SUNGHOON KIM, SIMON J. BLANCHARD, WAYNE S. DESARBO, and DUNCAN K.H. FONG*

Researchers have recently introduced a finite mixture Bayesian regression model to simultaneously identify consumer market segments (heterogeneity) and determine how such segments differ with respect to active regression coefficients (variable selection). This article introduces three extensions of this model to incorporate managerial restrictions (constraints). The authors demonstrate with synthetic data that the new constrained finite mixture Bayesian regression models can be used to identify and represent several constrained heterogeneous response patterns commonly encountered in practice. In addition, they show that the proposed models are more robust against multicollinearity than traditional methods. Finally, to illustrate the proposed models' usefulness, the authors apply the proposed constrained models in the context of a service quality (SERVPERF) survey of National Insurance Company's customers.

Keywords: Bayesian regression models, market segmentation, variable selection, managerial constraints, finite mixture models, heterogeneity

Online Supplement: http://dx.doi.org/10.1509/jmr.13.0056

Implementing Managerial Constraints in Model-Based Segmentation: Extensions of Kim, Fong, and DeSarbo (2012) with an Application to Heterogeneous Perceptions of Service Quality

A large number of data analytic procedures, including correlation, regression, structural equation models, partial least squares, and latent class methods, can be used to analyze survey data such as consumer service evaluations (e.g., Brady and Cronin 2001; Ladhari 2009). It is well documented that when data come from consumer surveys with limited sample sizes and a relatively large number of potential predictors that are highly intercorrelated, there are often problems of multicollinearity and coefficient instability with inflated standard errors (see, e.g., Brown, Churchill, and Peter 1993; Drolet and Morrison 2001).

Recently, several marketing scholars have introduced models that can reduce these concerns by simultaneously incorporating respondent heterogeneity and variable selection. For example, Chandukala et al. (2011; see also Chandukala, Edwards, and Allenby 2011) extend the individual-level variable selection approach by Gilbride, Allenby, and Brazell (2006) to the segment-level case by incorporating covariates that help identify active regression coefficients. More recently, Kim, Fong, and DeSarbo (2012) proposed a Bayesian model (hereinafter referred to as the "unconstrained

*Sunghoon Kim is a doctoral candidate (e-mail: suk220@psu.edu), Wayne S. DeSarbo is Smeal Distinguished Professor of Marketing (e-mail: wsd6@psu.edu), and Duncan K.H. Fong is Professor of Marketing and Statistics (e-mail: i2v@psu.edu), Smeal College of Business, Pennsylvania State University. Simon J. Blanchard is Assistant Professor of Marketing, McDonough School of Business, Georgetown University (e-mail: sjb247@georgetown.edu). This research was funded in part by Pennsylvania State University's Smeal College of Business. Fred Feinberg served as associate editor for this article.

model") that simultaneously performs segmentation and determines the optimal subset of independent variables per segment. The current research provides several extensions of the unconstrained model to incorporate managerial restrictions as well as to examine and test for patterns of heterogeneity in a confirmatory manner.

By allowing for consumer heterogeneity through the identification of segments, researchers must also consider the critical link between the segmentation and subsequent resource allocation decisions (Mahajan and Jain 1978; Montoya-Weiss and Calantone 1999; Winter 1979). Despite its intuitive importance, the actionability of segments is one of Kotler's (1999) most overlooked effective segmentation criteria. Here, "actionability" refers to the ability of model solutions to satisfy organizational, strategic, financial, technological, promotional, and other constraints. The marketing literature has long acknowledged the importance of considering implementation issues in marketing models (e.g., Little 1970). For example, Wind (1978) suggests that it is important to evaluate the expected market response, management objectives, and resources when evaluating market segment solutions.

Segment-level models can lead to solutions with many significant variables that are difficult to implement. As an illustration, consider a latent-class regression that identifies numerous significant predictors with coefficients differing between segments: how can a manager realistically implement the results? This issue is particularly evident in the case of what has perhaps been the most popular approach for measuring service quality over the past three decades: the SERVQUAL methodology developed by Zeithaml, Parasuraman, and Berry (1990). SERVQUAL contains 22 questions measuring service quality on five dimensions: reliability (the ability to perform the service dependably and accurately), assurance (the service provider's knowledge and the confidence they instill), tangibles (the facilities, written materials, and other physical evidence of the service), empathy (the level of attention given to consumers), and responsiveness (the ability to respond to customers' needs on a timely basis). Despite its extensive use, the SERVQUAL scale has been criticized on both conceptual and methodological grounds. According to Jain and Gupta (2004), the major objections to SERVQUAL involve the use of gap scores, the length of the questionnaire, the demands placed on respondents, the predictive power of the instrument, and the validity of the five-dimensional structure across varied application areas (see also Babakus and Boller 1992; Cronin and Taylor 1992). Perhaps the most damaging criticism involves the use of this disconfirmation gap framework: most studies have found a poor fit between the SERVQUAL service quality average scores and an overall service quality as measured through a single-item scale (Babakus and Boller 1992; Jain and Gupta 2004). Cronin and Taylor (1992) propose SERVPERF as an alternative measurement system. It eliminates the expectation aspect and focuses on the 22 performance items and an overall service quality measure. Many researchers have documented the superior performance of SERVPERF over SERVQUAL (e.g., Babakus and Boller 1992; Brady and Cronin 2001; Cronin and Taylor 1992; Jain and Gupta 2004).

The Appendix presents the 22 SERVQUAL/SERVPERF-based items from Parasuraman, Berry, and Zeithaml (1991) adapted to National Insurance Company ("National"), an application we explore in this article. We demonstrate that a possible solution to a regression applied to these 22 variables would likely lead not only to multicollinearity due to highly correlated items but also to many significant predictors that lack face validity, which marketing managers cannot use.

To avoid obtaining models that managers cannot easily implement, Mahajan and Jain (1978), Winter (1979), and DeSarbo and Grisaffe (1998) propose conceptual and mathematical segmentation frameworks that enable the incorporation of managerial and budget constraints. Others have placed parameter estimation restrictions to ensure that the interpretation of the final solution is simple and intuitive (Gordon 1996; Sriram 1990). We also believe that adding constraints directly in the analysis can increase the actionability of the solutions. Here, we present three types of constraints that can be beneficial for managers when using a segment-level regression that simultaneously performs variable selection. Specifically, we extend the Kim, Fong, and DeSarbo (2012) unconstrained model for segment-level variable selection to be more actionable by incorporating three specific sets of managerial constraints: common factors constraints (which require identical active predictors across segments but allow the coefficients to differ), distinctive factors constraints (which prohibit two or more segments from sharing the same active predictors), and dimension constraints (in which each segment can have, at most, one predictor among each a priori defined latent dimension). Although other constraints could be implemented, we believe the aforementioned to be particularly useful in a wide range of studies. We illustrate their usefulness through an empirical example in the following section.

To our knowledge, no latent structural model exists that simultaneously allows variable selection and accommodates such managerial constraints. In the next section, we describe the technical details of the constrained models and evaluate their performance using synthetic data.

THE PROPOSED CONSTRAINED BAYESIAN REGRESSION MODELS

Let Y_i be the dependent variable: consumer i's overall evaluation of a store, service, product, and so on. We assume that this overall consumer's evaluation can be described through a multiple regression specification such that

$$Y_i = \underline{X}_i^T \underline{\beta}_{H_i} + \varepsilon_i,\ i = 1, \ldots, n, \tag{1}$$

where $\underline{X}_i^T$ is a row vector of dimension P containing values of all the key predictors as well as an intercept term for respondent i, H_i = k represents an index identifying that consumer i belongs to market segment k, $\underline{\beta}_{H_i}$ is a (column) vector of segment-level regression coefficients, and the error terms ε_i are posited to be independently and normally distributed as $N(0, \sigma^2)$. We obtain a finite mixture regression model when the segment indicator variable $H_i \in \{1, \ldots, K\}$ is assumed to follow a discrete distribution with positive parameters $d_1, \ldots, d_K$:

(2) $$\pi(H_i|\underline{d}) = \text{discrete}(d_1, ..., d_K),$$

where discrete(·) denotes a discrete distribution, that is, $\pi(H_i = k|\underline{d}) = d_k$ and $\Sigma_{k=1}^{K} d_k = 1$. In our Bayesian approach, these parameters follow a Dirichlet distribution:

(3) $$\pi(d_1, ..., d_K) = \text{Dirichlet}(\alpha_1, ..., \alpha_K).$$

In addition, we employ a spike and slab prior on $\underline{\beta}_k$, k = 1, ..., K (see Kim, Fong, and DeSarbo 2012). For each segment-level coefficient β_{kp}, p = 1, ..., P, we assume that

(4) $$\begin{cases} \pi(\beta_{kp}|Z_{kp} = 1) = N(0, \tau_p^2) \\ \pi(\beta_{kp} = 0|Z_{kp} = 0) = 1 \end{cases},$$

where Z_{kp} is a Bernoulli random variable. When $Z_{kp} = 1$, the pth variable in the kth segment is selected and its coefficient (β_{kp}) is assumed to follow a normal distribution with zero mean, as is commonly assumed in the variable selection literature. When $Z_{kp} = 0$, the pth variable is not selected for the kth segment, and its coefficient (β_{kp}) is set to zero. Thus, the binary latent variable Z_{kp} is introduced to indicate whether variable p has an impact on the dependent variable for segment k. For the variances σ^2 and τ_p^2, we employ the commonly used inverse Gamma distributions with s_1, s_2, s_{p1}, and s_{p2} as hyperparameters:

(5) $$\pi(\sigma^2) = \text{InvGamma}(s_1, s_2), \text{ and}$$

(6) $$\pi(\tau_p^2) = \text{InvGamma}(s_{p1}, s_{p2}).$$

This unconstrained model from Kim, Fong, and DeSarbo (2012) simultaneously identifies segments of consumers who share same active predictors (within each segment); however, there are no constraints regarding variable selection and the estimation of regression coefficients for each segment. In this sense, the model is exploratory, and we expect it to work reasonably well overall in many applications, particularly when there are no a priori managerial issues that would require the use of constraints. However, it may not be the most parsimonious representation and may derive solutions that are not, in the end, actionable for managers. Web Appendix I-A shows the derivation of the associated full conditional distributions, and Web Appendix II-A summarizes the steps of the Markov chain Monte Carlo (MCMC) estimation for this unconstrained model.

The Dimension-Constrained Model

Managers are often presented model solutions that include correlated items across many a priori defined dimensions. When it is not feasible for managers to focus on multiple items from each dimension, they may want to know which they should prioritize in their allocation decisions. For example, with respect to four items from the tangibles dimension in SERVQUAL/SERVPERF, managers may not have sufficient resources (e.g., financial, organizational, temporal) to both modernize the equipment (item 1) and simultaneously improve the general look of the facilities (item 2) for all segments of consumers. For such cases, we can ensure that, at most, one variable can be selected with each defined latent dimension while allowing each segment to vary both which variable is active per selected dimension (if any) as well as the variable's effect (coefficient).

To implement such constraints, we group variables into Q mutually exclusive dimensions (higher-order factors) as predetermined by the manager. Let m_q, q = 1, ..., Q denote the number of items in each dimension ($\Sigma_{q=1}^{Q} m_q = P$), and

$$\underline{Z}_{kq} = \left[Z_{k\left(\sum_{j=1}^{q-1} m_j + 1\right)}, Z_{k\left(\sum_{j=1}^{q-1} m_j + 2\right)}, ..., Z_{k\left(\sum_{j=1}^{q} m_j\right)}\right], \quad \forall q = 1, ..., Q.$$

When $m_q = 1$, $\underline{Z}_{kq}$ is the single variable

$$Z_{k\left(\sum_{j=1}^{q-1} m_j + 1\right)},$$

and we assume it is equal to 1 with a probability of τ,

(7) $$\pi(\underline{Z}_{kq}|\tau) = \text{Bernoulli}(\tau).$$

When $m_q > 1$, because, at most, one variable can be selected with each dimension, we introduce a Bernoulli random variable r_{kq} and assume the following prior for $\underline{Z}_{kq}$:

(8) $$\begin{cases} \pi(\underline{Z}_{kq}|r_{kq} = 1) = \text{Multinomial}(1, \underline{w}_q) \\ \pi(\underline{Z}_{kq} = \underline{0}|r_{kq} = 0) = 1 \end{cases},$$

where $\underline{w}_q = (w_{q1}, ..., w_{qm_q})$ and $\Sigma_{l=1}^{m_q} w_{ql} = 1$. Thus, when $r_{kq} = 1$, one variable from the qth dimension will be selected because $\underline{Z}_{kq}$ follows a multinomial distribution with parameters 1 and $\underline{w}_q$. When $r_{kq} = 0$, all values in $\underline{Z}_{kq}$ are set as zero, and none of the variables are selected. Then, we assume an exchangeable prior for r_{kq} with parameter τ following a Beta distribution:

(9) $$\pi(r_{kq}) = \text{Bernoulli}(\tau), \text{ and}$$

(10) $$\pi(\tau) = \text{Beta}(t_1, t_2),$$

with t_1 and t_2 as hyperparameters. Finally, $\underline{w}_q$ is assumed to follow a Dirichlet distribution:

(11) $$\pi(w_{q1}, ..., w_{qm_q}) = \text{Dirichlet}(\theta_{q1}, ..., \theta_{qm_q}).$$

We note that all the full conditional distributions of this model are standard probability distributions, so we use a Gibbs sampling algorithm to generate a random sample from the joint posterior distribution. Web Appendix I-B shows the derivation of the associated full conditional distributions, and Web Appendix II-B delineates the specific steps of our proposed MCMC estimation for this dimension constrained model. Finally, we note that the unconstrained model can be obtained when $m_q = 1$ and Q = P are set in the dimension-constrained model.

The Common Factors Model

Although managers often recognize the presence of consumer heterogeneity, they may not have sufficient resources to improve on all of the identified elements lacking for each targeted segment. Rather than run an aggregate model masking respondent heterogeneity, we argue that constraints could instead be imposed on the selection and estimation of heterogeneous coefficients by introducing common factors constraints in which the most important predictors are identified for the entire sample (i.e., across all derived segments) but the corresponding coefficients are permitted to vary across segments. The common factors model can be obtained from the unconstrained model by setting $Z_{kp} = Z_p$ for all k. Namely, each variable is either activated for all segments or not at all, but when activated, the regression coefficients are free to vary between segments. Web Appendix I-C shows the derivation of the associated full conditional distributions, and Web Appendix II-C delineates the specific steps of our proposed MCMC estimation for the common factors model.

The Distinctive Factors Model

It may be of utmost importance to managers that the segments respond differently to alternative marketing mixes (Brusco, Cradit, and Stahl 2002)—for example, when a manager is interested in selecting a combination of communication media to match segment-level reactivity (Montoya-Weiss and Calantone 1999). In practice, adding distinctive factor constraints into the model would prohibit the activation of one variable for more than one segment at a time (except the intercept). The general setup of this model is the same as that of the dimension-constrained model, with the following major differences. Let $\underline{Z}_p = (Z_{1p}, ..., Z_{Kp})$ and let β_{k1} be the intercept. We assume the following:

(12) $$\pi(Z_{k1}|\tau) = \text{Bernoulli}(\tau), \text{ for all k, and}$$

(13) $$\begin{cases} \pi\left(\underline{Z}_p \middle| r_p = 1\right) = \text{Multinomial}(1, \underline{w}) \\ \pi\left(\underline{Z}_p = \underline{0} \middle| r_p = 0\right) = 1 \end{cases},$$

for p = 2, ..., P, where $\underline{w} = (w_1, ..., w_K)$ and $\Sigma_{k=1}^{K} w_k = 1$. Note that we introduce a Bernoulli random variable r_p to decide whether $\underline{Z}_p = \underline{0}$. When $r_p = 1$, $\underline{Z}_p$ is assumed to follow a multinomial distribution with parameters 1 and $\underline{w}$; when $r_p = 0$, $\underline{Z}_p$ is set as zero. Then, we assume an exchangeable prior for r_p with parameter τ following a Beta distribution:

(14) $$\pi(r_p) = \text{Bernoulli}(\tau), \text{ and}$$

(15) $$\pi(\tau) = \text{Beta}(t_1, t_2),$$

with t_1 and t_2 as hyperparameters. Finally, we assume that $\underline{w}$ follows a Dirichlet distribution:

(16) $$\pi(w_1, ..., w_K) = \text{Dirichlet } (\theta_1, ..., \theta_K).$$

Web Appendix I-D shows the derivation of the associated full conditional distributions, and Web Appendix II-D presents the steps of the proposed MCMC estimation for this distinctive factors model. Given that all the full conditional distributions are standard probability distributions, a Gibbs sampler can be used to generate a random sample from the joint posterior distribution.

Additional Considerations

Label switching. One issue to consider in such Bayesian finite mixture models is that of label switching (see Jasra, Holmes, and Stephens 2005; Sperrin, Jaki, and Wit 2010), whereby multiple solutions can provide the same function value in a nonidentified model. In our implementation, we first simulate from the nonconstrained posterior distribution and then impose identifiability constraints on the generated MCMC sample. After we have simulated an MCMC sample from a nonconstrained posterior distribution, we can impose any ordering constraint on this sample after the simulations have been completed for estimation purposes (see Stephens 1997). This postprocessing approach alleviates any concern of an adverse effect on simulation by imposing a constraint on the support of the posterior because the simulations are performed from the nonconstrained posterior distribution. After the simulation is completed, the $\underline{\beta}_k$ are relabeled for each MCMC iteration according to a constraint—for example, $\beta_{1p} < \beta_{2p} < ... < \beta_{Kp}$ for a given component p. Then, the other associated parameters $\underline{Z}$, $\underline{H}$, and $\underline{d}$ (excluding $\underline{Z}$ for the common factors model and including $\underline{r}$ for the dimension-constrained model) are reordered accordingly to match that of $\underline{\beta}_k$.

Model selection. Unless a user has an a priori set of managerial constraints to impose on the model (in which case, he or she needs to run only that model), assistance is needed in deciding which of the four specifications, as well as the number of segments, are the most appropriate given the data. This requires model selection heuristics. To compare models M_l and M_{l*}, we recommend the use of the Bayes factor (the ratio of the two marginal likelihoods). To find the marginal likelihood $\pi(y|M_k)$, k = 1, ..., K, we use the basic marginal likelihood identity suggested by Chib (1995) in logarithmic form for computational efficiency:

(17) $$\ln[\pi(\underline{Y}|M_k)] = \ln[f(\underline{Y}|M_k, \theta^*)] + \ln[\pi(\theta^*|M_k)] - \ln[\pi(\theta^*|M_k, \underline{Y})],$$

where $\theta^* = \{\underline{H}^*, \underline{\beta}^*, \sigma^{2*}\}$ are set equal to the posterior modes. In this case, $\ln[f(\underline{Y}|M_k, \theta^*)]$ and $\ln[\pi(\theta^*|M_k)]$ can be computed analytically, and for $\ln[\pi(\theta^*|M_k, \underline{Y})]$, we employ a data augmentation scheme to estimate this quantity. Given that the logarithm of the Bayes factor is equal to the difference of the log marginal likelihoods (LMLs), we can similarly employ the LML as a model selection heuristic and select the model with the greatest value. To identify significant variables in the four Bayesian models, we adopted the ratio of the posterior odds and prior odds (i.e., the odds ratio with a cutoff value of 20; see Jeffreys 1961) because we assumed Bernoulli distributions. Web Appendix III provides the computational details of the odds ratios and variable selection procedures for the proposed models.

Robustness of the procedures. Web Appendix IV provides the details of a Monte Carlo analysis designed to examine the performance and robustness of the newly proposed constrained Bayesian finite mixture regression models versus latent class regression (FlexMix; Grün and Leisch 2008) and Bayesian latent class regression (RegmixMH; Benaglia et al. 2009) on a variety of performance measures. Namely, we manipulated four independent factors in a full factorial design: (1) the underlying model generating the data (e.g., pattern of heterogeneity and active variables), (2) the level of collinearity among the independent variables in the model, (3) the number of segments, and (4) the sample size. The performance or dependent variables examined in this study were the recovery of the correct pattern of heterogeneity, segment memberships, and coefficients. We found that the proposed models accurately capture the pattern of heterogeneity under the appropriate data structure. All four constrained Bayesian models are more robust against multicollinearity than FlexMix and RegmixMH (for details, see Web Appendix IV).

ANALYSIS OF A SERVPERF SURVEY FOR NATIONAL INSURANCE COMPANY

In exchange for money, time, and effort, service customers expect value from a company's goods, labor, professional skills, facilities, networks, and systems; however, they do not typically take ownership of any of the physical elements involved (Lovelock 2005). The extensive economic importance of the service industry has led to an increased interest in problems regarding service marketing (Kotler and Keller 2012), and the resulting research has produced several major findings, especially in the area of service quality assessment (Rust and Huang 2012). Marketing strategists have found that firms with comparatively higher levels of quality typically reap higher market share and return (Philips, Chang, and Buzzell 1983) and obtain lower costs and higher profit margins (Crosby 1984). Yet the positive relationship between service quality and shareholder value can be masked when customers have heterogeneous perceptions of service quality (Grewal, Chandrashekaran, and Citrin 2010). It is thus important to measure service quality and derive its key drivers accurately while taking heterogeneity into account. Furthermore, consumer evaluations such as those obtained from SERVQUAL/SERVPERF studies have a strong impact on future purchase intentions, and yet we still question how marketing managers can identify which subsets of the 22 items are most important. Are managers able to use the results in their own managerial context? A focus on actionability has not been present in the realm of service quality models.

To illustrate how the constrained proposed procedures can help identify response patterns and heterogeneity in service quality evaluations, we use the SERVPERF mail survey data from Parasuraman, Grewal, and Krishnan (1991), collected from a random sample mailing to 1,000 policyholders of National Insurance Company. In the survey, participants first answered questions regarding the five service quality dimensions of SERVPERF from the 22 questionnaire items (independent variables) shown in the Appendix. Second, they answered an overall service quality question (dependent variable), "How would you rate the overall service quality of National and its employees?" on a ten-point scale ranging from "extremely poor" (1) to "extremely good" (10). Third, participants answered general questions about their relationship with the company and provided demographic information. We used the 191 returned surveys for our analyses.

We first examined the correlations (not provided here) between the 22 SERVPERF items and the dependent variable. Of the 253 pairwise correlations, 247 (98%) are significant at $p < .05$ and all of the 22 independent variables are highly positively correlated with the dependent variable. The large number of positive correlations among the independent variables suggests the presence of severe multicollinearity, which is typical for such service quality surveys.

Results from Traditional Approaches

Table 1 presents the estimated coefficients from four traditional procedures (aggregate regression, stepwise regression, FlexMix, and RegmixMH) for the 191 respondents. We discuss each of the traditional procedures in detail in the following subsections.

Aggregate and stepwise regressions. The aggregate regression solution is strongly affected by multicollinearity, with a maximum condition index of 71.4. In addition to the intercept, four items are significant: Reliability 3, Empathy 1, and Assurance 2 have a significant positive effect on overall service quality. It is somewhat unsettling to observe that Reliability 5 has a significant negative coefficient, which implies that if the company were to maintain records with more errors, customers' perceived overall service quality would increase.

Stepwise regression significantly reduces the problem of multicollinearity in the data (maximum condition index = 19.62). In addition to variables selected with a positive coefficient by the aggregate regression, Empathy 2, Reliability 1, and Responsiveness 1 also have significant and positive effects. Yet does this aggregate variable selection mask segment-level differences?

Results from FlexMix. We ran FlexMix, a traditional latent class regression approach, for $K = 1 \ldots 5$ segments. Given the incidence of many local optima, we ran the procedure multiple times for each value of K and selected the best solution for each K. As Table 2 shows, the recommendations from the information heuristics are mixed. Whereas the consistent Akaike information criterion (CAIC) suggests the presence of no heterogeneity ($K = 1$), the other information criteria (Akaike information criterion [AIC], Bayesian information criterion [BIC], and the Akaike information criterion with 3 as the penalty weight [AIC3]) suggest that a larger number of segments ($K > 5$) is more appropriate. As such, we are not able to use information criteria reliably to select the best number of segments for FlexMix. Given that the proposed Bayesian procedures identify a solution with $K = 2$ segments (discussed subsequently), we present the FlexMix $K = 2$ segment solution along with the FlexMix $K = 5$ segment solution suggested by AIC, BIC, and AIC3. The FlexMix results presented in Table 1 represent the best local optima we could obtain.

We first turn to the $K = 2$ solution for FlexMix. Of the 22 SERVPERF items, 8 have significant effects on overall service quality for members of Segment 1 (representing 67% of the sample). We note that three of the eight independent variables have significant negative coefficients that are difficult to explain properly to any manager in this service quality context. For example, Reliability 4 ("National provides its services at the time it promises to do so") seems to have a negative effect and would *decrease* perceptions of service quality if improved. Members of Segment 2 (representing 36% of the sample) possess 12 significant variables. Just as with Segment 1, the four resulting negative coefficients lack face validity. With respect to the $K = 5$ solution by FlexMix, an excessive number of variables are selected, many with negative coefficients. Namely, of 104 selected variables, 39 variables have negative coefficients. All coefficients can be reasonably expected to be positive in this particular application.

Results from traditional Bayesian mixture regression (RegmixMH). In the two rightmost columns in Table 1, we present the $K = 2$ solution from RegmixMH, a Bayesian latent class regression procedure without variable selection. Considering the unbalanced mixture proportions (i.e., 2%, 98%) and three negative coefficients, the solution is also of questionable value.

Table 1
RESULTS FROM TRADITIONAL METHODS

Independent Variables	*Ordinary Least Squares Regression*	*Stepwise Regression*	*FlexMix (K = 2)*		*FlexMix (K = 5)*					*RegmixMH*	
			k = 1	*k = 2*	*k = 1*	*k = 2*	*k = 3*	*k = 4*	*k = 5*	*k = 1*	*k = 2*
Intercept	**–.93**	**–1.05**	**–1.58**	**1.61**	.29	**–1.40**	**1.66**	**–2.28**	**–1.29**	.62	**2.00**
Reliability 1	.14	**.21**	**.49**	**–.28**	**.27**	**–.04**	**–.10**	**–.52**	**.50**	**2.44**	**1.75**
Reliability 2	.16	.00	**.40**	.04	**.20**	**.29**	**.11**	**.51**	**–.15**	**2.00**	**2.45**
Reliability 3	**.17**	**.25**	–.1	**.30**	**.42**	**.13**	**.48**	**.28**	**–.08**	.40	**–.62**
Reliability 4	.11	.00	**–.31**	**.47**	.05	**.07**	**.28**	**1.01**	**–.29**	–.12	**.33**
Reliability 5	**–.10**	.00	–.1	**–.26**	**–.15**	–.12	**.07**	–.08	–.07	**.76**	.01
Empathy 1	**.23**	**.32**	**.47**	**.21**	**.40**	.13	**.24**	**.20**	**.16**	**2.12**	**.22**
Empathy 2	.08	**.1**	–.02	.06	**–.24**	**.13**	**.11**	**.13**	**.01**	.34	–.06
Empathy 3	–.03	.00	.00	–.07	**.26**	**–.25**	**.07**	**–.36**	**.08**	–.09	**–.16**
Empathy 4	.06	.00	–.044	**.21**	**–.22**	**.14**	**.15**	**.05**	**.15**	**2.30**	.17
Empathy 5	.08	.00	.17	.03	**.29**	**.22**	**–.44**	**.37**	.00	**2.17**	**.13**
Tangibles 1	–.01	.00	.03	**.43**	**.11**	**–.44**	**.70**	**–.29**	**.19**	**3.82**	.01
Tangibles 2	–.12	.00	–.08	**–.83**	**.15**	**.26**	**–.93**	**–.08**	**–.40**	**.93**	.00
Tangibles 3	.06	.00	.06	–.11	**–.09**	**.20**	**–.21**	**.27**	**–.08**	**.66**	**.19**
Tangibles 4	.12	.00	.15	.2	**–.24**	**.43**	.00	**.52**	**.36**	**1.56**	**–1.07**
Responsiveness 1	.07	.00	**.47**	**–.45**	**–.56**	**.19**	**–.27**	.00	**.28**	.31	**.63**
Responsiveness 2	–.09	.00	**–.31**	**.26**	**–.21**	**.01**	**.59**	**–.70**	**.07**	–1.16	**.15**
Responsiveness 3	.03	.00	**–.40**	.2	**.17**	**–.20**	**–.05**	**–.38**	**–.09**	**.90**	**.43**
Responsiveness 4	.07	**.20**	.13	–.05	**.07**	.01	**.24**	**.41**	**.09**	–.90	–.07
Assurance 1	.1	.00	.00	**.57**	.00	**.24**	**–.13**	**.23**	**–.15**	**1.19**	.02
Assurance 2	**.42**	**.53**	**.51**	.07	**.51**	**.41**	**.23**	**–.12**	**.96**	**2.00**	–.07
Assurance 3	–.12	.00	.12	–.2	**–.41**	**–.44**	**–.19**	**.07**	**.20**	**1.75**	.01
Assurance 4	.17	.00	–.03	**.35**	**.46**	**.39**	**.37**	**.23**	**–.05**	**2.45**	**.19**
Mixture	N.A.	N.A.	67%	36%	24%	17%	21%	19%	20%	2%	98%

Notes: For ordinary least squares regression, stepwise regression, and FlexMix, boldfaced cells denote significance at $p < .05$; for the RegmixMH, boldfaced cells do not include zero in a 95% highest posterior density credible set. N.A. = not applicable.

Table 2

MODEL SELECTION HEURISTICS FOR FLEXMIX

	K = 1	K = 2	K = 3	K = 4	K = 5
AIC	489.28	433.33	333.38	225.74	**133.97**
AIC3	513.28	482.33	407.38	324.74	**257.97**
BIC	567.33	592.7	574.05	547.72	**537.25**
CAIC	**591.33**	641.7	648.05	646.72	661.25

Notes: Boldfaced cells denote the solution that the criterion selects.

Results from the Proposed Bayesian Finite Mixture Methods

For each of the four Bayesian models proposed, we performed 10,000 iterations and discarded 5,000 as a burn-in period. Here, to identify significant variables in the four Bayesian models, we also adopted the odds ratio with a cutoff value of 20 (see Jeffreys 1961) because we assumed Bernoulli distributions. Inspection of the trace plots also suggests that convergence was obtained before the start of the iterations that were kept (see Figure V-1 in Web Appendix V). Given that no a priori constraints were known, we use model selection information (i.e., LML; see Table 3) to select the best solution for each of the four proposed models as well as the overall best solution across all models. The LML numbers suggest that the distinctive factors model with K = 2 segments (LML: –145.6) is the best solution across all models and numbers of segments.

Table 4 presents the K = 2 solutions that the four proposed Bayesian models obtained alongside the K = 1 solutions for the common factors and dimension-constrained models (their best solutions, as identified in Table 3). We note that the common factors K = 2 solution shows unbalanced mixture proportions (3% and 97%) and that the dimension-constrained model selected none of the variables in the K = 2 solution. We obtained somewhat similar results to the distinctive factors solution for the unconstrained model with K = 2. Both models represent a data structure of distinctive items across segments, but the distinctive factors model provides a better fit and more intuitive results, as we describe subsequently.

For the K = 2 solution from the distinctive factors model, Segment 1 is the smallest group (14%): a segment of less-satisfied users with the potential for higher loyalty with improved retention strategies. Specifically, we find that members of this segment have the lowest average overall service quality score (M = 6.70 vs. 7.98; t(198) = 2.60, p = .01) and are less likely to recommend the company to a friend (70.4% vs. 84.5%; $\chi^2(1) = 3.74$, $p = .05$). The finding that a higher proportion of Segment 1 members had experienced more problems than those of Segment 2 (51.9% vs. 29.9%; $\chi^2(1) = 5.06$, $p = .02$) may explain the lower service quality score and recommendation response. The segment also includes fewer loyal customers, with only 51.9% who have been with the company for five years or more (vs. 73.8% in Segment 2; $\chi^2(1) = 5.38$, $p < .01$). In addition, we regressed the overall service quality score and the consumers' willingness to recommend National on their tenure with the company, an indicator of segment membership, and their interaction. We found that members of Segment 1 who have been with National for a longer time (i.e., more than five years) are more likely than those of Segment 2 to have higher perceptions of service quality and a willingness to recommend National (both $ps < .05$).

Table 3

LML RESULTS FOR THE MODEL SELECTION OF THE SERVPERF APPLICATION

LML	K = 1	K = 2	K = 3	K = 4	K = 5
Unconstrained model	–202.8	*–179.4*	–307.0	–401.6	–588.3
Common factors model	*–244.3*	–315.2	–342.2	–389.7	–401.3
Dimension-constrained model	*–282.9*	–312.5	–411.8	–453.3	–487.7
Distinctive factors model	–226.8	***–145.6***	–263.0	–357.9	–600.1

Notes: Italicized cells represent the best model for each model. The boldfaced cell indicates the best overall model.

Regarding the managerial implications for National, Segment 1 customers require intensive care during the early stages of their relationship with the company; National must focus on the key drivers to convert people in this segment into long-term, loyal customers. In terms of key drivers, the insurance company should focus on Empathy 5 ("Employees of National understand your specific needs"; $\beta_{1,11} = .56$) and Reliability 2 ("When you have a problem, National shows a sincere interest in solving it"; $\beta_{1,3} = .33$) for this segment. Post hoc analyses also reveal that members of this segment rate the company worse than those of Segment 2 on both understanding their needs ($M_1 = 4.81$, $M_2 = 5.49$; t(189) = 2.02, $p = .05$) and having a sincere interest in solving their problem ($M_1 = 5.00$, $M_2 = 5.63$; t(189) = 1.91, p = .06). Although the company scored lower on a host of other SERVPERF indicators, it can increase these consumers' service experience by ensuring that they feel that the company understands their needs and works diligently to solve their problems. Members of Segment 2 represent the majority of highly satisfied National customers, with an average service quality score of 7.98. The key drivers for Segment 2 involve Reliability 1 ("When National promises to do something, it does so"; $\beta_{2,2} = .27$), Empathy 1 ("National treats you with care"; $\beta_{2,7} = .22$), and Assurance 2 ("You feel safe in your transactions with National"; $\beta_{2,21} = .52$).

We also point out that the solution offered by the unconstrained model, which has poorer fit, provided different findings. Namely, for Segment 1 (13% of the sample), only Empathy 5 ($\beta_{1,11} = .68$) is the main driver of overall service satisfaction. For Segment 2 (87% of the sample), the model identifies five key drivers, three of which are related to reliability. The significant drivers are Reliability 1 ($\beta_{2,2} = .24$), Reliability 3 ($\beta_{2,4} = .20$), Reliability 5 ($\beta_{2,6} = -.10$), Empathy 1 ($\beta_{2,7} = .24$), and Assurance 2 ($\beta_{2,21} = .58$). The two derived segments from this solution are somewhat similar to those of the distinctive factor model solution that we selected, with an adjusted Rand index (Hubert and Arabie 1985) of .70 and a 94% match. For Segment 1 of both solutions, understanding customer needs (Empathy 5) is important. Yet the distinctive factor model identifies a second driver of satisfaction for this lower satisfaction group, regarding the company's having a sincere interest in fixing their problem. We would expect this to be important given that members of Segment 1 have experienced more problems than those of Segment 2 (51.9% vs. 29.9%; $\chi^2(1)$ =

Table 4
SOLUTIONS FOR THE PROPOSED BAYESIAN MODELS FOR THE SERVPERF APPLICATION

	Unconstrained		*Common Factors (K = 1)*	*Common Factors (K = 2)*		*Dimension Constrained (K = 1)*	*Dimension Constrained (K = 2)*		*Distinctive Factors*[a]	
	k = 1	*k = 2*	*k = 1*	*k = 1*	*k = 2*	*k = 1*	*k = 1*	*k = 2*	*k = 1*	*k = 2*
Intercept	−.09	.01	−.17	−.04	.00	**5.03**	**2.94**	**8.87**	−.25	−.03
Reliability 1	.05	**.24**	.00	.02	.16	.01	.01	.00	.00	**.27**
Reliability 2	.09	.01	.15	.01	.08	.02	.04	.00	**.33**	.00
Reliability 3	.04	**.20**	**.29**	**.15**	**.21**	**.11**	.03	.00	.00	.00
Reliability 4	.00	.01	.00	.00	.00	.02	.05	.00	.00	.00
Reliability 5	−.01	**−.10**	−.01	−.04	−.03	.00	.01	.00	.00	.00
Empathy 1	.05	**.24**	**.33**	**.74**	**.43**	.03	.02	.00	.00	**.22**
Empathy 2	.00	.01	.00	.00	.00	.00	.00	.00	.00	.00
Empathy 3	.00	.00	.00	.00	.00.	.00	.00	.00	.00	.00
Empathy 4	.00	.00	.00	.00	.00	.00	.02	.00	.00	.00
Empathy 5	**.68**	.13	.00	.00	.00	.04	.02	.00	**.56**	.00
Tangibles 1	.02	.02	.00	.00	.00	.00	.00	.00	.00	.00
Tangibles 2	−.06	−.04	.00	.00	.00	.00	.00	.00	.00	.00
Tangibles 3	.00	.00	.00	.00	.00	.09	.01	.00	.00	.00
Tangibles 4	.00	.00	.00	.00	.00	.01	.00	.00	.00	.00
Responsiveness 1	.1	.04	.00	.00	.00	.00	.01	.00	.00	.00
Responsiveness 2	−.01	.00	.00	.00	.00	.00	.00	.00	.00	.00
Responsiveness 3	.07	.02	.00	.00	.00	.01	.04	.00	.00	.00
Responsiveness 4	.23	.03	.08	.00	.00	.04	.02	.00	.00	.00
Assurance 1	.07	.01	.00	.00	.00	.00	.02	.00	.00	.00
Assurance 2	.08	**.58**	**.60**	**.33**	**.38**	.06	.05	.00	.00	**.52**
Assurance 3	.01	.00	.00	.00	.00	.00	.01	.00	.00	.00
Assurance 4	.00	.01	.00	.00	.00	.01	.01	.00	.00	.00
Mixture	13%	87%	N.A.	3%	97%	N.A.	24%	76%	14%	86%

[a]Represents the selected solution.
Notes: Boldfaced entries indicate odds ratio >20. N.A. = not applicable.

5.06, $p = .02$). For members of Segment 2, Empathy 1, Reliability 1, and Assurance 2 are significant in both solutions. The unconstrained model also identifies Reliability 3 and 5 as important. The coefficient for Reliability 5 is especially problematic because it is negative, which implies that adding errors to National records would improve customer satisfaction. In summary, although both solutions offer drivers that are distinctive across segments, the solution from the distinctive factors model provides a better LML value and also has greater face validity.

DISCUSSION

In this article, we extend the Kim, Fong, and DeSarbo (2012) unconstrained model for segment-level variable selection to be more actionable by managers. We do so by proposing three Bayesian constrained models: the common factors model (characterized by common active predictors across segments, although the coefficients may differ), the distinctive factors model (in which segments must have different active predictors), and the dimension-constrained model (in which, at most, one predictor can be active per segment among each a priori defined latent dimension). The added constraints augment the unconstrained model by incorporating possible managerial restrictions that would otherwise render the derived solutions difficult and/or impractical to implement. These models can be especially helpful when there is high potential for instability due to small sample sizes, multicollinearity, and/or many independent variables, as is typical in service quality studies.[1]

Managers can use the four Bayesian models in two ways. When a manager has a priori constraints that need to be reflected in the final solution, he or she can focus on the corresponding constrained model and use model selection heuristics to determine the number of segments and the heterogeneous response patterns present in the data simultaneously. When a manager does not have a priori constraints that need to be considered, he or she can use model selection heuristics to choose not only the number of latent segments and the heterogeneous response patterns in the data but also which of the four models is most appropriate. Through our empirical analysis of a SERVPREF survey, we demonstrate that using the constrained models in conjunction with model selection heuristics can provide solutions with both a better fit and an interpretation with greater face validity than traditional analyses and an unconstrained model.

For future studies, researchers could extend our constrained approaches to constrained variable selection normative segmentation models. For example, stochastic variable selection can be constrained, reflecting a desire to minimize implementation costs or maximize profits. Another area for further research lies in applying these proposed Bayesian models into discrete models such as binary or multinomial choice models (e.g., for conjoint analyses). Moreover, it would be worthwhile to incorporate both latent structures and variable selection to situations in which there are multiple correlated dependent variables. Such a situation could occur, for example, if an investigator wanted to examine factors that influence not only service quality but also satisfaction with the service representatives or price–quality inferences. In addition, we could generalize this approach to partial least squares regression or structural equations modeling to take measurement error directly into account; alternatively, we could use Bayesian decision theory to incorporate managerial constraints into the loss function (see Gilbride, Lenk, and Brazell 2008). Finally, these models can also be used in customer satisfaction studies, which would extend their impact and usability.

[1]The R codes for all four proposed Bayesian models are available on the first author's website. Contact the first author for additional details.

APPENDIX: NATIONAL INSURANCE COMPANY'S SERVQUAL/SERVPERF DIMENSIONS AND ITEMS

All questions are seven-point Likert-type items measured from "strongly disagree"(1) to "strongly agree" (7).

Reliability

1. When National promises to do something, it does so.
2. When you have a problem, National shows a sincere interest in solving it.
3. National performs the service right the first time.
4. National provides its services at the time it promises to do so.
5. National maintains error-free records.

Empathy

1. National treats you with care.
2. National has operating hours convenient to all its policyholders.
3. National has employees who give you personal attention.
4. National has your best interest in mind.
5. Employees of National understand your specific needs.

Tangibles

1. National has modern-looking equipment.
2. National's physical facilities are visually appealing.
3. National's employees are neat appearing.
4. Materials associated with service (such as pamphlets or statements) are visually appealing at National.

Responsiveness

1. Employees of National tell you exactly when services will be performed.
2. Employees of National give you prompt service.
3. Employees of National are always willing to help you.
4. Employees of National are never too busy to respond to your requests.

Assurance

1. The behavior of employees of National instills confidence in you.
2. You feel safe in your transactions with National.
3. Employees of National are consistently courteous with you.
4. Employees of National have the knowledge to answer your questions.

REFERENCES

Babakus, Emin and Gregory W. Boller (1992), "An Empirical Assessment of the SERVQUAL Scale," *Journal of Business Research*, 24 (3), 253–68.

Benaglia, Tatiana, Didier Chauveau, David R. Hunter, and Derek S. Young (2009), "mixtools: An R Package for Analyzing Finite Mixture Models," *Journal of Statistical Software*, 32 (6), 1–29.

Brady, Michael K. and Joseph J. Cronin (2001), "Some New Thoughts on Conceptualizing Perceived Service Quality: A Hierarchical Approach," *Journal of Marketing*, 65 (July), 34–49.

Brown, Tom J., Gilbert A. Churchill Jr., and Paul J. Peter (1993), "Improving the Measurement of Service Quality," *Journal of Retailing*, 69 (1), 127–39.

Brusco, Michael J., Dennis J. Cradit, and Stephanie Stahl (2002), "A Simulated Annealing Heuristic for a Bicriterion Partitioning Problem in Market Segmentation," *Journal of Marketing Research*, 39 (February), 99–109.

Chandukala, Sandeep R., Jeffrey P. Dotson, Jeff D. Brazell, and Greg M. Allenby (2011), "Bayesian Analysis of Hierarchical Effects," *Marketing Science*, 30 (1), 123–33.

——, Yancy D. Edwards, and Greg M. Allenby (2011), "Identifying Unmet Demand," *Marketing Science*, 30 (1), 61–73.

Chib, Siddhartha (1995), "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, 90 (432), 1313–21.

Cronin, Joseph J., Jr., and Steven A. Taylor (1992), "Measuring Service Quality: A Reexamination and Extension," *Journal of Marketing*, 56 (July), 55–68.

Crosby, Phillip B. (1984), *Quality Without Tears: The Art of Hassle-Free Management*. New York: McGraw-Hill.

DeSarbo, Wayne S. and Douglas Grisaffe (1998), "Combinatorial Optimization Approaches for Constrained Market Segmentation: An Application to Industrial Market Segmentation," *Marketing Letters*, 9 (2), 115–34.

Drolet, Aimee L. and Donald G. Morrison (2001), "Do We Really Need Multiple-Item Measures in Service Research?" *Journal of Service Research*, 3 (3), 196–204.

Gilbride, Timothy J., Greg M. Allenby, and Jeff D. Brazell (2006), "Models for Heterogeneous Variable Selection," *Journal of Marketing Research*, 43 (August), 420–30.

——, Peter J. Lenk, and Jeff D. Brazell (2008), "Market Share Constraints and the Loss Function in Choice-Based Conjoint Analysis," *Marketing Science*, 27 (6), 995–1011.

Gordon, A.D. (1996), "A Survey of Constrained Classification," *Computational Statistics and Data Analysis*, 21 (1), 17–29.

Grewal, Rajdeep, Murali Chandrashekaran, and Alka V. Citrin (2010), "Customer Satisfaction Heterogeneity and Shareholder Value," *Journal of Marketing Research*, 47 (August), 612–26.

Grün, Bettina and Friedrich Leisch (2008), "FlexMix Version 2: Finite Mixtures with Concomitant Variables and Varying and Constant Parameters," *Journal of Statistical Software*, 28 (4), 1–35.

Hubert, Lawrence and Phipps Arabie (1985), "Comparing Partitions," *Journal of Classification*, 2 (1), 193–218.

Jain, Sanjay K. and Garima Gupta (2004), "Measuring Service Quality: SERVQUAL vs. SERVPERF Scales," *Vikalpa: The Journal for Decision Makers*, 29 (2), 25–37.

Jasra, A., C.C. Holmes, and D.A. Stephens (2005), "Markov Chain Monte Carlo Methods and the Label Switching Problem in Bayesian Mixture Modeling," *Statistical Science*, 20 (1), 50–67.

Jeffreys, Harold (1961), *Theory of Probability*. Oxford, UK: Clarendon Press.

Kim, Sunghoon, Duncan K.H. Fong, and Wayne S. DeSarbo (2012), "Model-Based Segmentation Featuring Simultaneous Segment-Level Variable Selection," *Journal of Marketing Research*, 49 (October), 725–36.

Kotler, Phillip (1999), *Marketing Management: Analysis, Planning, Implementation, and Control*, 9th ed. Upper Saddle River, NJ: Prentice Hall.

—— and Kevin L. Keller (2012), *Marketing Management*, 14th ed. Upper Saddle River, NJ: Prentice Hall.

Ladhari, Riadh (2009), "A Review of Twenty Years of SERVQUAL Research," *International Journal of Quality and Service Sciences*, 1 (2), 172–98.

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

Lovelock, Christopher (2005), "Service Portraits in Service Research: A Critical Review," *International Journal of Service Industry Management*, 16 (1), 107–121.

Mahajan, Vijay and Arun K. Jain (1978), "An Approach to Normative Segmentation," *Journal of Marketing Research*, 15 (August), 338–45.

Montoya-Weiss, Mitzi and Roger J. Calantone (1999), "Development and Implementation of a Segment Selection Procedure for Industrial Markets," *Marketing Science*, 18 (3), 373–95.

Parasuraman, A., Leonard L. Berry, and Valarie A. Zeithaml (1991), "Refinement and Reassessment of the SERVQUAL Scale," *Journal of Retailing*, 67 (4), 420–50.

——, Dhruv Grewal, and R. Krishnan (1991), *Marketing Research*. Boston: Houghton Mifflin.

Philips, Lynn W., Dae R. Chang, and Robert D. Buzzell (1983), "Product Quality, Cost Position, and Business Performance: A Test of Some Key Hypotheses," *Journal of Marketing*, 47 (Spring), 26–43.

Rust, Roland T. and Ming-Hui Huang (2012), *Service Marketing: Insights and Directions*. Cambridge, MA: Marketing Science Institute.

Sperrin, Matthew, Thomas Jaki, and Ernst Wit (2010), "Probabilistic Relabelling Strategies for the Label Switching in Bayesian Mixture Models," *Statistics and Computing*, 20 (3), 357–66.

Sriram, N. (1990), "Clique Optimization: A Method to Construct Parsimonious Ultrametric Trees from Similarity Data," *Journal of Classification*, 7 (1), 33–52.

Stephens, M. (1997), "Bayesian Methods for Mixtures of Normal Distributions," doctoral dissertation, University of Oxford.

Wind, Yoram (1978), "Issues and Advances in Segmentation Research" *Journal of Marketing Research*, 15 (August), 317–37.

Winter, Frederick W. (1979), "A Cost-Benefit Approach to Market Segmentation," *Journal of Marketing*, 43 (October), 103–111.

Zeithaml, Valarie A., A. Parasuraman, and Leonard L. Berry (1990), *Delivering Quality Service: Balancing Customer Perceptions and Expectations*. New York: Simon & Schuster.