Department of Social Systems and Management

Discussion Paper Series

No.1258

A Bayesian simultaneous demand and supply model for aggregate data in a differentiated product market

by

Yutaka Yonetani, Yuichiro Kanazawa, Satoshi Myojo
and Stephen John Turnbull

April 2010

UNIVERSITY OF TSUKUBA
Tsukuba, Ibaraki 305-8573
JAPAN

A Bayesian simultaneous demand and supply model for aggregate data in a differentiated product market

Yutaka Yonetani, Yuichiro Kanazawa, Satoshi Myojo
and Stephen John Turnbull[1]

[1]Yutaka Yonetani is a research fellow in National Institute of Science and Technology Policy (NISTEP), Ministry of Education, Culture, Sports, Science and Technology, Chuo-Godo-Chosha 7 East 16th floor, 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo 100-0013, Japan. His e-mail address is yonetani@nistep.go.jp.

Yuichiro Kanazawa is Professor in Statistics in Department of Social Systems and Management, Graduate School of Systems and Information Engineerings, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan. His e-mail address is kanazawa@sk.tsukuba.ac.jp.

Satoshi Myojo is Associate Professor in Faculty of Economics, Graduate School of Economics, Kobe University, 2-1 Rokkodai-cho, Nada-ku, Kobe, Hyogo 657-8501. His e-mail address is myojo@econ.kobe-u.ac.jp.

Stephen John Turnbull is Associate Professor in Economics in Department of Social Systems and Management, Graduate School of Systems and Information Engineerings, University of Tsukuba. His e-mail address is turnbull@sk.tsukuba.ac.jp.

Abstract

In this paper, we propose a Bayesian simultaneous demand and supply model for aggregate data in a differentiated product market. The proposed method treats price endogeneity and consumer heterogeneity as well as requires only aggregate data. In the Bayesian estimation, we use an MCMC algorithm including the data augmentation, Gibbs sampler and Metropolis-Hastings algorithm. Our likelihood for the demand and supply model is directly derived from the endogenous sales volume and price unlike a past similar framework. To show validity of our proposed method, we perform an analysis of simulated data, and apply our method to data from the U.S. automobile market.

1 Introduction

In making inferences about consumer preferences, we would like to make use of data on choices by individuals, along with their heterogeneous personal characteristics (disaggregate data). However, often we must perform such inferences under the handicap that personal characteristics may not be linked to choice data. That is, we only have information about the marginal distributions of characteristics or choices (aggregate data). Models of this type were pioneered by Berry, Levinsohn and Pakes (1995; henceforth BLP, 1995).

An important aspect of most applications that must be incorporated in the model is price endogeneity. It is known that ignoring price endogeneity leads to estimation bias (Berry, 1994; Villas-Boas and Winer, 1999). Therefore, modern analyses in marketing and economics often model firms' pricing behavior simultaneously with consumers' purchasing behavior.

In the literature, there are both frequentist and Bayesian research streams. The frequentists include Berry (1994), BLP (1995), Sudhir (2001), Petrin (2002), Berry, Levinsohn and Pakes (2004; henceforth BLP, 2004), Myojo (2007) and Myojo and Kanazawa (2010). Bayesian analysis was employed by Yang et al. (2003), Romeo (2007), Jiang et al. (2009) and Musalem et al. (2009). The Yang et al. (2003) Bayesian paper introduced both limited and full information models for disaggregate data. The Yang et al. (2003) method was extended to aggregate data by Jiang et al. (2009) and Musalem et al. (2009) in the limited information framework, and by Romeo (2007) in the full information case.

In this paper, we propose another Bayesian estimation method for the full information model with aggregate data. To show validity of our proposed method, we perform an analysis of simulated data, and apply our method to data from the U.S. automobile market.

The Bayesian framework offers us four advantages. First, in spite of the complexity of models used, we can construct an exact posterior distribution for the parameters, which does not depend on specific distributional assumptions for the random processes. Second, because the posterior distribution is exact, finite sample inference may be conducted without resorting to asymptotic methods of dubious applicability. Third, when appropriate, we can incorporate existing knowledge, both formal and heuristic, about the parameters in the Bayesian prior distribution. Fourth, the Bayesian framework allows more general distributions of utility and cost parameters. We can make more natural assumptions about these parameters.

Compared with the limited-information methods, the full-information methods can improve the estimation if their supply side pricing modelings are appropriate for markets we investigate. Additionally, the full-information methods avoid a controversial problem of choosing instrumental variables in the limited-information methods. Among the full-information methods, Romeo (2007) used a pseudo-likelihood of mean utility which is calculated by a contraction mapping using aggregate sales volume in each Markov Chain Monte Carlo (MCMC) iteration to follow the frequentists' Generalized Method of Moment (GMM). On the other hand, our method derives a likelihood directly from aggregate sales volume and thus is constructed in one unified computational framework of the MCMC without the contraction mapping.

This paper is organized as follows. In Section 2, we specify our simultaneous demand and supply model. Then we develop our Bayesian estimation method for the model, using the data augmentation (Tanner and Wong, 1987), Gibbs sampler (Geman and Geman, 1984) and Metropolis-Hastings algorithm (Hastings, 1970). In Sections 3 and 4, we implement the simulation and empirical studies respectively. Conclusions and discussions are presented in Section 5.

2 Simultaneous demand and supply model and its Bayesian estimation

We explain our model in subsections 2.1 and 2.2. We adopt the BLP (1995) model of a market for a differentiated indivisible good. We explain our Bayesian estimation in subsection 2.3.

2.1 Demand Model

Consumers buy one unit of the good, chosen from among J products indexed by $j = 1, \dots, J$, or may choose not to buy any of the products. We model consumers who do not purchase as substituting an outside good, with index $j = 0$. The demand model is derived from utility maximization by heterogeneous consumers. The observed characteristics of product j are given by a vector $\boldsymbol{x}_j = (x_{j1}, \dots, x_{jQ})$. Product j's price is denoted by p_j and unobserved product characteristics are summarized by ξ_j. For the outside good $j = 0$, we assume that $\boldsymbol{x}_0 = \mathbf{0}$, $p_0 = 0$ and $\xi_0 = 0$. Consumer i's income is denoted by y_i, and net income if product j is purchased is $y_i - p_j$. Consumer i's idiosyncratic preference for product j is also denoted by ε_{ij}. Then consumer i's utility function u_{ij} for product j is specified as

$$u_{ij} = u_{ij}(y_i, p_j, \boldsymbol{x}_j, \xi_j, \varepsilon_{ij}; \boldsymbol{\theta}_i) = \alpha_i \log(y_i - p_j) + \boldsymbol{x}_j \boldsymbol{\beta}_i + \xi_j + \varepsilon_{ij}, \tag{1}$$

where $\boldsymbol{\theta}_i = (\alpha_i, \boldsymbol{\beta}_i')'$ are respectively his/her marginal utility for $\log(y_i - p_j)$ and a $Q \times 1$ marginal utility vector for $\boldsymbol{x}_j$.

Researchers observe neither consumers' idiosyncratic preference nor the unobserved characteristics, so we specify a probabilistic model of discrete consumer choice. Let s_{ij} denote the probability that consumer i will be observed to choose product j. Assuming that ε_{ij} in (1) is independent of the other terms and is independently and identically Gumbel (type I extreme value) distributed across consumers and alternatives, we derive the logistic choice probability for s_{ij} as

$$s_{ij} = s_{ij}(y_i, \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}; \boldsymbol{\theta}_i) = \frac{\exp(\alpha_i \log(y_i - p_j) + \boldsymbol{x}_j \boldsymbol{\beta}_i + \xi_j)}{\sum_{k=0}^{J} \exp(\alpha_i \log(y_i - p_k) + \boldsymbol{x}_k \boldsymbol{\beta}_i + \xi_k)}, \tag{2}$$

where $\boldsymbol{p} = (p_1, \dots, p_J)'$, $\boldsymbol{X} = (\boldsymbol{x}_1', \dots, \boldsymbol{x}_J')'$ and $\boldsymbol{\xi} = (\xi_1, \dots, \xi_J)'$.

The market share function for product j is obtained by aggregating s_{ij} in (2) with respect to y_i and $\boldsymbol{\theta}_i$ over the population of consumers as

$$s_j^0 = s_j^0(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}) = \int \int s_{ij} f^0(y_i, \boldsymbol{\theta}_i) dy_i \boldsymbol{d\theta}_i \tag{3}$$

where $f^0(y_i, \boldsymbol{\theta}_i)$ is the joint population density of y_i and $\boldsymbol{\theta}_i$.

2.2 Supply Model

We assume that there are F firms in a multiproduct Bertrand oligopoly, where each firm produces an exclusive subset of the J products and knows the true market share function (3) for its own products. Each firm sets price for each of its products according to the pricing strategy that maximizes the total profit from its products.

We specify the profit for each firm indexed by $f = 1, \dots, F$ as

$$\Pi_f = \sum_{j \in f} M s_j^0(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi})(p_j - c_j)$$

where M is the potential market size and c_j denotes unit cost for product j. The first order condition to obtain the profit maximizing prices $\boldsymbol{p}^*$ is

$$\boldsymbol{p}^* = -\left\{ \left(\frac{\boldsymbol{\partial G}}{\boldsymbol{\partial p}} \right)' \right\}^{-1} \boldsymbol{s}^0 + \boldsymbol{c}, \tag{4}$$

assuming the inverse above exists, where $\boldsymbol{s}^0 = (s_1^0, \dots, s_J^0)'$ and $\boldsymbol{c} = (c_1, \dots, c_J)'$; $(\boldsymbol{\partial G}/\boldsymbol{\partial p}) = (\boldsymbol{\partial s}/\boldsymbol{\partial p}) * \boldsymbol{\delta}$ with their (j, k) elements being $\partial G_j/\partial p_k$, $\partial s_j/\partial p_k$ and δ_{jk} respectively; and $\delta_{jk} = 1$ if prices for products j and k are set by the same firm and $\delta_{jk} = 0$ otherwise.

We assume that marginal cost is constant for each product. The cost shifters of product j are given by a vector $\boldsymbol{z}_j = (z_{j1}, \dots, z_{jS})$ and unobserved cost characteristics are summarized by η_j. We specify the unit cost vector $\boldsymbol{c}$ in (4) as

$$\log [\boldsymbol{c}] = \boldsymbol{Z\gamma} + \boldsymbol{\eta} \tag{5}$$

where $\boldsymbol{Z} = (\boldsymbol{z}_1', \dots, \boldsymbol{z}_J')'$ and $\boldsymbol{\eta} = (\eta_1, \dots, \eta_J)'$, and $\boldsymbol{\gamma}$ is a $(S \times 1)$ coefficient vector. This specification enforces positive cost. Then substituting $\boldsymbol{p}^* + \{(\boldsymbol{\partial G}/\boldsymbol{\partial p})'\}^{-1}\boldsymbol{s}^0$ from (4) for $\boldsymbol{c}$ in (5), we obtain a pricing equation as

$$\log \left[\boldsymbol{p}^* + \left\{ \left(\frac{\boldsymbol{\partial G}}{\boldsymbol{\partial p}} \right)' \right\}^{-1} \boldsymbol{s}^0 \right] = \boldsymbol{Z\gamma} + \boldsymbol{\eta}, \tag{6}$$

where we write $\boldsymbol{p}^* = \boldsymbol{p}^*(\boldsymbol{s}^0, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\delta}, \boldsymbol{\xi}, \boldsymbol{\eta}; \boldsymbol{\gamma})$.

2.3 Bayesian estimation

We assume that researchers observe aggregated sales volume data $\boldsymbol{v}^o = (v_1^o, \dots, v_J^o)'$ and prices $\boldsymbol{p}^o = (p_1^o, \dots, p_J^o)'$ for the J products but not individual purchase incidence. With known market size I, we will specify a conditional distribution of sales volume $(v_0, \boldsymbol{v}') = (v_0, v_1, \dots, v_J)$ given $\boldsymbol{p}$. Theoretically, we would use $I = M$. To avoid computational problems with large M, we choose I appropriately. To obtain the number of product j sold in randomly drawn I consumers, we operationalize

$$v_j = \text{int}\left(I \cdot \frac{v_j^o}{M} + 0.5\right)$$

for $j = 1, \dots, J$, where int$(\cdot)$ is the integral part in the expression $(\cdot)$. Then the number of consumers choosing the outside good $j = 0$ in the I consumers is $v_0 = I - \sum_{j=1}^J v_j$. Let $\boldsymbol{v} = (v_1, \dots, v_J)'$.

We assume $(v_0, \boldsymbol{v}') = (v_0, v_1, \dots, v_J)$ follows a multinomial distribution with index I and category probabilities $(s_0^0, \boldsymbol{s}^{0\prime}) = (s_0^0, s_1^0, \dots, s_J^0)$,

$$f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}) = \frac{I!}{v_0! \cdots v_J!}(s_0^0)^{v_0} \cdots (s_J^0)^{v_J}. \tag{7}$$

This constitutes an assumption that we may aggregate individual choice probabilities s_{ij} to a representative probability s_j^0. Because the taste parameter $\boldsymbol{\theta}_i$ is unobserved, we do not have the joint population distribution of y_i and $\boldsymbol{\theta}_i$ in (3). Therefore, to construct s_j^0, we replace it with a simulation estimator from the randomly drawn I individuals,

$$s_j(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}; \boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{I}\sum_{i=1}^{I} s_{ij}, \tag{8}$$

where $\boldsymbol{y} = (y_1, \dots, y_I)'$ is drawn from the empirical income distribution and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_I)$ is drawn from its posterior distribution. Then we replace

(7) with

$$f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}; \boldsymbol{y}, \boldsymbol{\theta}) = \frac{I!}{v_0! \cdots v_J!} s_0^{v_0} \cdots s_J^{v_J}. \tag{9}$$

Let $\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}; \boldsymbol{y}, \boldsymbol{\theta}) = (s_1, \ldots, s_J)'$. We assume that the unobserved cost characteristics η_j for $j = 1, \ldots, J$ follow

$$\eta_j|\sigma_s^2 \sim N(0, \sigma_s^2).$$

Since the pricing equation (6) is implicit in $\boldsymbol{p}$, we use the transformation of variables $\boldsymbol{\eta} = \mathbf{log}[\boldsymbol{p} + \{(\boldsymbol{\partial G}/\boldsymbol{\partial p})'\}^{-1}\boldsymbol{s}] - \boldsymbol{Z\gamma}$ with a joint distribution $\prod_{j=1}^J N(0, \sigma_s^2)$ for $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_J)'$ to derive the distribution of $\boldsymbol{p}$ as

$$\begin{aligned} &f(\boldsymbol{p}|\boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{\delta}, \boldsymbol{\xi}; \boldsymbol{y}, \boldsymbol{\theta}, \gamma, \sigma_s^2) \\ &= (2\pi\sigma_s^2)^{-\frac{J}{2}} \left|\left| \left(\frac{\boldsymbol{\partial \eta}}{\boldsymbol{\partial p}}\right) \right|\right| \exp\left[-\frac{1}{2\sigma_s^2} \sum_{j=1}^J \left[\log \left[p_j + \left\{ \left(\frac{\boldsymbol{\partial G}}{\boldsymbol{\partial p}}\right)' \right\}_{j\cdot}^{-1} \boldsymbol{s} \right] - \boldsymbol{z}_j \boldsymbol{\gamma} \right]^2 \right] \end{aligned} \tag{10}$$

where $||(\boldsymbol{\partial \eta}/\boldsymbol{\partial p})||$ is the Jacobian and $\{(\boldsymbol{\partial G}/\boldsymbol{\partial p})'\}_{j\cdot}^{-1}$ is the jth row of $\{(\boldsymbol{\partial G}/\boldsymbol{\partial p})'\}^{-1}$.

In terms of consumers' marginal utilities $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$, we assume

$$\boldsymbol{\theta}_i|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \sim MVN(\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}) \tag{11}$$

for $i = 1, \ldots, I$ as in Yang et al. (2003), where $\bar{\boldsymbol{\theta}}$ is a $Q \times 1$ mean vector and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is a $Q \times Q$ variance-covariance matrix. We also assume the unobserved product characteristics ξ_j for $j = 1, \ldots, J$ follow

$$\xi_j|\sigma_d^2 \sim N(0, \sigma_d^2). \tag{12}$$

To obtain the joint posterior of the parameters $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$, $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2, we assume conjugate priors of $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2 as follows.

$$\bar{\boldsymbol{\theta}} \sim MVN(\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}, \boldsymbol{V}_{\bar{\boldsymbol{\theta}}}), \tag{13}$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}} \sim IW_{g_{\boldsymbol{\theta}}}(\boldsymbol{G}_{\boldsymbol{\theta}}), \tag{14}$$

$$\boldsymbol{\gamma} \sim MVN(\bar{\gamma}, \boldsymbol{V}_{\boldsymbol{\gamma}}), \tag{15}$$

$$\sigma_d^2 \sim IG_{g_d/2}(G_d/2), \tag{16}$$

$$\sigma_s^2 \sim IG_{g_s/2}(G_s/2). \tag{17}$$

The $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, $\bar{\gamma}$, $\boldsymbol{V}_{\boldsymbol{\gamma}}$, g_d, G_d, g_s and G_s are hyperparameters.

Let us omit the exogenous $\boldsymbol{y}$, $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{\delta}$ for notational simplicity. Multiplying the distributions of $\boldsymbol{v}$, $\boldsymbol{p}$, $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$ and $\boldsymbol{\xi}$ in (9) through (12), and the priors of $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2 in (13) through (17) so far, we obtain

$$\begin{aligned} f(\boldsymbol{\xi}, \boldsymbol{\theta}, \bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\gamma}, \sigma_d^2, \sigma_s^2 | \boldsymbol{v}, \boldsymbol{p}, \boldsymbol{\xi}) \propto\ & f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}; \boldsymbol{\theta}) f(\boldsymbol{p}|\boldsymbol{\xi}; \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma_s^2) \\ & \times \left[\prod_{i=1}^{I} f(\boldsymbol{\theta}_i|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})\right] \left[\prod_{j=1}^{J} f(\xi_j|\sigma_d^2)\right] \\ & \times f(\bar{\boldsymbol{\theta}}) f(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}) f(\boldsymbol{\gamma}) f(\sigma_d^2) f(\sigma_s^2). \end{aligned} \tag{18}$$

Therefore, we obtain the joint posterior of the parameters by averaging (18) over $\boldsymbol{\xi}$ as

$$f(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\gamma}, \sigma_d^2, \sigma_s^2 | \boldsymbol{v}, \boldsymbol{p}) = \int f(\boldsymbol{\xi}, \boldsymbol{\theta}, \bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\gamma}, \sigma_d^2, \sigma_s^2 | \boldsymbol{v}, \boldsymbol{p}) \boldsymbol{d\xi}. \tag{19}$$

However, it is difficult to calculate the integral (19) analytically due to intricately embedded $\boldsymbol{\xi}$ in the integrand. Therefore, we apply the data augmentation (Tanner and Wong, 1987) to (19), in which we further apply the Gibbs sampler (Geman and Geman, 1984), and then obtain an MCMC algorithm in Appendix A. In the MCMC algorithm, we generate random draws of $\boldsymbol{\xi}$, $\boldsymbol{\theta}$, $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma_{\theta}}$, γ, σ_d^2 and σ_s^2 from their conditional posteriors in Appendix B. Since the conditional posteriors of $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ have nonstandard parametric forms, we apply the Metropolis-Hastings algorithm of the third method in Chib and Greenberg (1995) to them: For the conditional posterior of $\boldsymbol{\xi}$, we first generate proposal draws of $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_J)'$ from $\prod_{j=1}^{J} f(\xi_j|\sigma_d^2)$ in (22) which is a mixture of J identical normal distributions. Then we evaluate the acceptance probability for those proposal draws by acceptance probability of ratio of $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}; \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma_s^2)$ with proposal and current values for $\boldsymbol{\xi}$. For the conditional posterior of $\boldsymbol{\theta}$, we follow a similar way.

3 Simulation study

In this section, we implement simulation studies to validate our proposed method. Specifically, we check if our proposed method can recover true parameter values. Also we are interested in the speed and stability of convergence of parameter distributions, including relative speed of convergence among the parameters. We also call attention to some implementation issues such as nonpositive cost and computational zero likelihood problems.

3.1 Simulation design

We construct data for simulated markets such that computational problems like multicollinearity do not arise in order to focus on convergence behavior.

We set the market size $M = 100{,}000$. On the demand side, we specify consumer i's utility u_{ij} for product j as

$$u_{ij} = \alpha_i \log(y_i - p_j) + \beta_{i1} x_{j1} + \cdots + \beta_{i5} x_{j5} + \xi_j + \varepsilon_{ij} \tag{20}$$

where $\bar{\boldsymbol{\theta}} = (\bar{\alpha}, \bar{\boldsymbol{\beta}}')' = (\bar{\alpha}, \bar{\beta}_1, \ldots, \bar{\beta}_5)' = (3, 2, \ldots, 2)'$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}} = \mathrm{diag}(\sigma^2_{\alpha}, \sigma^2_{\beta_1}, \ldots, \sigma^2_{\beta_5}) = 0.1\boldsymbol{E}_6$. Given $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, we generate true $\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{100{,}000}$ randomly from $MVN(\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ in the market. On the supply side, we choose to use a subset of characteristics that consumers value as cost shifters while allowing demand-side-specific and supply-side-specific variables. Concretely, we specify product j's pricing equation with cost shifters of $\boldsymbol{z}_j = (z_{j1}, \ldots, z_{j4}, z_{j5}) = (x_{j1}, \ldots, x_{j4}, z_{j5})$ as

$$\log \left[p_j + \left\{ \left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}} \right)' \right\}^{-1}_{j\cdot} \boldsymbol{s} \right] = \gamma_1 x_{j1} + \cdots + \gamma_4 x_{j4} + \gamma_5 z_{j5} + \eta_j \tag{21}$$

where $\boldsymbol{\gamma} = (\gamma_1, \ldots, \gamma_5)' = (1, \ldots, 1)'$. Notice that $Q = 5$ and $S = 5$.

In the simulated market, there are five firms each of which sells an exclusive set of two products. This means the market offers $J = 10$ products. We randomly generate values for $x_{11}, \ldots, x_{10,1}$, and then for for $x_{12}, \ldots, x_{10,2}$ from $N(0, 0.1^2)$. If these two have a correlation less than 0.05, we retain both. Otherwise, keep discarding and regenerating $x_{12}, \ldots, x_{10,2}$ until the correlation with $x_{11}, \ldots, x_{10,1}$ is less than 0.05. We randomly generate values for $x_{13}, \ldots, x_{10,3}$ from $N(0, 0.1^2)$, and check if its correlations with two accepted $x_{11}, \ldots, x_{10,1}$ and $x_{12}, \ldots, x_{10,2}$ are less than 0.05. In this way, we generate $x_{j1}, \ldots, x_{j5}$ for $j = 1, \ldots, 10$ so that any two of the five sets of $(x_{11}, \ldots, x_{10,1})', \ldots, (x_{15}, \ldots, x_{10,5})'$ have a correlation less than 0.05 to avoid the multicollinearity problem in (20). We follow a similar process to generate $(z_{15}, \ldots, z_{10,5})'$ which has a correlation less than 0.05 with any set of the other four sets of $(x_{11}, \ldots, x_{10,1})', \ldots, (x_{14}, \ldots, x_{10,4})'$. We set $\boldsymbol{x}_0 = \mathbf{0}$ for the outside good.

We set true values for the variance parameters of the unobserved product and cost characteristics respectively as $\sigma_d^2 = 10^{-4}$ and $\sigma_s^2 = 10^{-4}$. Given σ_d^2 and σ_s^2, we randomly generate a true value for each component of $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{10})'$ from $N(0, \sigma_d^2)$ and that of $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_{10})'$ from $N(0, \sigma_s^2)$ until they have correlations less than 0.05 with all of the observed product characteristics and cost shifters respectively. Note that $\xi_0 = 0$ for the outside good $j = 0$.

We obtain positive values for incomes $y_1, \ldots, y_{100{,}000}$ randomly from the log normal distribution with mean 1 and standard deviation 0.1. We finally determine a pair of equilibrium values for market shares $\boldsymbol{s}^0$ and prices $\boldsymbol{p}^*$ for the 10 products, using the Newton-Raphson method with ten dimensional nonlinear simultaneous equations in (6). We note that $p_0 = 0$ for the outside good $j = 0$. We also note that $\boldsymbol{v}^o = M\boldsymbol{s}^0$.

3.2 MCMC implementation

We now check if our proposed method can recover the true parameter values. We randomly draw $I = 1{,}000$ from the $M = 100{,}000$ consumers for the estimation. To obtain reliable results according to Gelman (1996), we run five independent MCMC sequences with different sets of initial parameter values. See Appendix C.1 for actual initial parameter values used. Note that these initial values are designed to avoid not only nonpositive values for costs in $\boldsymbol{c}$ but also the likelihood with computationally zero which can be time consuming. Each sequence has 30,000 iterations.

We also set hyperparameter values for $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, $\bar{\gamma}$, $\boldsymbol{V}_{\boldsymbol{\gamma}}$ g_d, G_d, g_s and G_s. See Appendix C.1 for actual hyperparameter values used. Of these values, $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, g_d and G_d are designed to avoid the nonpositive costs.

We assess the convergence of the MCMC algorithm by inspecting a time-series plot of draws for each parameter from the five MCMC sequences. We also check if the 95% posterior inverval from the last halves of draws in the five MCMC sequences includes its true value for each parameter.

We summarize the results in time-series plots in Figures 1 and 2 and summary statistics in Table 1. We confirmed the convergences for all of the parameters to their true values. Each of their 95% posterior intervals included the corresponding true value.

There are three points to be noted on the summary statistics in Table 1. First, the posterior standard deviation 0.056 of $\bar{\alpha}$ was far smaller than the posterior standard deviations $(0.44, 0.58, 0.48, 0.43, 0.47)'$ of $\bar{\boldsymbol{\beta}} = (\bar{\beta}_1, \ldots, \bar{\beta}_5)'$. Second, the posterior mean 0.26 of σ^2_α was smaller than the posterior means $(0.44, 0.34, 0.35, 0.35, 0.35)$ of $(\sigma^2_{\beta_1}, \ldots, \sigma^2_{\beta_5})$. Third, the posterior standard deviation 0.13 of σ^2_α was much smaller than the posterior standard deviations $(0.51, 0.28, 0.42, 0.29, 0.44)$ of $(\sigma^2_{\beta_1}, \ldots, \sigma^2_{\beta_5})$. In the following, we will first describe how the first and third facts are consequences of the second fact. Then we will explain why the second fact arose.

As for the first fact, let $\widehat{\bar{\alpha}} = \sum_{i=1}^{1000} \alpha_i/1000$ and $\widehat{\bar{\beta}}_q = \sum_{i=1}^{1000} \beta_{iq}/1000$. Since $\sqrt{var[\widehat{\bar{\alpha}}]} = \sqrt{\sigma^2_\alpha/1000}$ and $\sqrt{var[\widehat{\bar{\beta}}_q]} = \sqrt{\sigma^2_{\beta_q}/1000}$ approximate the posterior standard deviations of $\bar{\alpha}$ and $\bar{\beta}_q$ respectively, the fact that the posterior mean of σ^2_α was less than that of $\sigma^2_{\beta_q}$ for $q = 1, \ldots, 5$ is the reason for the first fact.

As for the third fact, we note that the theoretical posterior standard deviations of σ_α^2 and $\sigma_{\beta_q}^2$ are $\sqrt{var[\sigma_\alpha^2]} = \sqrt{2E[\sigma_\alpha^2]^2/1001}$ and $\sqrt{var[\sigma_{\beta_q}^2]} = \sqrt{2E[\sigma_{\beta_q}^2]^2/1001}$ respectively. Therefore, the fact that the posterior mean of σ_α^2 was less than that of $\sigma_{\beta_q}^2$ for $q = 1, \ldots, 5$ is also the reason for the third fact.

Now we turn our attention to the reason why the posterior mean of σ_α^2 was smaller than that of $\sigma_{\beta_q}^2$ for $q = 1, \ldots, 5$. In our proposed method, β_{iq} exists only in the components of s_{ij} formula in (2). On the other hand, α_i appears not only in the components of s_{ij} but also in the other components of the pricing equation. This formulation worked to generate few negative values for α_i which was the coefficient for $\log(y_i - p_i)$ and thus took postive values for almost all consumers. This means our proposed method restricted the range α_i could take, which in turn induced the second fact of the posterior mean of σ_α^2 being smaller than that of $\sigma_{\beta_q}^2$ for $q = 1, \ldots, 5$. If we had set a higher true value for $\bar{\alpha}$, that is, the true $\alpha_1, \ldots, \alpha_I$ generated from $MVN(\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ would have been also higher, then the posterior mean of σ_α^2 would have been higher and the posterior standard deviations of $\bar{\alpha}$ and σ_α^2 would have been also higher.

We also have some observations from the time-series plots in Figures 1 and 2. The time-series plots for $\bar{\alpha}$, $\boldsymbol{\gamma}$ and σ_α^2 were stable. The stabilities in the time-series plots for $\bar{\alpha}$ and σ_α^2 reflected their smaller standard deviations. We also needed much larger number of iterations for the MCMC algorithm to obtain reliable estimates for $\bar{\boldsymbol{\beta}}$ than that for $\bar{\alpha}$ or $\boldsymbol{\gamma}$.

Figure 1: Time series plots for $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ in the simulation study

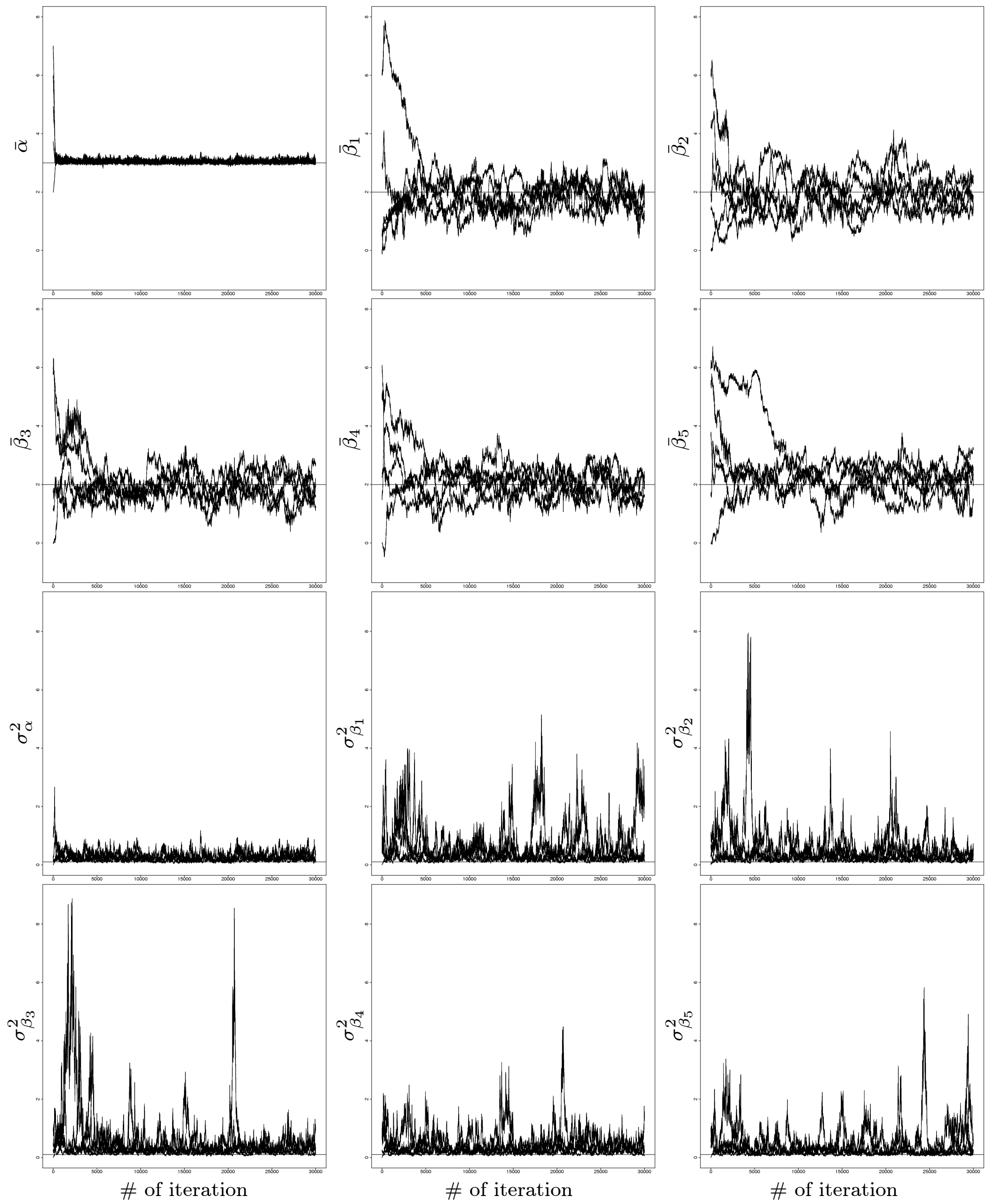


Figure 2: Time series plots for $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2 in the simulation study

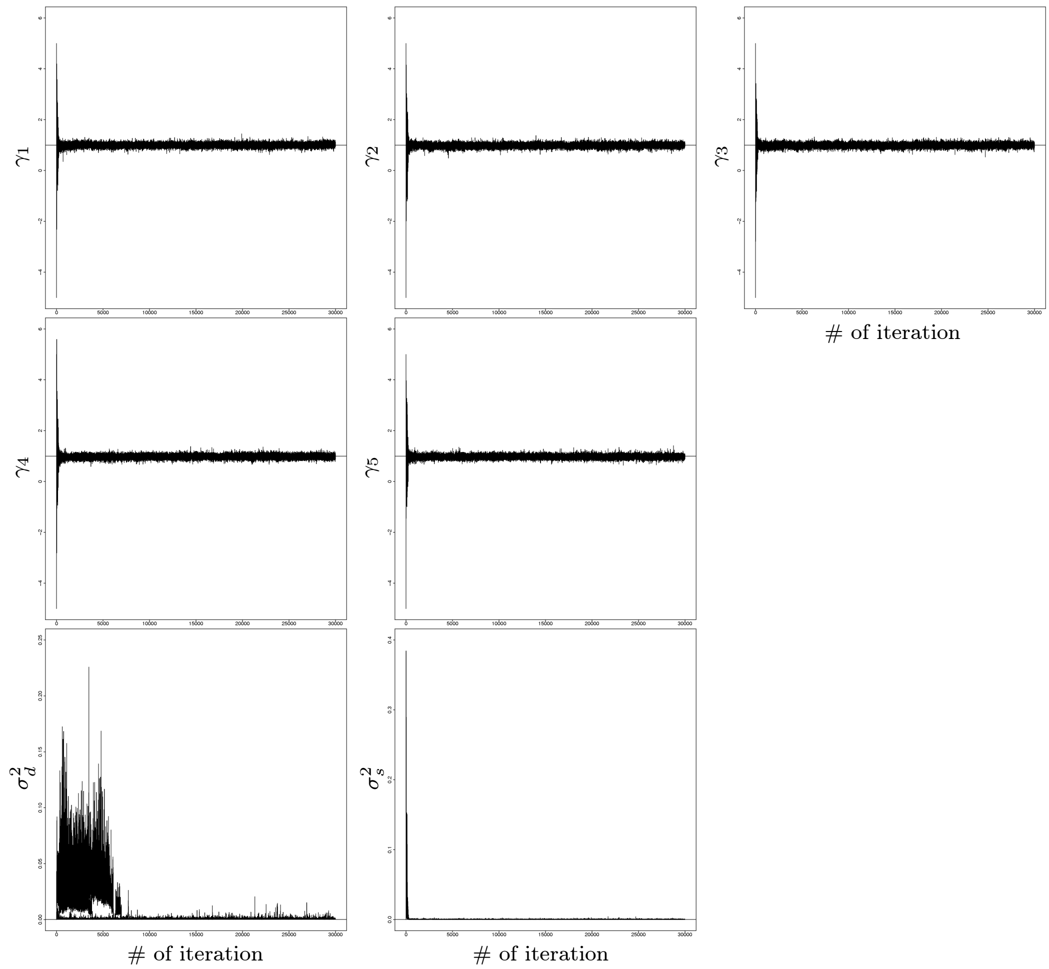


Table 1: Posterior means, standard deviations and quantiles (2.5%, 50% and 97.5%) in the simulation study

Parameter	Mean	Std.Dev.	2.5%	50%	97.5%	True value
$\bar{\alpha}$	3.05	0.056	2.96	3.04	3.18	3
$\bar{\beta}_1$	1.84	0.44	0.93	1.86	2.67	2
$\bar{\beta}_2$	1.95	0.58	0.98	1.92	3.31	2
$\bar{\beta}_3$	1.89	0.48	0.94	1.87	2.78	2
$\bar{\beta}_4$	2.00	0.43	1.26	1.98	2.83	2
$\bar{\beta}_5$	2.16	0.47	1.17	2.21	3.01	2
σ^2_α	0.26	0.13	0.087	0.23	0.59	10^{-1}
$\sigma^2_{\beta_1}$	0.44	0.51	0.097	0.27	2.19	10^{-1}
$\sigma^2_{\beta_2}$	0.34	0.28	0.091	0.26	1.11	10^{-1}
$\sigma^2_{\beta_3}$	0.35	0.42	0.089	0.26	1.05	10^{-1}
$\sigma^2_{\beta_4}$	0.35	0.29	0.091	0.27	1.00	10^{-1}
$\sigma^2_{\beta_5}$	0.35	0.44	0.072	0.20	1.49	10^{-1}
γ_1	1.00	0.056	0.89	1.00	1.11	1
γ_2	0.99	0.056	0.87	0.99	1.10	1
γ_3	0.99	0.057	0.88	0.99	1.10	1
γ_4	0.98	0.056	0.87	0.98	1.09	1
γ_5	0.97	0.052	0.87	0.97	1.08	1
σ^2_d	0.00039	0.00048	0.000094	0.00027	0.0014	10^{-4}
σ^2_s	0.00021	0.00014	0.000069	0.00017	0.00057	10^{-4}

Note: The mean of $R^{(t)}_{\boldsymbol{\xi}^*}$ is 0.76 and that of $R^{(t)}_{\boldsymbol{\theta}^*}$ is 0.68.

3.3 Implementation issues and their remedies

As shown in Appendix D, when we use so-called diffuse priors for $\mathbf{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2 as well as $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\gamma}$ with the other settings being the same as those above, we can overestimate $\mathbf{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2. In summary, we can run into three problems of the nonpositive cost, computational zero likelihood and overestimations of $\mathbf{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2 in our proposed method.

The nonpositive cost problem is generated by choosing inappropriate sets of values for the hyperparameters $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, g_d and G_d and for $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\bar{\boldsymbol{\theta}}^{(0)}$, $\mathbf{\Sigma}_{\boldsymbol{\theta}}^{(0)}$ and $\sigma_d^{2(0)}$. The computational zero likelihood problem was produced by choosing inappropriate sets of values for $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$ and $\sigma_s^{2(0)}$.

We think that there are two causes of the overestimation of $\mathbf{\Sigma}_{\boldsymbol{\theta}}$:

(1) The posterior estimation method for $\boldsymbol{\theta}$ induced the overestimation of $\mathbf{\Sigma}_{\boldsymbol{\theta}}$. In the posterior estimation method for $\boldsymbol{\theta}$, we averaged all of the functions of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_I)$ with respect to $i = 1, \dots, I$ and then reduced the information about the variances and covariances of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_I)$.

(2) The small number $J = 10$ of products induced large posterior variances of σ_d^2 and σ_s^2 which in turn affected the posterior estimation method for $\mathbf{\Sigma}_{\boldsymbol{\theta}}$ indirectly through the posterior estimation method for $\boldsymbol{\theta}$. Specifically, the posterior estimation method for $\boldsymbol{\theta}$ was affected by a large posterior variance of σ_d^2 indirectly through $\boldsymbol{\xi}$ and was affected by a large posterior variance of σ_s^2 directly or indirectly through $\boldsymbol{\gamma}$.

We also think that there was one common cause of the overestimations of

σ_d^2 and σ_s^2: The small number $J = 10$ of products allowed the priors of σ_d^2 and σ_s^2 with low densities around their true values to retain their influences largely in their posteriors; and allowed their posterior means to be larger.

To avoid all of the three problems, we have to choose appropriate sets of values for the hyperparameters $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, g_d, G_d, g_s and G_s and for $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$, $\sigma_d^{2(0)}$ and $\sigma_s^{2(0)}$. It would be nice to have enough information on the parameters so that we are able to specify such appropriate sets of these values. Such information would be obtained from experts in the field, relevant theories and other datasets of relevance. However, it is not always possible to have the information available.

The nonpositive cost and computational zero likelihood problems give us information as to inappropriate sets of these values. Hence, we search for appropriate sets of these values by re-setting them and then re-running the MCMC algorithm several times based on the directions. This process is easy.

On the other hand, it is somewhat tricky for us to set priors of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2 informative enough to estimate them as well as the other parameters correctly. In the following, we suggest three possible criteria to obtain such information. First, meaningful informative prior of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ should produce right sign conditions for components of $\boldsymbol{\theta}_i$ for $i = 1, \ldots, I$ since our experiences show that $\bar{\boldsymbol{\theta}}$ can be correctly estimated with diffuse priors for all of the parameters. Second, meaningful informative priors of σ_d^2 and σ_s^2 should not produce extremely large or small acceptance probability $R_{\boldsymbol{\xi}^*}$ in the Metropolis-Hastings algorithm. The same can be said for informative priors of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ and σ_s^2 relative to $R_{\boldsymbol{\theta}^*}$. Third, meaningful informative priors of σ_d^2 and σ_s^2 should produce the orders of the theoretical prior means of σ_d^2 and

σ_s^2 smaller than the smallest orders of variances of observed product characteristics and of cost shifters respectively. This is because the variances σ_d^2 and σ_s^2 are those of unobserved product and cost characteristics, and their variabilities should be dominated by the variabilities of influential observed product characteristics and cost shifters respectively.

4 Empirical study

To show how practical the proposed method is, we first bring it to the 1995 U.S. new automobile market and obtain estimates for parameters. Given the estimates, we then provide forecasts for market shares for products in the 1996 market and examine their accuracy by comparing with observed market shares. We chose the U.S. automobile market for two reasons. First, it is a market for a differentiated indivisible product where aggregate sales volume data and product characteristic data are publicly available, while disaggregate purchase incidence data are not. While similar methods have been applied to markets for divisible product, one where a consumer purchases one unit of the product during the course of observation fits the model best. Second, it is one of the largest industries in the U.S. and thus has a strong influence on both domestic and international economies.

The choice of the specific year 1995 for the U.S. automobile market was dictated by the following two considerations. First, we use an empirical eight-year total cost of ownership (TCO) as the price variable for consumers.[2] Then

[2]The ownership period of eight years is derived from the median age for vehicle: 7.7 in 1995 and 7.9 in 1996 according to *Ward's Motor Vehicle Facts & Figures 1999.*

this restricts the date used to those for which we have at least eight years of data on non-price components of TCO. The year of 1996 is the most recent year for which we have TCO data, and we use 1996 data for an out-of-sample comparison. This means that the data for 1995 is the most recent we can use for our estimation.

The second consideration is that consumers' preferences were stable between the introduction of the minivan in 1985, and the introduction of hybrid electric vehicles (HEVs) in the 1999.[3]

4.1 Data

We obtain observed data from several sources. For consumers' incomes $\boldsymbol{y}$, we use data in *Integrated Public Use Microdata Series – Current Population Survey 1995* (*IPUMS-CPS 1995*) by Minnesota Population Center in University of Minnesota.[4] Most households received "Total Household Income" which is less than any of the estimated eight-year TCOs for the top 50 models in sales, which will be included in our estimation. Since automobiles are durable, we multiply each household's "Total Household Income" from the

[3]Although electric cars had been expected for some time, we note that the first HEV (in Japan) was the Toyota Prius in 1997, while the HEV was introduced to the U.S. as the Honda Insight in 1999 followed by the Prius in 2000. Sales for the Prius in North America increased from 5,800 in 2000 to 1,838,000 in 2007 according to a news release from TOYOTA (`http://www.toyota.co.jp/jp/news/08/May/nt08_032.html`, in Japanese). This suggests that U.S. consumers were probably not very aware of the future potential of the HEV in 1995 and 1996.

[4]The *IPUMS-CPS 1995* is publicly available on their website (`http://cps.ipums.org/cps/`) and has information on a joint distribution of a variety of consumer demographics for the U.S. population.

IPUMS-CPS 1995 by the same eight year life-span used to compute TCO. Their average annual income is \$42,953.

The market size M in 1995 is assumed to be

$$M = \frac{\text{(vehicles per household)} \times \text{(households)}}{\text{planned holding period}},$$

where the number 2 of vehicles per household and the number 104,212,000 of households are from *Consumer Expenditure Survey 1995* (henceforth *CEX 1995*).[5] The planned holding period of 8 years is chosen to match the period used in computing TCO. The observed sales volumes $\boldsymbol{v}^o$ are from *Ward's Automotive Yearbook 1996*.

The TCO ($\boldsymbol{p}$) for each 1995 model includes its acquisition cost, M&R cost, operating cost and resale value.[6] Acquisition cost is taken from the "mid" retail price from *Ward's Automotive Yearbook 1995*. Operating cost for each 1995 model is calculated from its mileage in *Ward's Automotive Yearbook 1995* and the U.S. city average gasoline prices in 1995 through 2002[7] with an assumption that each vehicle travels 12,000 miles per year during the eight years[8]. The resale value is the discounted average of the wholesale and retail values in 2003, eight years after a consumer purchases a new vehicle in 1995, from *Official Wisconsin Automobile Valuation Guide 2003*.[9]

[5]The number 2 of vehicles per household is a rounded value of 1.9 from the website of *CEX 1995* (`http://www.bls.gov/cex/1995/Standard/cusize.pdf`).

[6]Examination of insurance cost data shows it to be very driver- and location-dependent, even for the same make and model. We omit it from the computation.

[7]The U.S. city average gasoline prices in 1995 through 2002 are on the website of Energy Information Administration (`http://www.eia.doe.gov/emeu/aer/txt/stb0524.xls`).

[8]The number 12,000 miles each vehicle travels per year is an approximate average of 12,385 in 1995 and 11,813 in 1996 from *Ward's Motor Vehicle Facts & Figures 1999*.

[9]The discount rate used is the rate on 10-year U.S. government bonds for each year,

An estimated M&R cost is reported by *The Complete Car/Small Truck Cost Guide* every year. However, Puripunyavanich et al. (2004) showed that it had a serious overestimation for each model because their cost figures for repairs were based on service contract pricing for automobiles' mechanical breakdowns. We follow Puripunyavanich et al. (2004), and obtain an estimated eight-year M&R cost for each 1995 model by the sum of annual products of the following three estimates between 1995 and 2002: An average M&R cost per problem, an expected annual number of problems for each 1995 model, and a series of annual ratios of an M&R cost per problem for each 1995 model relative to a weighted average (in sales) M&R cost per problem. The third factor accounts for variation in M&R costs across models as they age. We used data for 166 models to obtain our estimated M&R cost.

We next explain the number J of models, observed product characteristics and cost shifters we include in our estimation. We obtained data on 160 models available in 1995, accounting for essential all sales of new automobiles in the U.S. To reduce the computational burden, we need to reduce the number of models treated. First, we combined twin models (and the occasional triplets) into a single model in cases where the price range indicated that

adjusted for inflation, in the 8-year period. The former rate is on the web site of Federal Researve Board (`https://www.federalreserve.gov/releases/h15/data/Annual/H15_TCMNOM_Y10.txt`). For the value of inflation is obtained by subtracting the increasing rate in the Consumer Price Index (CPI) from 1994 to 1995 on the web site of Bureau of Labor Statistics in U.S. Department of Labor (`ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt`).

Table 2: Values for the TCO components for the 1995 small cars ($)

Segment	Model	Acquisition	M&R	Operation	Resale	TCO	Sales rank
Small Car	Ford Escort	11,647	1,772	3,287	2,146	14,560	8
	Mercury Tracer						
	Toyota Corolla	13,546	1,580	3,522	3,330	15,318	12
	Geo/Chev. Prizm						
	Honda Civic	12,360	1,243	2,901	3,835	12,669	14
	Chev. Cavalier	13,046	2,530	4,084	2,710	16,950	15
	Pont. Surfire						
	Saturn SC/SL/SW Wagon	12,355	3,001	3,945	3,293	16,008	18

Table 3: Values for the TCO components for the 1995 midium cars ($)

Segment	Model	Acquisition	M&R	Operation	Resale	TCO	Sales rank
Midium Car	Ford Taurus	20,184	2,208	4,988	2,932	24,447	4
	Mercury Sable						
	Honda Accord	20,490	1,617	3,945	5,454	20,598	7
	Toyota Camry	22,015	1,809	4,696	5,566	22,955	9
	Pont. Grand Am	15,669	2,302	4,483	3,130	19,324	10
	Olds. Achievae						
	Buick Skylark						
	Pont. Grand Prix	18,891	2,890	5,191	3,580	23,392	11
	Buick Regal						
	Olds Cutlass Supreme						

Table 4: Values for the TCO components for the 1995 midium cars (continued; $)

Segment	Model	Acquisition	M&R	Operation	Resale	TCO	Sales rank
Midium Car	Chev. Lumina	18,239	2,874	5,350	3,256	23,207	13
	Chev. Monte Carlo						
	Ford Contour	14,831	2,718	4,109	2,531	19,128	23
	Mercury Mistique						
	Chev. Corsica	14,485	2,099	3,945	1,824	18,705	27
	Nissan Altima	19,259	1,478	4,696	3,807	21,626	32
	Dodge Stratus	17,876	2,505	4,931	2,722	22,589	35
	Chrysler Cirrus						
	Nissan Maxima	21,989	1,612	5,451	5,229	23,823	37
	Mazda 626	20,035	1,751	4,696	3,414	23,068	47

Table 5: Values for the TCO components for the 1995 large cars ($)

Segment	Model	Acquisition	M&R	Operation	Resale	TCO	Sales rank
Large Car	Buick Lesable	23,444	2,535	5,368	4,610	26,737	16
	Pont. Vonnevile						
	Olds. 88						
	Dodge Intrepid	21,528	2,433	4,989	3,258	25,691	26
	Chrysler Concorde						
	Eagle Vision						
	Ford Crown Victria	22,895	1,844	5,801	4,335	26,205	28
	Mercury Grand Merquis						
	Cadillac Deville	40,035	3,347	7,154	6,642	43,894	43
	Chev Caprice	24,193	1,947	5,801	9,011	22,931	44
	Buick Roadmaster						
	Lincoln Town Car	39,140	1,794	5,801	6,249	40,486	45

Table 6: Values for the TCO components for the 1995 sporty cars and minivans ($)

Segment	Model	Acquisition	M&R	Operation	Resale	TCO	Sales rank
Sporty Car	Ford Thunderbird LX	23,037	2,296	6,285	4,392	27,226	31
	Mercury Cougar						
	Chev. Camaro	20,834	3,043	6,025	5,515	24,387	33
	Pont. Firebird						
	Ford Mustang	17,410	2,696	5,801	7,661	18,246	34
Minivan	Dodge Caravan	24,755	2,522	5,567	3,957	28,888	3
	Plym. Voyager						
	Chrysler Town & Country						
	Ford Windstar	19,780	2,945	5,801	2,993	25,533	25
	Chev. Astro	20,584	2,221	6,575	3,411	25,969	30
	GMC Safari						
	Mercury Villager	24,639	2,645	5,801	4,042	29,043	36
	Nissan Quest						
	Chev. Lumina Van	18,830	2,578	5,270	3,099	23,579	46
	Pont. Trans Sport						
	Olds. Silhoutte						
	Ford Aerostar	21210	2019	5801	2881	26149	50

Table 7: Values for the TCO components for the 1995 pickup trucks ($)

Segment	Model	Acquisition	M&R	Operation	Resale	TCO	Sales rank
Pickup Truck	Chev. C/K Pickup	18,656	1,770	6,575	5,332	21,669	1
	GMC Siera						
	Ford F-Series	17,606	1,603	6,575	4,995	20,788	2
	Ford Ranger	16,225	1,355	4,793	4,156	18,217	6
	Mazda B-Series						
	Dodge Ram	19,504	2,393	7,586	5,603	23,879	19
	Chev. S-10	15,374	1,738	6,164	3,978	19,297	20
	GMC Sonoma						
	Dodge Neon	11,740	2,446	3,401	2,123	15,463	22
	Plyms. Neon						
	Nissan Pickup	14,309	1,320	5,479	4,144	16,964	38
	Dodge Dakota	15,820	2,395	6,575	3,330	21,459	39
	Toyota Tacoma	17,708	1,400	5,191	5,248	19,051	48

Table 8: Values for the TCO components for the 1995 SUVs ($)

Segment	Model	Acquisition	M&R	Operation	Resale	TCO	Sales rank
SUV	Ford Explorer	23,500	1,898	5,479	4,425	26,452	5
	Chev. Blazer	20,778	2,382	5,801	4,532	24,429	17
	GMC Jimmy						
	Jeep Grand Cherokee	29,255	2,641	6,164	5,014	33,046	21
	Olds. Cutlass Ciera	16,945	2,650	4,893	3,090	21,399	24
	Buick Century						
	Nissan Sentra	14,002	1,210	3,127	2,746	15,593	29
	Nissan 200X						
	Jeep Cherokee	18,267	2,155	5,801	3,779	22,444	40
	Chev. Tahoe	26,995	2,270	8,218	7,540	29,943	41
	GMC Yukon						
	Chev. Suburban	24,867	2,185	8,218	7,172	28,097	42
	GMC Suburban						
	Isuzu Rodeo	21,358	2,449	6,164	3,575	26,395	49
	Honda Passport						

they were very close substitutes up to a small brand effect.[10] This left us with 117 models. Of these, we include the top $J = 50$ in sales, which account for 83.80% of the total sales for the 117 models and 44.61% of the potential market size M respectively. Of the 50 models, 39 are from U.S. manufacturers and 11 are from Japanese manufacturers. We summarize the TCO breakdown for the 50 models, grouped by category, in Tables 2 through 8.

Data on product characteristics in $\boldsymbol{X}$, $\boldsymbol{Z}$ and the manufacturer dummy variables in $\boldsymbol{\delta}$ are from *Ward's Automotive Yearbook 1995*, except for the predicted reliability on a five-point scale from *Consumer Reports April Annual Auto Issue 1996*. Observed product characteristics included in $\boldsymbol{x}_j$ for each model are size (length$\times$width), a measure of safety (a dummy indicating whether dual air bags are available standard or optional), three dummies for minivan, pickup truck and SUV, and two dummies indicating the country of origin of manufacturers (Japan and U.S.). The cost shifters in $\boldsymbol{z}_j$ are an intercept, the logarithm of observed sales volume v_j^o to capture economies of scale, a measure of acceleration (horsepower/weight), mileage, predicted reliability, and observed product characteristics used in $\boldsymbol{x}_j$.

There are three points to be noted. First, the selection of product characteristics and cost shifters follows the studies by BLP (1995), Sudhir (2001), Petrin (2002), BLP (2004) and Myojo (2007). Second, we included mileage and reliability in the cost shifters, but not in the observed product characteristics, because they are accounted for in the operating cost and M&R cost components, respectively, of TCO. Third, we used the intercept instead of

[10] A twin model is a very similar model under a different brand, such as the Ford Taurus and the Mercury Sable.

the U.S dummy in the pricing equation (6) while we used both Japan and U.S. dummies in consumers' utility (1). In other words, the outside good $j = 0$ is the baseline in the demand model, while U.S. models play the role of the baseline in the supply model. Notice that $Q = 7$ and $S = 11$.

4.2 MCMC estimation

We randomly obtain $I = 1{,}000$ households with their eight-year "Total Household Incomes" greater than the largest estimated eight-year TCO.[11] Then we run three independent MCMC sequences with different sets of initial parameter values. Each sequence has $T = 50{,}000$ iterations. Accroding to the implications from Section 3, we set hyperparameter and initial parameter values. Note that the implications include how we avoid the nonpositive cost, computational zero likelihood and overestimation problems. See Appendix C.2 for actual hyperparameter and initial parameter values used.

We assess the convergence of the MCMC algorithm by inspecting a time-series plot of draws for each parameter from the three sequences in Figures 3 through 6. Tables 9 and 10 report summary statistics with the 90% and 95% posterior invervals for each parameter from the last halves of draws in the three MCMC sequences. The averages of the acceptance probabilities of $R_{\boldsymbol{\xi}^*}$ and $R_{\boldsymbol{\theta}^*}$ were 0.82 and 0.76 respectively.

We used a so-called diffuse prior for $\boldsymbol{\gamma}$. Then we confirmed convergences for all of the components of $\boldsymbol{\gamma}$. Since we estimated $\boldsymbol{\gamma}$ correctly by using its diffuse prior in the simulation study in Section 3, we were confident of the

[11] We decide the number of sample consumers included in the estimation, taking account of the number $J = 50$ of models in the estimation as well as our computational burden.

results for $\boldsymbol{\gamma}$ in this empirical study.

On the other hand, the other priors were informative. Then $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma_\theta}$, σ_d^2 and σ_s^2 were estimated well as far as their priors were defined. Note that we set the hyperparameter values for these priors, according to the implications from Section 3.

We next explain the results for $\bar{\boldsymbol{\theta}}$ in detail. The 95% posterior interval for $\bar{\alpha}$ was above zero as expected. The 95% posterior intervals for $\bar{\beta}_{size}$, $\bar{\beta}_{safety}$, $\bar{\beta}_{minivan}$, $\bar{\beta}_{pickup}$ and $\bar{\beta}_{SUV}$ were above zero while those for $\bar{\beta}_{Japan}$ and $\bar{\beta}_{U.S.}$ were below zero. Our results for $\bar{\boldsymbol{\theta}}$ indicated that consumers' utility was positively affected by a product characteristic of size and dual air bags while it was negatively affected by consumers' expense of TCO. The minivans, pickup trucks and SUVs also enhanced consumers' utility. The negative signs for the Japan and U.S. dummies were measured against the outside good. Since the market share for the outside good based on our market size M was the highest value of 55.39% among $j = 0, \ldots, 50$, these negative signs were accordance with the data. There was no difference in consumers' preference for the Japanese models and the U.S. models because the 90% posterior intervals for $\bar{\beta}_{Japan}$ and $\bar{\beta}_{U.S.}$ overlapped.

In our Bayesian estimation, we assumed consumer heterogeneity to be the diagonal components of $\boldsymbol{\Sigma_\theta}$. Our results for the diagonal components of $\boldsymbol{\Sigma_\theta}$ indicated there existed individual differences in preference for the corresponding product characteristics.

As for the results for $\boldsymbol{\gamma}$, the 95% posterior interval for γ_{safety} and 90% posterior interval for γ_{SUV} were above zero while the 95% posterior interval for $\gamma_{mileage}$ was below zero. The results for γ_{safety} and γ_{SUV} indicated that it took cost to produce vehicles with dual air bags and that it took more cost to produce SUVs than the other models. As for the negative $\gamma_{mileage}$, since mileage for a vehicle in general is highly correlated with the number of cylinders and weight for it, the result reflected the fact that it took more cost to produce vehicles with a greater number of cylinders or heavier vehicles. Our results also implied that the Japanese models cost as much as the U.S. models.

4.3 Prediction of the market shares for the top 50 models in sales in the 1996 market

In order to examine accuracy of our estimates from the 1995 data, we provide forecasts for the market shares for the top 50 in sales of the 1996 models. Specifically, we calculate predicted values for the market shares (8) for the outside good and the 50 models ($j = 0, \ldots, 50$), using our estimates. To calculate them, we use estimated eight-year TCOs, observed product characteristics $\boldsymbol{X}$ and unobserved product characteristics $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{50})'$ generated from the posterior of $\boldsymbol{\xi}$ for the top fifty 1996 models in sales, and incomes $\boldsymbol{y}$ for randomly sampled 1,000 consumers from *IPUMS-CPS 1996* and their marginal utilities $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{1,000})$ from the posterior of $\boldsymbol{\theta}$. We then calculate 300 different sets of market shares for $j = 0, \ldots, 50$ in terms of different sets of $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$. Note that the 300 sets for $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ are randomly obtained from last halves of draws in the three MCMC sequences. Our predicted market shares are the means of the 300 sets of the market shares for $j = 0, \ldots, 50$.

The predicted and observed market shares for $j = 0, \ldots, 50$ are presented in Figure 7. We also mark and specify 14 models whose observed market shares were largely over/underestimated. In terms of the accuracy of the predicted market shares for $j = 0, \dots, 50$, the mean of the absolute percentage errors was 48.86% while the mean of the absolute deviations was 0.0037.

Except for the 12 largely over/underestimated models, our estimates predicted the observed market shares well for the other 36 models. Note that the underestimated 9 models had tended to be acheiving top ranks of sales at least in the past 5 years. Therefore, we could fail to capture the reputation of these popular models by each ξ_j. If we could capture it, the overestimations as well as the underestimations could improve. We can thus ascribe these over/underestimations to limitation of our posterior estimation method for $\boldsymbol{\xi}$ as well as that of data.

5 Conclusion and discussion

In this paper, we developed a Bayesian simultaneous demand and supply model for aggregate data in a differentiated indivisible product market. To predict consumers' purchasing behavior, our proposed method requires only aggregate data unlike the Yang et al. (2003) method which requires disaggregate data.

Our Bayesian estimation used the MCMC algorithm including the data augmentation, Gibbs sampler and Metropolis-Hastings algorithm. To take an advantage of conjugacy, we assumed multivariate normals for $\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_I$, $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\gamma}$ even when some of their values should be only positive or negative.

Figure 3: Time series plots for $\bar{\boldsymbol{\theta}}$ in the empirical study

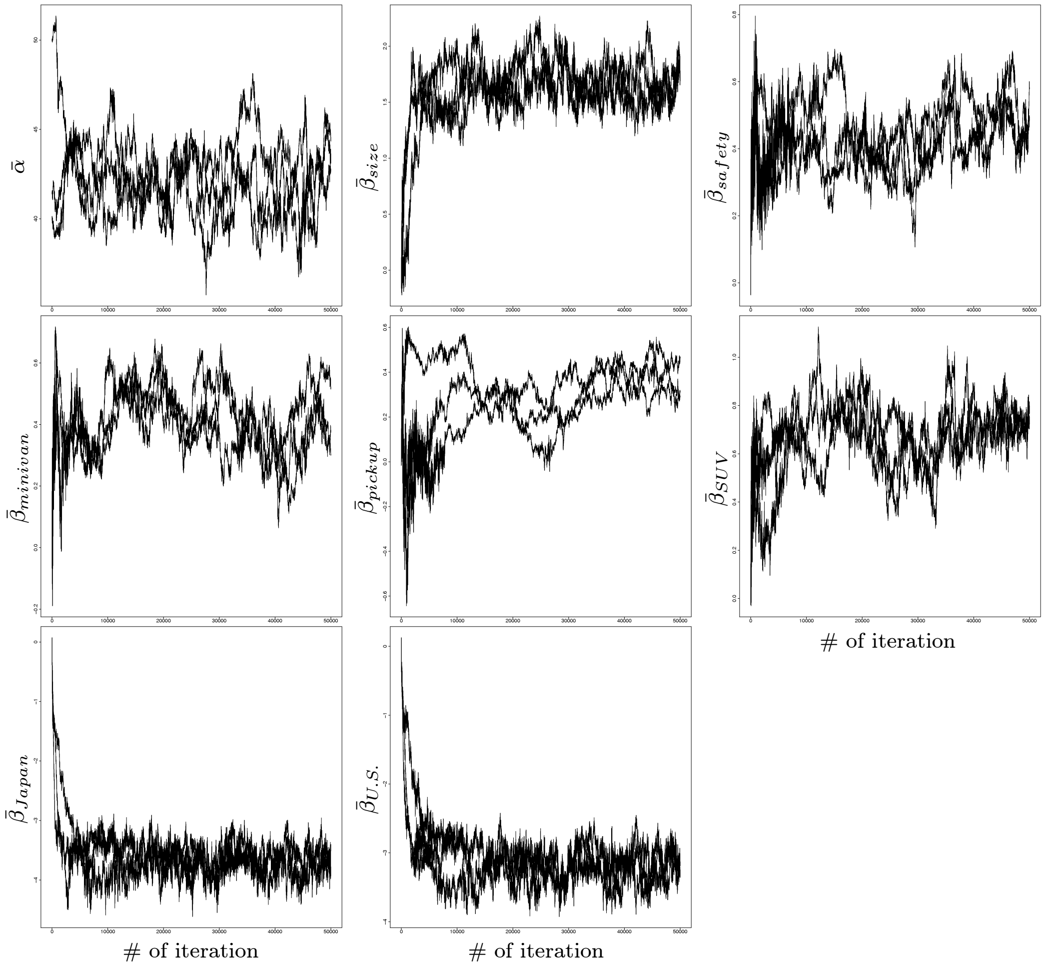


Figure 4: Time series plots for the diagonal components of $\mathbf{\Sigma}_{\boldsymbol{\theta}}$ in the empirical study

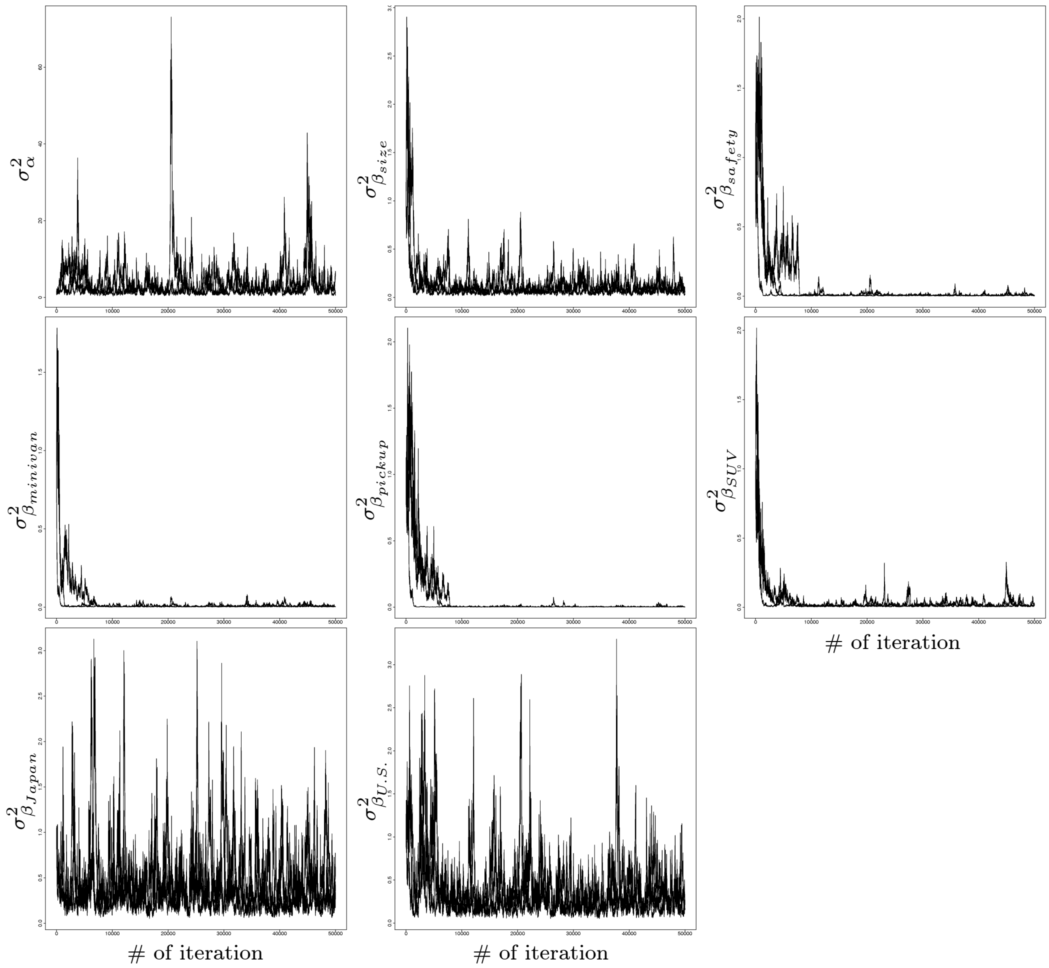


Figure 5: Time series plots for $\boldsymbol{\gamma}$ in the empirical study

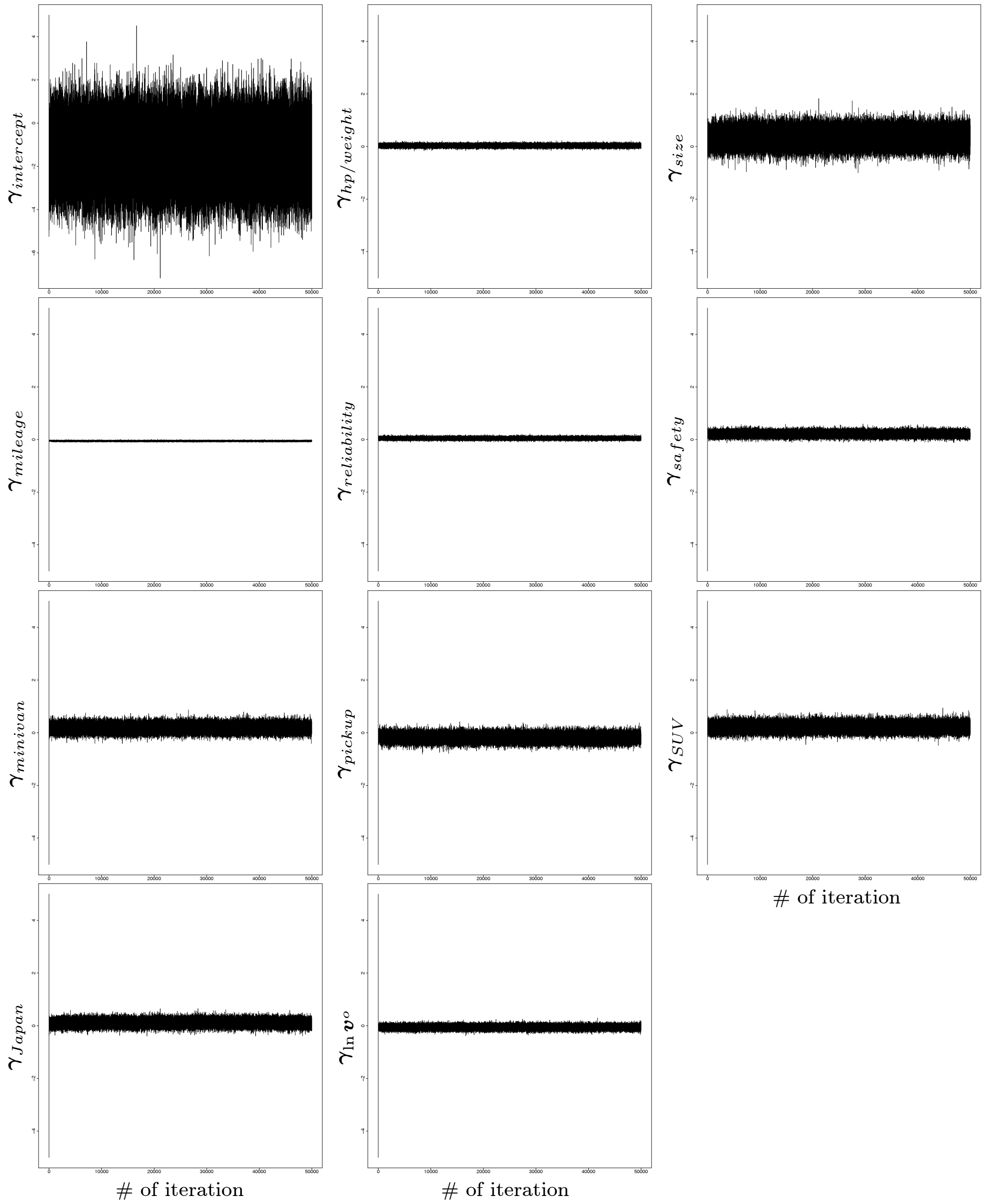


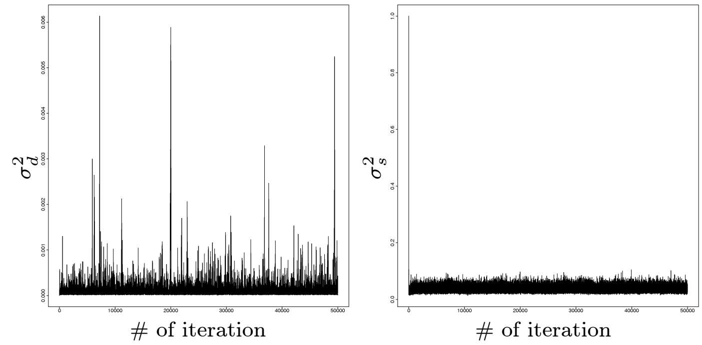
Table 9: Posterior means, standard deviations and quantiles for $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ (2.5%, 5%, 50%, 95% and 97.5%) in the empirical study

Parameter	Mean	Std.Dev.	2.5%	5%	50%	95%	97.5%
$\bar{\alpha}$	41.98	2.01	38.29	38.81	42.00	45.60	46.13
$\bar{\beta}_{size}$	1.67	0.16	1.39	1.43	1.66	1.95	2.00
$\bar{\beta}_{safety}$	0.45	0.087	0.27	0.30	0.45	0.59	0.62
$\bar{\beta}_{minivan}$	0.39	0.10	0.19	0.23	0.39	0.57	0.59
$\bar{\beta}_{pickup}$	0.33	0.10	0.064	0.16	0.34	0.47	0.49
$\bar{\beta}_{SUV}$	0.67	0.11	0.42	0.48	0.68	0.84	0.89
$\bar{\beta}_{Japan}$	−3.70	0.23	−4.15	−4.08	−3.69	−3.32	−3.24
$\bar{\beta}_{U.S.}$	−3.20	0.23	−3.66	−3.59	−3.20	−2.82	−2.75
σ^2_{α}	3.23	3.37	0.72	0.83	2.19	8.89	12.48
$\sigma^2_{\beta_{size}}$	0.099	0.065	0.029	0.033	0.082	0.22	0.27
$\sigma^2_{\beta_{safety}}$	0.0074	0.0065	0.0019	0.0022	0.0055	0.018	0.024
$\sigma^2_{\beta_{minivan}}$	0.0078	0.0067	0.0019	0.0022	0.0057	0.021	0.027
$\sigma^2_{\beta_{pickup}}$	0.0038	0.0041	0.00099	0.0012	0.0028	0.0091	0.014
$\sigma^2_{\beta_{SUV}}$	0.019	0.021	0.0046	0.0053	0.013	0.055	0.076
$\sigma^2_{\beta_{Japan}}$	0.40	0.30	0.11	0.12	0.31	1.01	1.23
$\sigma^2_{\beta_{U.S.}}$	0.31	0.21	0.093	0.11	0.25	0.68	0.80

Table 10: Posterior means, standard deviations and quantiles for $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2 (2.5%, 5%, 50%, 95% and 97.5%) in the empirical study

Parameter	Mean	Std.Dev.	2.5%	5%	50%	95%	97.5%
$\gamma_{intercept}$	−1.40	1.06	−3.48	−3.14	−1.40	0.34	0.68
$\gamma_{hp/weight}$	0.035	0.042	−0.047	−0.033	0.035	0.10	0.12
γ_{size}	0.35	0.27	−0.18	−0.089	0.35	0.79	0.88
$\gamma_{mileage}$	−0.055	0.013	−0.081	−0.077	−0.055	−0.035	−0.031
$\gamma_{reliability}$	0.049	0.035	−0.019	−0.0084	0.049	0.11	0.12
γ_{safety}	0.21	0.078	0.061	0.086	0.21	0.34	0.37
$\gamma_{minivan}$	0.19	0.13	−0.067	−0.023	0.19	0.40	0.44
γ_{pickup}	−0.19	0.12	−0.43	−0.39	−0.19	0.015	0.058
γ_{SUV}	0.22	0.13	−0.043	0.0024	0.22	0.44	0.49
γ_{Japan}	0.12	0.11	−0.095	−0.061	0.12	0.30	0.33
$\gamma_{\ln \boldsymbol{v}^o}$	−0.056	0.065	−0.18	−0.16	−0.056	0.051	0.072
σ_d^2	0.00011	0.00014	0.000024	0.000027	0.000071	0.00028	0.00040
σ_s^2	0.037	0.0085	0.024	0.025	0.035	0.052	0.057

Figure 6: Time series plots for σ_d^2 and σ_s^2 in the empirical study

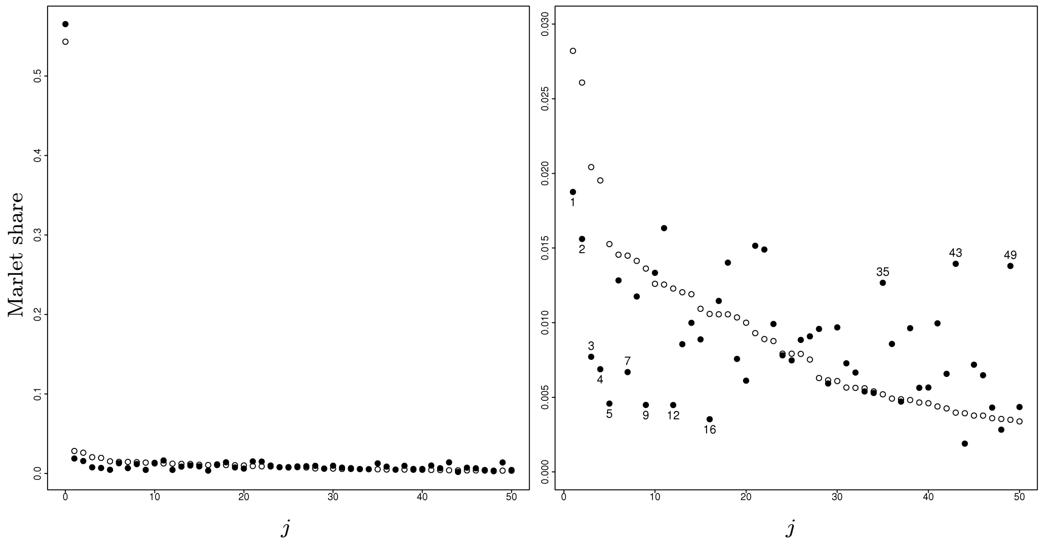


We can instead use log-normals and truncated normals for such parameters, but we lose conjugacy.

We brought our proposed method to simulated data and empirical data from the U.S. automobile market. In the simulation study, we note three problems of the nonpositive cost, computational zero likelihood and overestimations of $\mathbf{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2 our proposed method can run into, for which we proposed remedies. In the empirical study, it could be difficult for us to predict observed market shares for some products due to our limitation of our posterior estimation method for $\boldsymbol{\xi}$ as well as that of data.

In what follows, before discussing future research, we will briefly discuss some ideas on how to modify and improve our proposed method to overcome the nonpositive and overestimation problems by itself and to overcome the limitation of our posterior estimation method for $\boldsymbol{\xi}$. We note that our current MCMC algorithm has already had a mechanism to recover from the computational zero likelihood problem by itself in the Metropolis-Hastings algorithms for $\boldsymbol{\xi}$ and $\boldsymbol{\theta}$ though it can be time-consuming.

Figure 7: Predicted and observed market shares for the top 50 in sales of the 1996 models.



Note: The left figure is with the outside good $j = 0$ while the right is without it. The dots and circles indicate the predicted and observed market shares respectively.

Models with the number j marked on the right figure

1. Ford F-Series
2. Chevrolet C/K pickup / GMC Sierra
3. Dodge Caravan / Plym. Voyager / Chrysler Town & Country
4. Ford Taurus / Mercury Sable
5. Ford Explorer
7. Honda Accord
9. Toyota Camry
12. Chev. Blazer / GMC Jimmy
16. Jeep Grand Cherokee 4WD

35. Chev. Suburban / GMC Suburban
43. Dodge Dakota
49. Lincoln Town Car

We first need to incorporate a mechanism to recover from the nonpositive cost problem by itself. To modify the posterior estimation method for $\boldsymbol{\theta}$ which was a cause of the overestimation of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, a method proposed by Chen and Yang (2007) and Musalem et al. (2009) could be useful. In the method, they first augmented aggregate sales volume with disaggregate purchase incidence data. Then they estimated consumer i's $\boldsymbol{\theta}_i$ corresponding his/her augmented purchase incidence. The method could obtain a more precise set of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$ which would in turn generate a more precise posterior of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$.

Although a common cause of the overestimations of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2 was the small number J of products, there are not always enough products in a market in which we investigate. To increase the number of data on products, we can extend our model for cross-sectional data to that for panel data like Yang et al. (2003). This extention can lead our current unobserved product characteristics $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_J)'$ and cost characteristics $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_J)'$ to be $\boldsymbol{\xi}_n = (\xi_{1n}, \ldots, \xi_{Jn})'$ and $\boldsymbol{\eta}_n = (\eta_{1n}, \ldots, \eta_{Jn})'$ for $n = 1, \ldots, N$, the N being the number of times or locations of observation. This extention also enables us to define heterogenous variances as well as covariances for unobserved product and cost characteristics.

One possible reason for the over/underestimations of the observed market shares for some products was limitation of our posterior estimation method for $\boldsymbol{\xi}$. We need to improve our posterior estimation method for $\boldsymbol{\xi}$. We have two ideas to improve it. The first idea is to estimate the components of $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{50})'$ individually to be able to obtain a product-specific value more precisely. Note that we estimated $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_{50})'$ all at once in **MCMC1** through **MCMC3** in the MCMC algorithm in Appendix A, where it could be more difficult to estimate a product-specfic value for ξ_j for $j = 1, \dots, 50$ efficiently. We estimated $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{50})'$ all at once to reduce computational burden. If we improved our MCMC algorithm to estimate $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{50})'$ individually, we would increase computational burden when we use 50 or more products.

The second idea is to reconsider our assumption to set the prior of σ_d^2. The assumption was that all of the major influential product characteristics on consumers' utility were observed. We need better information to set an alternative prior of σ_d^2.

In future, we need more empirical studies in other differentiated indivisible product markets, using our proposed method. We also wish to have two kinds of comparison. One of them is to compare our proposed method with the past Bayesian methods (Romeo, 2007; Jiang et al., 2009; Musalem et al., 2009). The other is to compare our proposed method with the corresponding past frequentists' methods (BLP, 1995; Sudhir, 2001; Petrin, 2002; BLP, 2004; Myojo, 2007). It is generally said that a Bayesian framework facilitates exact and finite-sample inferences and requires no asymptotic theories while it requires to assume priors for parameters. We could verify these facts as shown in our simulation and empirical studies in Sections 3 and 4 respectively. Especially, we had to set priors for all of the parameters and then had to use additional information for some of the priors to obtain more valid and reliable results. We recognize the setting of priors is a Bayesian disadvantage if we had little confident information about parameters in the framework of the simultaneous demand and supply model. However, Myojo and Kanazawa (2010) showed that the frequentists' simultaneous demand

and supply model framework also required a series of strong assumptions for their asymptotic results to be valid. Additionally, since the asymptotic theories are derived in terms of the number of products, it may not be appropriate to apply the frequentists' framework to some of the markets with the limited number of products. We believe that it is important to uncover the frequentists' and Bayesian relative strengths and weaknesses in analyzing consumers' purchasing behavior in a differentiated indivisible product market with the framework of the simultaneous demand and supply model for aggregate data.

To predict consumers' purchasing behavior from aggregate data is useful because it is more difficult or costs more to obtain disaggregate data. We believe that our proposed method will be an important contribution to the literature of consumers' purchasing behavior.

References

Berry, S., J. Levinsohn, A. Pakes. 1995. Automobile prices in market equilibrium. *Econometrica* **63**(4) 841–890.

Berry, S., J. Levinsohn, A. Pakes. 2004. Differentiated products demand systems from a combination of micro and macro data: the new car market. *Journal of Political Economy* **112**(1) 68–105.

Berry, S. T. 1994. Estimating discrete-choice models of product differentiation. *RAND Journal of Economics* **25**(2) 242–262.

Chen, Y., S. Yang. 2007. Estimating disaggregate models using aggregate data through augmentation of individual choice. *Journal of Marketing Research* **44**(4) 613–621.

Chib, S., E. Greenberg. 1995. Understanding the Metropolis-Hastings algorithm. *American Statistical Association* **49**(4) 327–335.

Consumer Expenditure Survey. 1995. URL `http://www.bls.gov/cex/`. Bureau of Labor Statistics, U.S. Department of Labor.

Consumer Price Index. 1994–1995. URL `ftp://ftp.bls.gov/pub/special.requests/cpi/cpiai.txt`. Bureau of Labor Statistics, U.S. Department of Labor.

Consumer Reports April Annual Auto Issue. 1996. Consumer Union of U.S., Inc. Yonkers, N.Y.

Gelman, A. 1996. Inference and monitoring convergence. W. R. Gilks, S. Richardson, D. J. Spiegelhalter, eds., *Markov chain Monte Carlo in practice*. CAPMAN & HALL, 131–143.

Geman, S., D. Geman. 1984. Stochastic relaxation, Gibbs distributions, and the bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **PAMI-6**(6) 721–741.

Hastings, W. K. 1970. Monte carlo sampling methods using markov chains and their applications. *Biometrika* **57**(1) 97–109.

Integrated Public Use Microdata Series – Current Population Survey. 1995–1996. URL `http://cps.ipums.org/cps/`. Minnesota Population Center, University of Minnesota.

Jiang, R., P. Manchanda, P. E. Rossi. 2009. Bayesian analysis of random coefficient logit models using aggregate data. *Journal of Econometrics* **149** 136–148.

Musalem, A., E. T. Bradlow, J. S. Raju. 2009. Bayesian estimation of random-coefficients choice models using aggregate data. *Journal of Applied Econometrics* **24** 490–516.

Myojo, S. 2007. A study on the u.s. consumers' automobile preferences. Ph.D. thesis, Systems and Information Engineering: University of Tsukuba.

Myojo, S., Y. Kanazawa. 2010. On asymptotic properties of the parameters of differentiated product demand and supply systems when demographically-categorized purchasing pattern data are available. Working paper, University of Tsukuba.

Official Wisconsin Automobile Valuation Guide. 2003. The Wisconsin Automobile and Truck Dealers Association, Madison, Wisconsin.

Petrin, A. 2002. Quantifying the benefits of new products: the case of the minivan. *Journal of Political Economy* **110** 705–729.

Puripunyavanich, V., S. Myojo, Y. Kanazawa. 2004. Estimating the maintenance and repair cost in life cycle cost calculation: A case of automobile ownership in the u.s. *The Journal of Management Accounting, Japan* **13** 3–24.

Romeo, C. J. 2007. A Gibbs sampler for mixed logit analysis of differentiated product markets using aggregate data. *Computational Economics* **29** 33–68.

Sudhir, K. 2001. Competitive pricing behavior in the auto market: A structural analysis. *Marketing Science* **20**(1) 42–60.

Tanner, M. A., W. H. Wong. 1987. The calculation of posterior distribution by data augmentation. *Journal of the American Statistical Association* **82** 528–54.

The Complete Car Cost Guide. 1992–1994/1995–1999. IntelliChoice, Inc. San Jose/Campbell, CA.

The Complete Small Truck Cost Guide. 1992–1994/1995–1999. IntelliChoice, Inc. San Jose/Campbell, CA.

TOYOTA News Release on May 15th (in Japanese). 2008. URL `http://www.toyota.co.jp/jp/news/08/May/nt08_032.html`. TOYOTA.

U.S. City Average Gasoline Prices. 1995–2002. URL `http://www.eia.doe.gov/emeu/aer/txt/stb0524.xls`. Energy Information Administration.

U.S. Treasury Securities at 10-year maturity. 1995. URL `https://www.federal` reserve.gov/releases/h15/data/Annual/H15_TCMNOM_Y10.txt. Federal Researve Board.

Villas-Boas, J. M., R. S. Winer. 1999. Endogeneity in brand choice models. *Management Science* **45**(10) 1324–1338.

Ward's Automotive Yearbook. 1995–1996. Wards Communications. Detroit, MI.

Ward's Motor Vehicle Facts & Figures. 1999. Ward's Communications. Southfield, MI.

Yang, S., Y. Chen, G. M. Allenby. 2003. Bayesian analysis of simultaneous demand and supply. *Quantitative Marketing and Economics* **1** 251–275.

A MCMC algorithm

Let $\boldsymbol{\xi}^{(t)}$, $\boldsymbol{\theta}^{(t)}$, $\bar{\boldsymbol{\theta}}^{(t)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t)}$, $\boldsymbol{\gamma}^{(t)}$, $\sigma_d^{2(t)}$ and $\sigma_s^{2(t)}$ denote values for $\boldsymbol{\xi}$, $\boldsymbol{\theta}$, $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2 respectively at the tth iteration for $t = 0, \ldots,$ in which $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$, $\sigma_d^{2(0)}$ and $\sigma_s^{2(0)}$ especially denote their initial values we have to set; and let $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_I^*)$ and $\boldsymbol{\xi}^* = (\xi_1^*, \ldots, \xi_J^*)'$ denote proposal draws for $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$ and $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_J)'$ in their Metropolis-Hastings algorithms respectively. Our MCMC algorithm is as follows.

MCMC0 Set values for the hyperparameters $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, $\bar{\gamma}$, $\boldsymbol{V}_{\boldsymbol{\gamma}}$, g_d, G_d, g_s and G_s, and $\boldsymbol{\theta}^{(0)}$, $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$, $\sigma_d^{2(0)}$, $\sigma_s^{2(0)}$ and $\boldsymbol{\xi}^{(0)}$.

For $t = 1, \ldots,$

MCMC1 Generate each component of $\boldsymbol{\xi}^* = (\xi_1^*, \ldots, \xi_J^*)'$ randomly from $N(0, \sigma_d^{2(t-1)})$.

MCMC2 Calculate

$$R_{\boldsymbol{\xi}^*}^{(t)} = \begin{cases} \min\left(\frac{f(\boldsymbol{v},\boldsymbol{p}|\boldsymbol{\xi}^*,\boldsymbol{\theta}^{(t-1)},\boldsymbol{\gamma}^{(t-1)},\sigma_s^{2(t-1)})}{f(\boldsymbol{v},\boldsymbol{p}|\boldsymbol{\xi}^{(t-1)},\boldsymbol{\theta}^{(t-1)},\boldsymbol{\gamma}^{(t-1)},\sigma_s^{2(t-1)})}, 1\right) \\ \text{if the denominator } f(\boldsymbol{v},\boldsymbol{p}|\boldsymbol{\xi}^{(t-1)},\boldsymbol{\theta}^{(t-1)},\boldsymbol{\gamma}^{(t-1)},\sigma_s^{2(t-1)}) > 0, \\ \\ 1 \text{ otherwise.} \end{cases}$$

MCMC3 Set $\boldsymbol{\xi}^{(t)} = \boldsymbol{\xi}^*$ with probability $R_{\boldsymbol{\xi}^*}^{(t)}$ or $\boldsymbol{\xi}^{(t)} = \boldsymbol{\xi}^{(t-1)}$ with probability $1 - R_{\boldsymbol{\xi}^*}^{(t)}$.

MCMC4 Generate each component of $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^*, \dots, \boldsymbol{\theta}_I^*)$ randomly from $MVN(\bar{\boldsymbol{\theta}}^{(t-1)}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})$.

MCMC5 Calculate

$$R_{\boldsymbol{\theta}^*}^{(t)} = \begin{cases} \min\left(\frac{f(\boldsymbol{v},\boldsymbol{p}|\boldsymbol{\xi}^{(t)},\boldsymbol{\theta}^*,\boldsymbol{\gamma}^{(t-1)},\sigma_s^{2(t-1)})}{f(\boldsymbol{v},\boldsymbol{p}|\boldsymbol{\xi}^{(t)},\boldsymbol{\theta}^{(t-1)},\boldsymbol{\gamma}^{(t-1)},\sigma_s^{2(t-1)})}, 1\right) \\ \text{if the denominator } f(\boldsymbol{v},\boldsymbol{p}|\boldsymbol{\xi}^{(t)},\boldsymbol{\theta}^{(t-1)},\boldsymbol{\gamma}^{(t-1)},\sigma_s^{2(t-1)}) > 0, \\ \\ 1 \text{ otherwise.} \end{cases}$$

MCMC6 Set $\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^*$ with probability $R_{\boldsymbol{\theta}^*}^{(t)}$ or $\boldsymbol{\theta}^{(t)} = \boldsymbol{\theta}^{(t-1)}$ with probability $1 - R_{\boldsymbol{\theta}^*}^{(t)}$.

MCMC7 Generate $\bar{\boldsymbol{\theta}}^{(t)}$ from $f(\bar{\boldsymbol{\theta}}|\boldsymbol{\theta}^{(t)}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})$.

MCMC8 Generate $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t)}$ from $f(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}|\boldsymbol{\theta}^{(t)}, \bar{\boldsymbol{\theta}}^{(t)})$.

MCMC9 Generate $\boldsymbol{\gamma}^{(t)}$ from $f(\boldsymbol{\gamma}|\boldsymbol{\theta}^{(t)}, \sigma_s^{2(t-1)}, \boldsymbol{\xi}^{(t)}, \boldsymbol{p})$.

MCMC10 Generate $\sigma_s^{2(t)}$ from $f(\sigma_s^2|\boldsymbol{\theta}^{(t)}, \boldsymbol{\gamma}^{(t)}, \boldsymbol{\xi}^{(t)}, \boldsymbol{p})$.

MCMC11 Generate $\sigma_d^{2(t)}$ from $f(\sigma_d^2|\boldsymbol{\xi}^{(t)})$.

MCMC12 If random draws from the Metropolis-Hastings algorithm for $\boldsymbol{\theta}$ in **MCMC4** through **MCMC6**, from $f(\bar{\boldsymbol{\theta}}|\boldsymbol{\theta}^{(t)}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})$ in **MCMC7**, from $f(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}|\boldsymbol{\theta}^{(t)}, \bar{\boldsymbol{\theta}}^{(t)})$ in **MCMC8**, from $f(\boldsymbol{\gamma}|\boldsymbol{\theta}^{(t)}, \sigma_s^{2(t-1)}, \boldsymbol{\xi}^{(t)}, \boldsymbol{p})$ in **MCMC9**, from $f(\sigma_s^2|\boldsymbol{\theta}^{(t)}, \boldsymbol{\gamma}^{(t)}, \boldsymbol{\xi}^{(t)}, \boldsymbol{p})$ in **MCMC10** and from $f(\sigma_d^2|\boldsymbol{\xi}^{(t)})$ in **MCMC11** stabilize, then stop the iteration. Otherwise increase t by one and return to **MCMC1**.

B Conditional posteriors

The conditional posteriors we use in the MCMC algorithm in Appendix A are

$$
\begin{aligned}
& f(\boldsymbol{\xi}|\boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma_d^2, \sigma_s^2, \boldsymbol{v}, \boldsymbol{p}) \propto f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}; \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma_s^2) \left[\prod_{j=1}^{J} f(\xi_j|\sigma_d^2)\right] \\
& \propto s_0^{v_0} \cdots s_J^{v_J} \\
& \quad \times (\sigma_s^2)^{-\frac{J}{2}} \left|\left|\left(\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{p}}\right)\right|\right| \exp\left[-\frac{1}{2\sigma_s^2} \sum_{j=1}^{J} \left[\log\left[p_j + \left\{\left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}}\right)'\right\}_{j\cdot}^{-1} \boldsymbol{s}\right] - \boldsymbol{z}_j \boldsymbol{\gamma}\right]^2\right] \\
& \quad \times (\sigma_d^2)^{-\frac{J}{2}} \exp\left(-\frac{1}{2} \sum_{j=1}^{J} \xi_j^2\right), \qquad (22)
\end{aligned}
$$

$$
\begin{aligned}
& f(\boldsymbol{\theta}|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\gamma}, \sigma_s^2, \boldsymbol{\xi}, \boldsymbol{v}, \boldsymbol{p}) \propto f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}; \boldsymbol{\theta}, \boldsymbol{\gamma}, \sigma_s^2) \left[\prod_{i=1}^{I} f(\boldsymbol{\theta}_i|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})\right] \\
& \propto s_0^{v_0} \cdots s_J^{v_J} \\
& \quad \times (\sigma_s^2)^{-\frac{J}{2}} \left|\left|\left(\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{p}}\right)\right|\right| \exp\left[-\frac{1}{2\sigma_s^2} \sum_{j=1}^{J} \left[\log\left[p_j + \left\{\left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}}\right)'\right\}_{j\cdot}^{-1} \boldsymbol{s}\right] - \boldsymbol{z}_j \boldsymbol{\gamma}\right]^2\right]
\end{aligned}
$$

$$\times|\boldsymbol{\Sigma}_{\boldsymbol{\theta}}|^{-\frac{I}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^{I}(\boldsymbol{\theta}_i-\bar{\boldsymbol{\theta}})'\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}(\boldsymbol{\theta}_i-\bar{\boldsymbol{\theta}})\right\},$$

$$\bar{\boldsymbol{\theta}}|\boldsymbol{\theta},\boldsymbol{\Sigma}_{\boldsymbol{\theta}}\sim MVN((I\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}+\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}^{-1})^{-1}(I\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}\boldsymbol{\nu}+\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}^{-1}\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}),(I\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{-1}+\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}^{-1})^{-1}),$$

$$\boldsymbol{\Sigma}_{\boldsymbol{\theta}}|\boldsymbol{\theta},\bar{\boldsymbol{\theta}}\sim IW_{g_{\boldsymbol{\theta}}+I}\left(\sum_{i=1}^{I}(\boldsymbol{\theta}_i-\bar{\boldsymbol{\theta}})(\boldsymbol{\theta}_i-\bar{\boldsymbol{\theta}})'+\boldsymbol{G}_{\boldsymbol{\theta}}\right),$$

$$\boldsymbol{\gamma}|\boldsymbol{\theta},\sigma_s^2,\boldsymbol{\xi},\boldsymbol{p}\sim MVN((\boldsymbol{\Sigma}_{s*}^{-1}+\boldsymbol{V}_{\boldsymbol{\gamma}}^{-1})^{-1}(\boldsymbol{\mu}_{\boldsymbol{\gamma}*}+\boldsymbol{V}_{\boldsymbol{\gamma}}^{-1}\bar{\boldsymbol{\gamma}}),(\boldsymbol{\Sigma}_{s*}^{-1}+\boldsymbol{V}_{\boldsymbol{\gamma}}^{-1})^{-1}),$$

$$\sigma_d^2|\boldsymbol{\xi}\sim IG_{\frac{g_d+J}{2}}\left(\frac{1}{2}\left(\sum_{j=1}^{J}\xi_j^2+G_d\right)\right),$$

$$\sigma_s^2|\boldsymbol{\theta},\boldsymbol{\gamma},\boldsymbol{\xi},\boldsymbol{p}\sim IG_{\frac{g_s+J}{2}}\left(\frac{1}{2}\left(\sum_{j=1}^{J}\left[\log\left[p_j+\left\{\left(\frac{\partial\boldsymbol{G}}{\partial\boldsymbol{p}}\right)'\right\}_{j\cdot}^{-1}\boldsymbol{s}\right]-\boldsymbol{z}_j\boldsymbol{\gamma}\right]^2+G_s\right)\right),$$

where

$$\boldsymbol{\nu}=\frac{1}{I}\sum_{i=1}^{I}\boldsymbol{\theta}_i,\ \ \boldsymbol{\mu}_{\boldsymbol{\gamma}*}=\frac{1}{\sigma_s^2}\sum_{j=1}^{J}\boldsymbol{z}_j'\left[\log\left[p_j+\left\{\left(\frac{\partial\boldsymbol{G}}{\partial\boldsymbol{p}}\right)'\right\}_{j\cdot}^{-1}\boldsymbol{s}\right]\right],\ \ \boldsymbol{\Sigma}_{s*}^{-1}=\frac{1}{\sigma_s^2}\sum_{j=1}^{J}\boldsymbol{z}_j'\boldsymbol{z}_j.$$

C Hyperparameter and initial parameter values

C.1 Hyperparameter and initial parameter values for the simulation study

Hyperparameter values for the simulation study in Section 3 are

$$\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}=(\mu_{\bar{\alpha}},\mu_{\bar{\beta}_1},\dots,\mu_{\bar{\beta}_5})'=(20,0,\dots,0)',$$

$$
\begin{aligned}
&\boldsymbol{V}_{\bar{\boldsymbol{\theta}}} = 10^2\boldsymbol{E}_6,\\
&g_{\boldsymbol{\theta}} = 10,\\
&\boldsymbol{G}_{\boldsymbol{\theta}} = \mathrm{diag}(G_\alpha, G_{\beta_1}, \ldots, G_{\beta_4}, G_{\beta_5}) = \mathrm{diag}(1.2, 1.2, \ldots, 1.2, 0.9), \quad (23)\\
&\bar{\boldsymbol{\gamma}} = \boldsymbol{0},\\
&\boldsymbol{V}_{\boldsymbol{\gamma}} = 10^2\boldsymbol{E}_5,\\
&g_d = 5,\\
&G_d = 0.0012, \quad (24)\\
&g_s = 5,\\
&G_s = 0.0009. \quad (25)
\end{aligned}
$$

For initial parameter values, we have three sets as the large, middle and small sets. One of the five MCMC sequences has only the large set and another one has only the small set. Each of the remaining three MCMC seqeuences has a middle set with a uniformly generated random value for each parameter with the upper and lower bounds corresponding to the values in the forementioned large and small sets respectively. For the MCMC sequence with the large set, we have

$$
\begin{aligned}
&\bar{\boldsymbol{\theta}}^{(0)} = (\bar{\alpha}^{(0)}, \bar{\beta}_1^{(0)}, \ldots, \bar{\beta}_5^{(0)})' = (7, 6, \ldots, 6)', \ \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)} = \boldsymbol{E}_6,\\
&\boldsymbol{\gamma}^{(0)} = (5, \ldots, 5)', \ \sigma_d^{2(0)} = 10^{-2}, \ \sigma_s^{2(0)} = 10^{-2}.
\end{aligned}
$$

For the MCMC sequence with the small set, we have

$$
\begin{aligned}
&\bar{\boldsymbol{\theta}}^{(0)} = (\bar{\alpha}^{(0)}, \bar{\beta}_1^{(0)}, \ldots, \bar{\beta}_5^{(0)})' = (2, 0, \ldots, 0)', \ \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)} = 10^{-10}\boldsymbol{E}_6,\\
&\boldsymbol{\gamma}^{(0)} = (-5, \ldots, -5)', \ \sigma_d^{2(0)} = 10^{-10}, \ \sigma_s^{2(0)} = 10^{-10}.
\end{aligned}
$$

Given $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$ and $\sigma_d^{2(0)}$ in each MCMC sequence, we randomly generate each component of $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, \ldots, \boldsymbol{\theta}_{1,000}^{(0)})$ from $MVN(\bar{\boldsymbol{\theta}}^{(0)}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)})$ and that of $\boldsymbol{\xi}^{(0)} = (\xi_1^{(0)}, \ldots, \xi_{10}^{(0)})'$ from $N(0, \sigma_d^{2(0)})$.

C.2 Hyperparameter and initial parameter values for the empirical study

Hyperparameter values for the empirical study in Section 4 are

$$
\begin{aligned}
\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}} &= (\mu_{\bar{\alpha}}, \mu_{\bar{\beta}_{size}}, \mu_{\bar{\beta}_{safety}}, \mu_{\bar{\beta}_{minivan}}, \mu_{\bar{\beta}_{pickup}}, \mu_{\bar{\beta}_{SUV}}, \mu_{\bar{\beta}_{Japan}}, \mu_{\bar{\beta}_{U.S.}})' \\
&= (43.92, 1.73, 0.46, 0.43, 0.32, 0.68, -3.53, -3.18)', \\
\boldsymbol{V}_{\bar{\boldsymbol{\theta}}} &= \mathrm{diag}(V_{\bar{\alpha}}, V_{\bar{\beta}_{size}}, V_{\bar{\beta}_{safety}}, V_{\bar{\beta}_{minivan}}, V_{\bar{\beta}_{pickup}}, V_{\bar{\beta}_{SUV}}, V_{\bar{\beta}_{Japan}}, V_{\bar{\beta}_{U.S.}}) \\
&= \mathrm{diag}(3.12, 0.10, 0.017, 0.029, 0.030, 0.033, 0.17, 0.21), \\
g_{\boldsymbol{\theta}} &= 12, \\
\boldsymbol{G}_{\boldsymbol{\theta}} &= \mathrm{diag}(G_{\alpha}, G_{\beta_{size}}, G_{\beta_{safety}}, G_{\beta_{minivan}}, G_{\beta_{pickup}}, G_{\beta_{SUV}}, G_{\beta_{Japan}}, G_{\beta_{U.S.}}) \\
&= \mathrm{diag}(3.12, 0.12, 0.0083, 0.0073, 0.0041, 0.019, 0.50, 0.41), \\
\bar{\boldsymbol{\gamma}} &= \mathbf{0}, \\
\boldsymbol{V}_{\boldsymbol{\gamma}} &= 10^2 \boldsymbol{E}_{11}, \\
g_d &= 5, \\
G_d &= 0.0003, \\
g_s &= 5, \\
G_s &= 0.03.
\end{aligned}
$$

Notice that the prior of $\boldsymbol{\gamma}$ with $\bar{\boldsymbol{\gamma}}$ and $\boldsymbol{V}_{\boldsymbol{\gamma}}$ is diffuse while the other priors with the corresponding hyperparameters are informative.

We next explain initial parameter values. For $\bar{\boldsymbol{\beta}}^{(0)}$ in $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$, $\sigma_d^{2(0)}$ and $\sigma_s^{2(0)}$, we use the same values for all of the three MCMC sequences.

Specifically, we set $\bar{\boldsymbol{\beta}}^{(0)} = \mathbf{0}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)} = \boldsymbol{E}_8$, $\sigma_d^{2(0)} = 10^{-10}$ and $\sigma_s^{2(0)} = 1$. For $\bar{\alpha}^{(0)}$ in $\bar{\boldsymbol{\theta}}^{(0)}$ and $\boldsymbol{\gamma}^{(0)}$, we have three sets as the large, middle and small sets. One of the three MCMC sequences has only the large set and another one has only the small set. The remaining MCMC seqeuence has only the middle set with a uniformly generated random value for each parameter with the upper and lower bounds corresponding to the values in the forementioned large and small sets respectively. For the MCMC sequence with the large set, we have $\bar{\alpha}^{(0)} = 60$ and $\boldsymbol{\gamma}^{(0)} = (5, \ldots, 5)'$. For the MCMC sequence with the small set, we have $\bar{\alpha}^{(0)} = 50$ and $\boldsymbol{\gamma}^{(0)} = (-5, \ldots, -5)'$. Given $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$ and $\sigma_d^{2(0)}$ in each MCMC sequence, we randomly generate each component of $\boldsymbol{\xi}^{(0)} = (\xi_1^{(0)}, \ldots, \xi_{50}^{(0)})'$ from $N(0, \sigma_d^{2(0)})$ and that of $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, \ldots, \boldsymbol{\theta}_{1,000}^{(0)})$ from $MVN(\bar{\boldsymbol{\theta}}^{(0)}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)})$.

D An example simulation generating overestimations

We show overestimations of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2 when we use so-called diffuse priors for them as well as $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\gamma}$. To obtain the diffuse priors for $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2, we reset the scale hyperparameter $\boldsymbol{G}_{\boldsymbol{\theta}}$ in the inverse Wishart prior of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ in (14) to be $\mathrm{diag}(3, \ldots, 3, 3)$ instead of the original $\mathrm{diag}(1.2, \ldots, 1.2, 0.9)$ in (23); the scale hyperparameter G_d in the inverse gamma prior of σ_d^2 in (16) to be 0.03 instead of the original 0.0012 in (24); and the scale hyperparameter G_s in the inverse gamma prior of σ_s^2 in (17) to be 0.03 instead of the original $G_s = 0.0009$ in (25).

We summarize the results of the MCMC in time-series plots in Figures 8 and 9 and summary statistics in Table 11. We confirmed the convergences for the components of $\bar{\boldsymbol{\theta}} = (\bar{\alpha}, \bar{\boldsymbol{\beta}}')'$ and $\boldsymbol{\gamma}$ to their true values. Each of their 95% posterior intervals also included the corresponding true value. We were not able to confirm that the diagonal components of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, σ_d^2 and σ_s^2 converged to their true values as far as their time-series plots and summary statistics were concerned. Their true values were out of the corresponding 95% posterior intervals and our proposed method overestimated them.

Figure 8: Time series plots for $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ from the MCMC estimation given the diffuse priors for all of the parameters

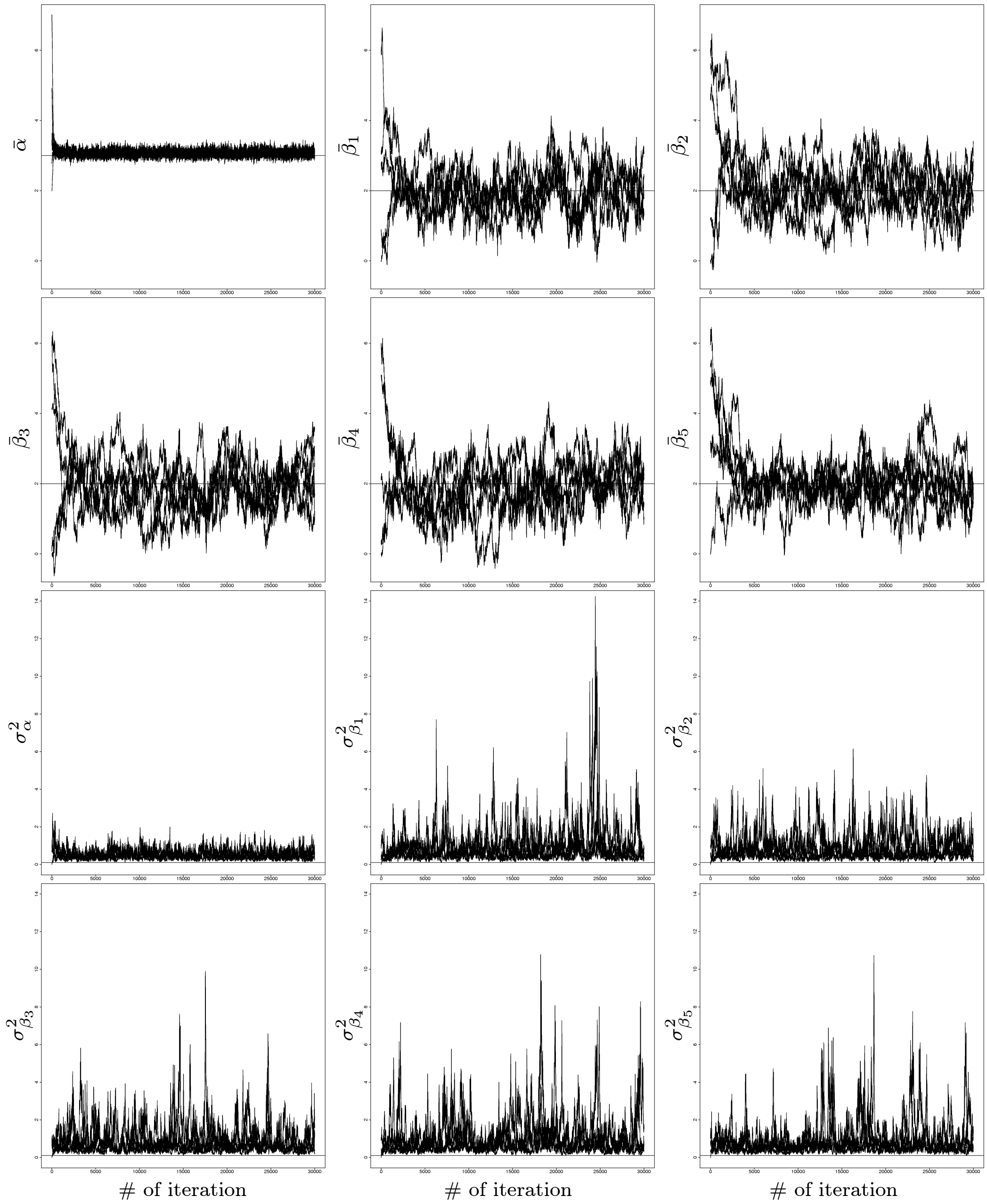


Figure 9: Time series plots for $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2 from the MCMC estimation given the diffuse priors for all of the parameters

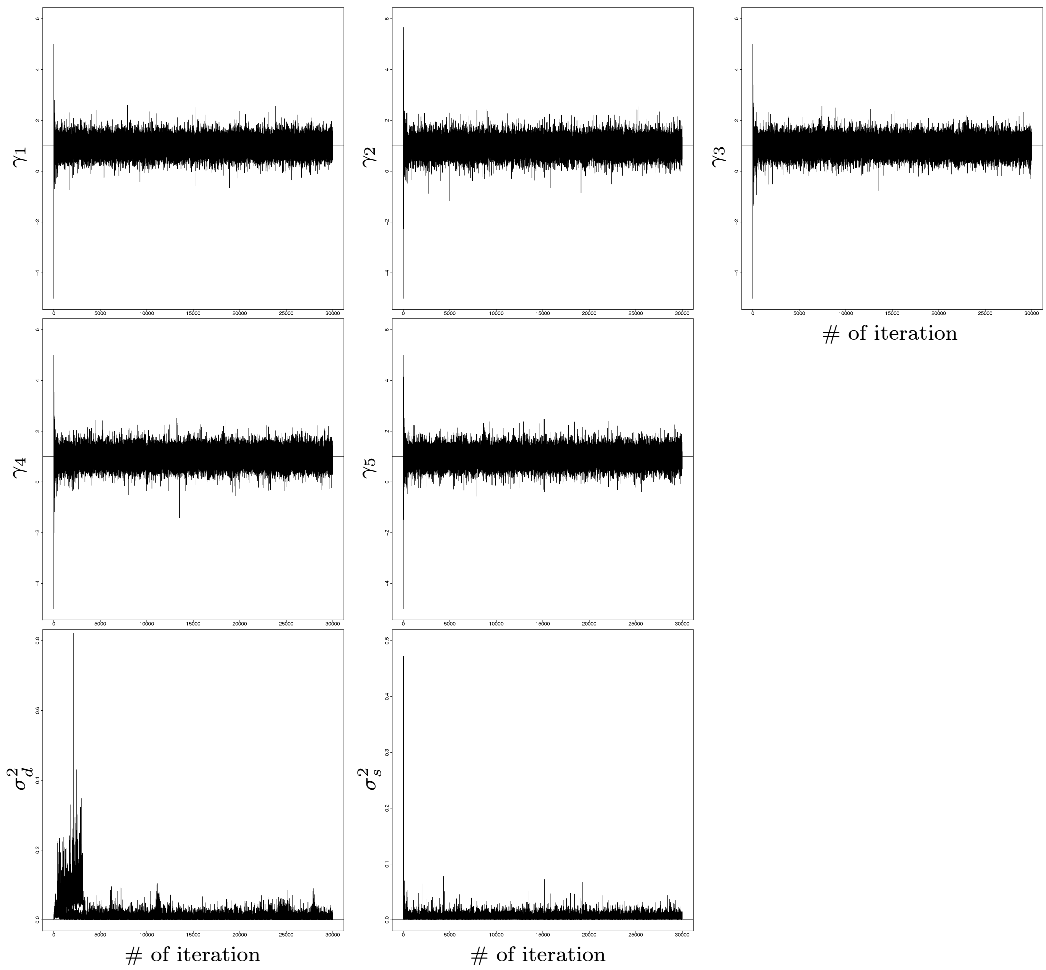


Table 11: Posterior means, standard deviations and quantiles (2.5%, 50% and 97.5%) from the MCMC estimation given the diffuse priors for all of the parameters

Parameter	Mean	Std.Dev.	2.5%	50%	97.5%	True value
$\bar{\alpha}$	3.06	0.088	2.90	3.06	3.25	3
$\bar{\beta}_1$	1.99	0.62	0.84	1.99	3.23	2
$\bar{\beta}_2$	2.01	0.60	0.84	1.99	3.22	2
$\bar{\beta}_3$	1.91	0.61	0.83	1.89	3.23	2
$\bar{\beta}_4$	2.00	0.61	0.89	1.98	3.29	2
$\bar{\beta}_5$	1.94	0.62	0.82	1.91	3.39	2
σ^2_α	0.50	0.22	0.21	0.45	1.05	10^{-1}
$\sigma^2_{\beta_1}$	0.93	0.91	0.24	0.66	3.20	10^{-1}
$\sigma^2_{\beta_2}$	0.79	0.55	0.23	0.62	2.29	10^{-1}
$\sigma^2_{\beta_3}$	0.78	0.63	0.21	0.60	2.45	10^{-1}
$\sigma^2_{\beta_4}$	0.95	0.87	0.22	0.68	3.39	10^{-1}
$\sigma^2_{\beta_5}$	0.84	0.76	0.24	0.60	2.99	10^{-1}
γ_1	0.99	0.22	0.54	0.99	1.44	1
γ_2	0.98	0.23	0.52	0.97	1.43	1
γ_3	0.98	0.23	0.53	0.98	1.44	1
γ_4	0.98	0.23	0.53	0.98	1.44	1
γ_5	0.98	0.22	0.54	0.98	1.42	1
σ^2_d	0.0073	0.0053	0.0022	0.0058	0.021	10^{-4}
σ^2_s	0.0044	0.0027	0.0016	0.0037	0.011	10^{-4}

Note: The mean of $R^{(t)}_{\boldsymbol{\xi}^*}$ is 0.35 and that of $R^{(t)}_{\boldsymbol{\theta}^*}$ is 0.81.