are presented in §4. Section 5 contains a discussion of the results, and §6 offers concluding remarks and discusses limitations of our analysis.

2. Model

Hierarchical effects models of consumer behavior and advertising are particularly important in the context of high-involvement goods and services in which substantial prolonged consideration occurs prior to purchase. An initial stage of information search is typically posited as a set of considered brands is formed, additional information is sought as preferences are developed, brands are ranked, and choices are made. If consumers do engage in a staged decision process, it is likely that their sensitivities to advertising and other experiential activities differ according to their relative stage in that process.

Our models for investigation of hierarchical effects relate an intent-to-purchase measure (y) to a series of intended consumer actions (X), brand beliefs (B), and media exposure variables (Z). We use capital letters (e.g., X) to denote a set of responses by collections of individuals and lowercase letters (e.g., x_i) to denote the response from a specific individual indexed by the subscript i. The intent-to-purchase measure, y, is a single item and is represented in lowercase (y and y_i). We consider a variety of relationships among these variables and report on the best-fitting results for our data below. We emphasize, though, that our approach can accommodate any hierarchical model structure and is not confined to our specific model development.

We describe our modeling approach in detail using one of the models: $[y \mid X][X \mid B][B \mid Z][Z]$. We use this model merely for illustration purposes and test other alternative forms, e.g., $[y \mid X, B, Z][X \mid B][B \mid Z][Z]$, where we incorporate lower-level direct effects into the heterogeneity specification. We assume that the different density factors ($[y \mid X]$, $[X \mid B]$) are normally distributed. In this model ($[y \mid X][X \mid B][B \mid Z][Z]$), we believe the effects of media exposure (Z) on purchase intention (y) are mediated first through beliefs (B) and then from beliefs to a series of intended actions (X). We introduce heterogeneous response segments in the first factor in the model by assuming a finite mixture of likelihoods:

$$y_i \mid x_i, \{\phi_k, \beta_k, \sigma_k^2\} \sim \sum_{k=1}^{K} \phi_k N(x_i'\beta_k, \sigma_k^2), \tag{1}$$

where k indexes the response segments with relative size ϕ_k and $\sum_k \phi_k = 1$. Equation (1) implies that an arbitrarily selected respondent belongs to one of K response segments, each of which is related to the consumer actions through segment-specific regression coefficients β_k and error variance σ_k^2. Equation (1) is needed when there exists some respondents who are "in the market" and are influenced by the intended action variables (X), whereas other respondents are new to the market and are not influenced by these variables. We show below that an advantage of introducing heterogeneity at this point of the model is that it gives greater weight to the relationship between y and X in segment formation. Our model, however, is general and can accommodate broader weights in segment formation, including the relationship among all the variables.

The brand belief variables (B) are related to intended consumer action variables (X) through independent auxiliary multivariate regression models for each response segment:

$$x_i \mid b_i, \Gamma_k, \Sigma_k, (s_i = k) \sim N(\Gamma_k b_i, \Sigma_k). \tag{2}$$

Likewise, media exposure variables (Z) are related to brand belief variables (B) through independent auxiliary multivariate regression models for each segment:

$$b_i \mid z_i, \Delta_k, \Omega_k, (s_i = k) \sim N(\Delta_k z_i, \Omega_k), \tag{3}$$

where the regression coefficients (Γ_k and Δ_k) and error covariance matrix (Σ_k and Ω_k) are segment-specific. In this model, the response segments (k) are determined entirely by the relationship between consumer actions (X) and purchase intention (y) in Equation (1). Equations (2) and (3) are only used to characterize these segments. This assumption, however, is not a requirement of our general approach to study hierarchical effects and is made in anticipation of our empirical results that are tested by comparing model fits.

An advantage of our assumed segment structure described by Equations (1)–(3) is that it does not necessarily require the presence of media effects (Δ) for a staged decision process to exist. Some media may be ineffective at driving some of the brand beliefs, and some of the brand beliefs may be ineffective in driving the consumer actions. We believe that the absence of this relationship (i.e., elements of Γ and Δ equal to zero) should not be used as a basis for identifying the latent segments. We recognize that this is a strong assumption and therefore propose alternative models below in which the assumption is relaxed.

Equations (1)–(3) specify an indirect effect of media advertising on purchase intention that is mediated first by brand beliefs (B) and then by intended actions (X). We can estimate media advertising's total effect on purchase intention using the chain rule of differentiation:

$$\begin{aligned}\frac{\partial y}{\partial z'} &= \sum_{k=1}^{K} \varphi_k \left(\frac{\partial y}{\partial x_k'} \frac{\partial x}{\partial b_k'} \frac{\partial b}{\partial z_k'} \right) \\ &= \sum_{k=1}^{K} \varphi_k (\beta_k' \Gamma_k \Delta_k).\end{aligned} \tag{4}$$

An immediate result from Equation (4) is that specific elements of the coefficient vectors (β_k) and matrices (Γ_k and Δ_k) need to be nonzero for their product to be nonzero. Thus, attempting to explore a direct link between purchase intention (y) and media exposure (Z) may result in a finding of no effect, even when media do affect some aspects of intermediate beliefs and behavior. Our model can reflect intermediate effects through the coefficient matrices Γ_k and Δ_k.

Our approach to studying hierarchical effects involves comparison of the joint distribution of the data, $[y, X, B, Z]$, under alternative factorizations. Comparison among factorizations can be done in terms of the marginal density of data, a Bayesian measure of model fit. Computing the joint marginal density requires the specification of the marginal density of model factors that are not conditionally related to other variables. For example, the model $[A, B, C] = [A \mid B][B \mid C][C]$ requires specification of the marginal density of C, $[C]$. We propose simple specifications based on the multivariate normal distribution:

$$x_{i,k} \sim \text{MVN}(\mu_{x,k}, \Psi_{x,k}), \quad b_{i,k} \sim \text{MVN}(\mu_{b,k}, \Psi_{b,k}), \quad z_{i,k} \sim \text{MVN}(\mu_{z,k}, \Psi_{z,k}), \tag{5}$$

where the μs denote the mean vectors and the Ψs are covariance matrices associated with the marginal distributions, which are assumed to be multivariate normal. In the analysis reported below, we tested the assumption of using multivariate normal distribution by performing a robust test: the E-statistic (energy) test of multivariate normality (Székely and Rizzo 2005). For the E-statistic test, we used *mvnorm.etest* in the R statistical package, obtaining an E-statistic value of 1.21 and an associated test p-value of >0.5 (a significant test statistic in this case would indicate that the data are not multivariate normal). Thus, we find support for the multivariate normal assumption.

The proposed model is highly parameterized as a result of the response segments (k) and the segment-specific coefficient matrices Γ_k and Δ_k characterizing the conditional relationships among variables. We deal with the issue of model dimensionality using the Bayesian variable selection developed by George and McCulloch (1993) and extended by Brown et al. (1998) for multivariate regression. Bayesian variable selection is implemented by assuming a mixture prior distribution for the model coefficients that places mass close to zero for the mixture components corresponding to the variable not being selected. Thus, posterior estimates based on these priors are centered away from zero if the data provide fairly strong support of a nonzero effect. Otherwise, the posterior is concentrated near zero. In the analysis reported below and detailed in the electronic companion, we specify the prior such that effect sizes smaller than 0.05 are shrunk to zero. We tried different values for shrinkage point per George and McCulloch (1993) and found that our results are robust to its selection. The choice of the shrinkage was such that the value of effect sizes can be safely replaced by zero. We believe this to be a flexibility of our model, as managers can exploit this by varying the shrinkage to what best suits their needs. Additional details of the derivation, alternative approaches, estimation algorithms, and pointers to recent research on Bayesian variable selection, as well as guidance for selection of tuning parameters, are provided in George and McCulloch (1993, 1997), Brown et al. (1998), and George (2000).

Details for estimating models of hierarchical effects are provided in an electronic companion. Also contained in the electronic companion is the result of a simulation study that provides evidence that model parameters can be successfully recovered and that departures from the true model are reliably detected by estimates of the log marginal likelihood, a Bayesian measure used to test hypotheses. We also calculated the log marginal density (Newton–Raftery approximation) for all the models considered and found it to be consistent with the log marginal likelihood calculations in determining the best-fitting model (details of these estimates can be obtained from the authors upon request).

3. Empirical Study

We investigate the presence of hierarchical effects using data from a national brand-tracking study of a leading luxury automobile. Study participants completed an extensive questionnaire designed to elicit their attitudes and opinions toward the focal brand; their likelihood of purchasing a vehicle (y); intended consumer actions (X), such as taking a test drive in the next six months; brand beliefs (B), such as durability and manufacturing quality; and their exposure to the focal brand through a variety of distinct media sources (Z), such as television, radio, Internet, and direct mail. A total of 6,177 observations were available for analysis.

Descriptive statistics for the data are reported in Table 1. Likelihood of purchase (y), intermediate actions (X), and brand beliefs (B) were collected on an 11-point scale, with 0 indicating "not at all" and 10 indicating "extreme" likelihood of purchase or action within the next six months. Media exposure variables (Z) were self-reported recall measures of the number of exposures to the brand via the indicated media during the prior six months. Summary statistics reported in Table 1 suggest that there is sufficient variation in the data for analysis. In addition, the data contain a broad set of media variables under partial control of management. The goal of our analysis is to explore conditional relationships between the batteries of variables X, B, and Z and their impact on likelihood of purchase (y).

4. Results

We begin our analysis by investigating homogenous and naïve models. The homogenous model

Table 1 **Descriptive Statistics** ($N = 6{,}177$)

Variable	Mean	Std. dev.
Dependent variable (y^*)		
Likelihood of purchase	3.10	3.44
Intermediate action variables (X^*)		
Go to a retailer	2.26	3.20
Seek info directly	2.58	3.33
Seek info from objective source	2.66	3.36
Recommend to a friend	3.48	3.68
Read mail	4.36	3.79
Take a test drive	2.32	3.21
Brand belief variables (B^*)		
Durability	7.72	2.33
Security	7.53	2.33
Excitement	6.40	2.48
Design	6.75	2.45
Innovation	6.98	2.39
Manufacturing quality	7.84	2.28
Overall impression of the brand	7.77	2.33
Media exposure variables (Z)		
Magazine advertisement	3.65	6.25
News advertisement	3.81	8.70
Radio advertisement	3.80	12.24
Television advertisement	7.15	12.65
Sponsor advertisement	0.69	2.53
Internet advertisement	0.86	3.96
Direct mailing	0.35	1.69
Brochure	0.44	2.03
Focal brand's website	0.45	2.43
Dealership website	0.23	1.16
Independently published article	0.71	2.33
Independent website	0.40	2.43
Display in public place	1.32	5.05

*Variables measured on a 0–10 scale.

Table 2 **Model and Submodel Fit Statistics**

	Hierarchical models	Log marginal likelihood
1	$[y \mid X][X][B][Z]$	−10,907
2	$[y \mid B][X][B][Z]$	−12,147
3	$[y \mid Z][X][B][Z]$	−12,308.6
4	$[y \mid X, B][X][B][Z]$	−10,709
5	$[y \mid X, Z][X][B][Z]$	−10,610
6	$[y \mid B, Z][X][B][Z]$	−13,341
7	$[y \mid X, B, Z][X][B][Z]$	−9,941.9
8	$[y, X, B, Z]$	−17,817
7(1)	$[y \mid X, B, Z][X \mid B][B \mid Z][Z]$	**−9,859.3**
7(2)	$[y \mid X, B, Z][X \mid B, Z][B][Z]$	−9,906.2
7(3)	$[y \mid X, B, Z][X \mid B][B][Z]$	−9,937.4
7(4)	$[y \mid X, B, Z][X \mid Z][B][Z]$	−9,998.2

assumes the presence of one response segment ($K = 1$). The naïve models include (i) identifying response segments using the marginal distribution of the dependent variable and (ii) estimating homogenous regression models. We find that any kind of homogenous or naïve model does not capture the impact of media on purchase intent; i.e., media effects are estimated to be insignificant. The details of these results are provided in the Online Appendix B of the electronic companion.

We now turn to our proposed analysis based on alternative factorizations of the joint distribution of the data. Table 2 reports the log marginal likelihood of the various factorizations of the data where we attempt to understand a direct relationship between purchase intention and all other variables. We compute the log marginal likelihood based on Chib (1995). The details of the computation of log marginal likelihood are provided in the electronic companion. We investigate all possible combinations of the impact of X, B, and Z on y. Our preliminary analysis indicates that model 7, which postulates direct effects between intended actions (X), brand beliefs (B), and media recall (Z) on purchase intention (y), fits the data best.

We find large gains in the log marginal likelihood when the intended action variables (X) are included as covariates in the conditional distribution of purchase intention in Equation (1). We also find that a model of the joint distribution that does not employ any factorization $[y, X, B, Z]$ exhibits poor fit to the data. Thus, response segments formed in terms of only y and X lead to better fit than response segments based on all the data.

We compare the fit of a finite mixture SEM against two different model specifications. The finite mixture SEM (for variables y, X, B, and Z) is specified with four segments (similar to our proposed models) and includes three latent variables. The details of the modeling approach and the WinBUGS code are provided in Lee (2008, Chapter 11). The trace plot for the mixture weights of the SEM model for the automobile data is provided in Figure 1 of Online Appendix B of the electronic companion. The finite mixture SEM was run for 10,000 iterations, and the last 5,000 iterations were used for analysis. We compare this to two different specifications of our proposed hierarchical models ($[y \mid X, B, Z][X][B][Z]$ and $[y \mid X, B, Z][X \mid B][B \mid Z][Z]$) with four segments, and encounter a large decrease in log marginal density, from $-17{,}883$ and $-16{,}732$, respectively, for our models to $-18{,}974$ for the finite mixture SEM. Because we compare a heterogeneous SEM model with our proposed heterogeneous response segment models, we can deduce that the use of a latent factor structure to measure dependencies among variables results in a large loss of information. Our proposed methodology provides the benefit of incorporating flexible conditional structures.

Finally, we also compare our hierarchical model against simple graphical models. One of the issues that we run into with the existing graphical model packages is the lack of support for large numbers of variables and observations. Also, most of the graphical models do not account for heterogeneity and variable selection, and this is an area of significant interest in current statistics literature (Jones et al. 2005). We used the mimR (mixed interaction models) package for undirected Gaussian graphical modeling in R. This package supports relatively larger data sets but does not provide the flexibility to incorporate different response structures accounted for in the model. The *mim* function in mimR is used to specify the graphical model for the automobile data. We compare the graphical model $[y \leftarrow X \leftarrow B \leftarrow Z]$ with our proposed model $[y \mid X][X \mid B][B \mid Z][Z]$. The *summary* and *properties* commands provide information about the model fit for the simple graphical model. We find that for our data, a simple graphical model (that does not account for heterogeneous response segments and variable selection) does much worse (log marginal density of $-24{,}092$ for the graphical model compared with $-17{,}558$ for our model). Thus, simple graphical analysis conducted with existing software misses important characteristics of the data.

We continue our analysis by examining refinements of the best-fitting models. The fit of these models is also reported in Table 2. The fit statistic indicates that the joint distribution of X, B, and Z is best factored hierarchically as $[X, B, Z] = [X \mid B][B \mid Z][Z]$. Thus, we find evidence for both direct and indirect media effects—direct in the sense that media (Z) directly influences purchase intention (y) and indirect in the sense that media influence on brand beliefs, and brand beliefs influence intended actions (see also Orth and Marchi 2007). The measurement of the influence of media on purchase intentions must therefore account for both direct and indirect effects.

Parameter estimates for our best-fitting model, model 7(1), are displayed in Tables 3–5. We report parameter estimates that have more than 95% of posterior mass away from zero. Table 3 reports parameter estimates $\{\beta_k\}$ for the normal mixture model in Equation (1), with brand beliefs (B) and media recall (Z) also included as explanatory variables. We find support for four latent segments of respondents ($K = 4$), where the explanatory variables are estimated to have different effects across segments and are found to be much larger than those obtained through the pooled analysis with $K = 1$ (see Table 1 in Online Appendix B of the electronic companion). Also reported at the bottom of Table 3 is the average value of purchase intention for each segment. The latent segments are ordered in the table so that segment 1 has the smallest average purchase intention (0.12 on an 11-point scale), and segment 4 has the largest (6.08). A detailed examination of the results is offered in §5.

Table 4 reports multivariate Bayesian variable selection parameter estimates $\{\Gamma_k\}$ for the auxiliary regression of the consumer actions (X) on brand beliefs (B). We only report estimates that have more than 95% of the posterior mass away from zero. Displayed are

Table 3 Direct Effects of Media, Brand Beliefs, and Intended Actions on Purchase Intention $\{\beta_k\}$

$[y \mid X, B, Z]$	Segment 1	Segment 2	Segment 3	Segment 4
β_X				
Intercept		0.76		
Go to a dealer			0.92	0.07
Seek info directly				0.1
Seek info from obj. source		−0.09		−0.07
Recommend to a friend				0.1
Read mail				
Take a test drive				
β_B				
Overall impression of each make				
Durability				
Security				
Excitement				
Design				
Innovation				
Manufacturing quality				0.23
β_Z				
Magazine ad	2.5	−0.40		
Newspaper ad	1	−0.24		0.15
Radio ad		0.1	0.02	0.06
TV ad				0.05
Sponsorship event				
Internet ad				
Direct mail		0.17		0.47
Brochure		−0.15		
Company website				
Dealer website				
Independent article		−0.08		
Independent website		0.16		
Public display				
σ	0.60	0.22	2.6	0.75
Mean: y	0.12	1.49	3.65	6.08
Weights: Φ_k	0.37	0.08	0.18	0.37

Note. Reported estimates have more than 95% of posterior mass away from zero.

the results for the four segments. Overall, we find large coefficient estimates. Consider, for example, the effect of the "overall impression of the make" (OIM) on consumer actions. OIM coefficients are found to have a large effect on the intended actions, with an increase of just one unit associated with an increase on the intention scale of approximately 0.50 for all relevant actions in segment 4. Also, OIM has a differential impact on different segments. Thus, we find that the effects of brand beliefs on intended consumer actions are large and are segment-specific. These intermediate effects are masked in pooled or homogenous analysis.

Table 5 reports multivariate Bayesian variable selection parameter estimates $\{\Delta_k\}$ for the auxiliary regression of brand beliefs (B) on media exposure (Z). We only report estimates that have more than 95% of

Table 4 **Effects of Brand Beliefs on Intended Actions** $\{\Gamma_k\}$

Γ for $[X \mid B]$	Go to a dealer	Seek info directly	Seek info from obj. source	Recommend to a friend	Read mail	Take a test drive
Segment 1						
Intercept	−0.23	−0.28		−1.31	−0.83	
Overall impression of each make	0.05	0.07	0.06	0.2	0.23	0.05
Durability						
Security			0.08			
Excitement						
Design						
Innovation				0.15	0.13	
Manufacturing quality						
Segment 2						
Intercept				−2.01		
Overall impression of each make	0.22	0.28	0.25	0.46	0.41	
Durability						
Security						
Excitement						0.21
Design						
Innovation						
Manufacturing quality	−0.31	−0.27				
Segment 3						
Intercept	−1.93	−1.57	−1.41	−3.23	−1.64	−1.67
Overall impression of each make	0.29	0.33	0.28	0.43	0.45	0.25
Durability						
Security	0.32	0.32	0.3		0.26	0.25
Excitement	0.3	0.27	0.24	0.3	0.2	0.27
Design					0.21	
Innovation						
Manufacturing quality	−1.34	−0.37	−0.32			−0.32
Segment 4						
Intercept				−2.96		
Overall impression of each make	0.33	0.32	0.34	0.59	0.49	0.28
Durability						
Security					0.15	
Excitement	0.24	0.19	0.16	0.17		0.2
Design				0.16	0.13	0.14
Innovation		0.15	0.17			
Manufacturing quality			−0.34	−0.24	−0.16	−0.3

Note. Reported estimates have more than 95% of posterior mass away from zero.

the posterior mass away from zero. Displayed are the results for the four segments. We again find large coefficient estimates. The effect of brochures, for example, is large for segment 1, where an increase of just one exposure is associated with an increase of approximately 0.62 to 1.02 across all brand beliefs.

Finally, Table 6 reports aggregated estimates of direct and indirect media affects across all segments. The estimates for the best-fitting model are based on the relationship $[y, X, B, Z] = [y \mid X, B, Z][X \mid B][B \mid Z][Z]$:

$$\begin{aligned}\frac{\partial y}{\partial z'} &= \sum_{k=1}^{K} \varphi_k \left(\frac{\partial y}{\partial z'_k} + \frac{\partial y}{\partial b'_k}\frac{\partial b_k}{\partial z'_k} + \frac{\partial y}{\partial x'_k}\frac{\partial x_k}{\partial b'_k}\frac{\partial b_k}{\partial z'_k} \right) \\ &= \sum_{k=1}^{K} \varphi_k (\beta'_{z_k} + \beta'_{b_k}\Delta_k + \beta'_{x_k}\Gamma_k\Delta_k).\end{aligned} \tag{6}$$

The first term in parentheses on the right side of Equation (6) is the direct, segment-specific effect of media exposure on purchase intention ($Z \rightarrow y$). The second term is the indirect effect of media exposure that operates through an effect on brand beliefs ($Z \rightarrow B \rightarrow y$). The third term is the indirect effect where media affects brand beliefs, which then affect intermediate actions and purchase intentions ($Z \rightarrow B \rightarrow X \rightarrow y$). We explore in detail the implication of coefficient, direct, and indirect effects estimates in the next section.

5. Discussion

Our analysis of media affects on brand beliefs, intended actions, and purchase intentions reveals the presence of direct and indirect effects that are large

Table 5 **Effects of Media Exposure on Brand Beliefs $\{\Delta_k\}$**

Δ for $[B \mid Z]$	Overall impression of each make	Durability	Security	Excitement	Design	Innovation	Manufacturing quality
			Segment 1				
Intercept	6.69	6.69	6.45	5.5	5.85	5.93	6.88
Magazine ad							
Sponsorship event		0.09	0.1				0.09
Internet ad			0.06				
Direct mail							
Brochure	0.89	0.62	0.66	0.85	1.02	0.88	0.66
Company website							
Dealer website			0.05			0.06	0.06
			Segment 2				
Intercept	7.97	7.84	7.54	6.4	6.79	6.99	7.91
Magazine ad							
Sponsorship event							
Internet ad							
Direct mail							
Brochure							
Company website							
Dealer website							
			Segment 3				
Intercept	7.94	7.83	7.61	6.67	7.02	7.1	7.94
Magazine ad							
Sponsorship event							
Internet ad							
Direct mail							
Brochure							
Company website							
Dealer website							
			Segment 4				
Intercept	8.23	8.08	7.88	6.77	7.16	7.3	8.15
Magazine ad		0.03	0.03				
Sponsorship event							
Internet ad							
Direct mail	0.13	0.13	0.16	0.15	0.11	0.12	0.12
Brochure							
Company website							
Dealer website	0.05					0.05	

Note. Reported estimates have more than 95% of posterior mass away from zero.

Table 6 **Aggregate Effect of Media Exposure (Z) on Purchase Intention (y)**

Media	$\frac{\partial y}{\partial Z'} = \sum_{k=1}^{K} \phi_k(\beta'_{z_k} + \beta'_{b_k}\Delta_k + \beta'_{x_k}\Gamma_k\Delta_k)$
Magazine ad	0.89
Newspaper ad	0.41
Radio ad	0.03
TV ad	0.02
Sponsorship event	0.01
Internet ad	0.00
Direct mail	0.20
Brochure	0.04
Company website	0.00
Dealer website	0.01
Independent article	−0.01
Independent website	0.01
Public display	0.00

in magnitude. In contrast, a naïve analysis based on aggregated effects indicates that media have little influence on any of the variables in our analysis. Because Tables 3, 4, and 5 are self-explanatory, we do not discuss all the effects (there are quite a few interesting effects). Instead, we highlight only the key takeaways and a few managerially interesting hierarchical effects among the variables. We believe that these hierarchical effects can provide managers with key insights about media exposure and consumer behavior (beliefs and actions).

5.1. Direct Effects on Purchase Intention $\{\beta_k\}$

We find that the segment-specific direct effects reported in Table 3 indicate media exposure (Z), brand beliefs (B), and consumer actions (X) to have a direct effect on purchase intent (y). Segments 1–4 are arranged from the lowest to highest likelihood of purchase values. In segment 1, the effect of magazine and newspaper advertising is large (with effect sizes of 2.5 and 1.0, respectively). Respondents in segment 2 are identified as being affected by a large number of media vehicles (magazine, newspaper, radio, direct mail, etc.). Respondents in segment 3 are not affected by media, and respondents in segment 4 are affected by direct mail and newspaper. From a managerial standpoint, direct effects on media can be summarized as follows: (i) segment 1, large effect sizes; (ii) segment 2, broad range of media; (iii) segment 3, no effect; and (iv) segment 4, small effects. Similarly, direct impact of brand beliefs (manufacturing quality) affects only respondents of segment 4 but does not have any effect for segments 1–3.

Overall, we find that direct effects of media are largest for respondents least likely to make a purchase in the next six months and are smallest for respondents who are most ready to buy. Thus, our results indicate that media advertising is most effective in the initial stages of the purchase process, possibly being associated with consideration set formation. The large effect sizes across different segments (compared with pooled or homogenous models) can aid marketing managers in media allocation across different media vehicles and can help them understand consumer behavior across different segments.

5.2. Effects of Brand Beliefs on Intended Actions $\{\Gamma_k\}$

The effects of brand beliefs on consumer actions reported in Table 4 indicate the presence of large and differential effects across response segments. Respondents' belief about the "overall impression of the make" is seen to be consistently associated with all of the intended actions examined in the study. The influence of this belief on intended actions is weakest in segment 1, where respondents have the lowest likelihood of making a purchase, and is largest in segments 3 and 4, where a respondent purchase is more likely. Thus, beliefs about overall impression can provide managers with a useful measure for predicting consumer desire to take the next step toward making a vehicle purchase.

Significant coefficients are more abundant in segments 3 and 4, which comprise respondents who are more likely to make a purchase in the next six months. For segment 3, numerous brand beliefs (i.e., OIM, security, excitement, and manufacturing quality) are found to play a significant role. If our model were only to account for structural heterogeneity $[y \mid X, B, Z]$ without incorporating a hierarchical component $[X \mid B]$, analysis of the data would suggest that brand beliefs are not significantly associated with purchase intent for these segments (see Table 3). In contrast, we find in segment 3 that brand beliefs are significantly associated with going to a dealer (see Table 4), which in turn has a significant association on purchase intent (0.92 in Table 3). Thus, an indirect hierarchical effect of brand beliefs on purchase intent is present.

Likewise, respondents in segment 4 demonstrate an interesting hierarchical effect associated with manufacturing quality (MQ). The significant MQ coefficients reported in Table 4 for segment 4 are all negative, implying that more positive beliefs of MQ are associated with a decrease in the likelihood of taking actions associated with seeking objective information, making recommendations, reading direct mail, and taking a test drive. It seems counterintuitive that higher MQ would have a negative impact. This can be partly explained by investigating our hierarchical model in more detail. From Table 3, we see that MQ has a direct positive impact on purchase intent (0.23; see Table 3) for this segment. Thus, although more positive beliefs about MQ are associated with respondents being less likely to take a test drive or read mail, the true impact of higher brand beliefs of MQ is twofold; it leads to less likely intended actions but higher values of intended purchase.

5.3. Effects of Media Exposure on Brand Beliefs $\{\Delta_k\}$

The effects of media exposure on brand beliefs reported in Table 5 indicate significant associations in segments 1 and 4 but not in segments 2 and 3. Respondents in segment 1 are influenced by numerous media vehicles (sponsorship event, Internet ad, brochure, and dealer website); the most significant is the impact of a brochure on an individual's brand beliefs (ranging from 0.62 to 1.02; see Table 5). For respondents in segment 4, magazine ads, direct mail, and the dealer website have a significant association with brand beliefs, with direct mail having the largest effects. Moreover, the effect sizes in segment 1 are generally larger than those reported for segment 4.

In general, we find that models that do not incorporate heterogeneous response segments fail to capture direct effects of media on purchase intent (compare Tables 3 and 5 to Table 1 in Online Appendix B of the electronic companion). Hierarchical effects measurement is also adversely affected (compare Tables 4 and 5 to Table 3 in Online Appendix B in the electronic companion). Bayesian model formulation involving the joint distribution of the data, coupled with structural heterogeneity and Bayesian variable selection, results in a richer description of the data. Finally, Table 6 demonstrates that our model captures significant effect sizes of the impact of advertising (after accounting for heterogeneous response segments) on purchase intention as opposed to a homogenous model (see Table 1 in Online Appendix B of the electronic companion).

We find that the indirect effects are diminished when combined together as in the second and third terms of Equation (6). This is primarily a result of multiplying the indirect effects together. For example, effect sizes as large as 0.2 for each of $Z \to B$, $B \to X$, and $X \to y$ would result in an overall effect size of $0.2 \times 0.2 \times 0.2$, i.e., 0.008. But this does not imply that we can ignore the impact of these indirect effects. An advertising campaign manager would be interested in demonstrating that the true effects of advertising are not just the direct effects but these indirect effects as well that, in essence, capture the total impact of the advertising. This adds further emphasis on using our proposed modeling approach that can capture these intermediate effects (Tables 3–5) and demonstrate the significance of these intermediate effects without using some form of aggregation. Also, our approach handles a common challenge in the analysis of industry data and hierarchical effects, i.e., a large number of variables present that reflect diverse aspects of behavior.

6. Conclusion

This paper presents a Bayesian approach for the analysis of hierarchical effects in cross-sectional data. We investigate a variety of models with varying conditional structures, response segments, and effect hierarchies. The model employs a mixture of likelihoods pertaining to latent segments of individuals that are structurally distinct and related to a set of intended consumer actions, brand beliefs, and media exposure variables. Brand beliefs and media exposure variables are also associated with these segments in an auxiliary specification. We deal with the high dimension of variables using Bayesian variable selection, and we find evidence of large effect sizes that are segment specific. We associate the latent segments with phases of a purchase decision using the average value of each segment's intent to purchase. The low purchase intention segment comprises individuals who have not yet made up their minds to purchase and who are found to be highly influenced by media exposure. We demonstrate that these effect sizes cannot be captured by pooled or homogenous models, or models with unreasonable partitioning (i.e., partitioning of the dependent variable). Moreover, we find that a pooled analysis that directly relates media exposure to purchase likelihood shows effect sizes to be near zero.

Hierarchical effects models are discussed widely in marketing and advertising textbooks. There is a long history of extended models of behavior in marketing, beginning with motivating conditions, moving to desired attributes and benefits, and ending with a purchase in the marketplace. Our model facilitates analysis among specific items across these components through the introduction of latent segments of respondents with qualitatively different responses. In contrast to numerous published articles that point to ineffective effects and the lack of a hierarchical decision structure (e.g., Weilbacher 2001, Vakratsas and Ambler 1999, Palda 1966), our analysis finds empirical support for the presence of a hierarchical purchase process for high-involvement products.

In addition to the results reported in this paper, our model could potentially be used to investigate the influence of lagged versus immediate hierarchical effects. If we assume that consumers move sequentially through the latent segments identified in our model, we could investigate the future effects of media exposure by incorporating across-segment calculations. For example, media exposure for respondents in segment 1 may move them to segment 2, for which indirect effects could then be added. Such analysis, however, is speculative in the absence of panel data that track specific respondents through time, allowing insights into the speed of movement across segments. The results reported in our analysis are correlational in nature, and although we find evidence in support of a staged decision process, our analysis does not offer conclusive proof of its existence. We leave this as a topic for future research. We also believe that a methodology for blocking different sets of variables, if developed, would be a powerful tool in conjunction with our proposed approach.

Another avenue for future research is the further development of mediation analysis. Researchers often refer to the notion of "partial mediation," where an intervening variable does not completely reflect the property of conditional independence (Kim 2007, Wood et al. 2008). If the variable B fully mediates the influence of C on A, then the conditional distribution of A given B $[A \mid B]$ is equal to the conditional distribution of A given B and C $[A \mid B, C]$. That is, C does not help explain A in the presence of B. One explanation for the presence of partial mediation is the presence of heterogeneous response segments, where full mediation is present for some but not all of the segments. Thus, the segments weights $\{\phi_k\}$ in Equation (1) can be used to assess the proportion of respondents for which mediation is full. Such a metric is absent in current approaches to mediation analysis and offers a new measure of mediation strength.

Marketing managers are often faced with the task of justifying advertising budgets across a variety of media such as television, radio, and direct mail. One reason for the presence of positive budget allocations to various media is the qualitatively distinct effects they have in the process of consumer decision making. A television advertisement, for example, may not be sufficiently persuasive to motivate prospects to make an immediate purchase. It could, however, prompt them to seek additional product information (e.g., *Consumer Reports*), which could then lead to purchase. In this case, the impact of the advertisement is mediated through an intermediate action and has a hierarchical, or intermediate, effect. Attempts to define advertising effectiveness as the direct link to purchase would understate its actual effect by overstating its role. We believe the investigation of hierarchical effects is a fruitful area of future research in marketing.

7. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mktsci.pubs.informs.org/.

References

Assael, H. 1993. *Marketing Principles and Strategy*, 2nd ed. Dryden Press, Fort Worth, TX.

Baron, R. M., D. A. Kenny. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic and statistical considerations. *J. Personality Soc. Psych.* **51**(6) 1173–1182.

Brown, P. J., M. Vannucci, T. Fearn. 1998. Multivariate Bayesian variable selection and prediction. *J. Roy. Statist. Soc. Ser. B* **60**(3) 627–641.

George, E. I. 2000. The variable selection problem. *J. Amer. Statist. Assoc.* **95**(452) 1304–1308.

George, E. I., R. E. McCulloch. 1993. Variable selection via Gibbs sampling. *J. Amer. Statist. Assoc.* **88**(423) 881–889.

George, E. I., R. E. McCulloch. 1997. Approaches for Bayesian variable selection. *Statistica Sinica* **7**(2) 339–373.

Holbrook, M. B., R. Batra. 1987. Assessing the role of emotions as mediators of consumer responses to advertising. *J. Consumer Res.* **14**(3) 404–420.

Iacobucci, D., N. Saldanha, X. Deng. 2007. A meditation on mediation: Evidence that structural equations models perform better than regressions. *J. Consumer Psych.* **17**(2) 139–153.

Jedidi, K., H. S. Jagpal, W. S. DeSarbo. 1997. Finite-mixture structural equation models for response-based segmentation and unobserved heterogeneity. *Marketing Sci.* **16**(1) 39–59.

Jones, B., C. Carvalho, A. Dobra, C. Hans, C. Carter, M. West. 2005. Experiments in stochastic computation for high-dimensional graphical models. *Statist. Sci.* **20**(4) 388–400.

Kim, H.-S. 2007. A multilevel study of antecedents and a mediator of employee–organization relationships. *J. Public Relations Res.* **19**(2) 167–197.

Lavidge, R. J., G. A. Steiner. 1961. A model for predictive measurements of advertising effectiveness. *J. Marketing* **25**(6) 59–62.

Letac, G., H. Massam. 2007. Wishart distributions for decomposable graphs. *Ann. Statist.* **35**(3) 1278–1323.

Orth, U. R., R. D. Marchi. 2007. Understanding the relationships between functional, symbolic, and experiential brand beliefs, product experiential attributes, and product schema: Advertising-trial interactions revisited. *J. Marketing Theory Practice* **15**(3) 219–233.

Palda, K. S. 1966. The hypothesis of a hierarchy of effects: A partial evaluation. *J. Marketing Res.* **3**(1) 13–24.

Strong, E. K., Jr. 1925. Theories of selling. *J. Appl. Psych.* **9**(1) 75–86.

Székely, G. J., M. L. Rizzo. 2005. A new test for multivariate normality. *J. Multivariate Anal.* **93**(1) 58–80.

Tanner, M. A., W. H. Wong. 1987. The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.* **82**(398) 528–540.

Vakratsas, D., T. Ambler. 1996. Advertising effects: A taxonomy and review of concepts, methods, and results from the academic literature. Working Paper 96–120, Marketing Science Institute, Cambridge, MA.

Vakratsas, D., T. Ambler. 1999. How advertising works: What do we really know? *J. Marketing* **63**(1) 26–43.

Weilbacher, W. M. 2001. Point of view: Does advertising cause a hierarchy of effects? *J. Advertising Res.* **41**(6) 19–26.

Whittaker, J. 1990. *Graphical Models in Applied Multivariate Statistics*. John Wiley & Sons, New York.

Wood, R. E., J. S. Goodman, N. Beckmann, A. Cook. 2008. Mediation testing in management research: A review and proposals. *Organ. Res. Methods* **11**(2) 270–295.

Zhao, X., J. G. Lynch Jr., Q. Chen. 2010. Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *J. Consumer Res.* **37**(2) 197–206.