数理解析研究所講究録
第 1703 巻 2010 年 91-111

An empirical study on consumers' purchasing behavior with a Bayesian simultaneous demand and supply model for market-level data in the U.S. automobile market

Yutaka YONETANI, Yuichiro KANAZAWA, Satoshi MYOJO and Stephen John Turnbull[1]

1 Introduction

In this paper, we will show an example of using a Bayesian estimation method for a simultaneous demand and supply model for market-level data which was proposed by Yonetani et al. (2010). To show the example, we also use data from the U.S. automobile market.

The method quantitatively investigates consumers' preference in a differentiated product market. The method also has the following three features as the past research had (Berry, 1994; Berry, Levinsohn and Pakes, 1995 (henceforth, BLP, 1995); Sudhir, 2001; Petrin, 2002; Berry, Levinsohn and Pakes, 2004 (henceforth, BLP, 2004); Myojo, 2007; Romeo, 2007; Jiang et al., 2008; Musalem et al., 2009; Myojo and Kanazawa, 2010). First, the method does not require consumer-level purchase incidence data but market-level sales volume data. In some markets, researchers can access to the latter much easier than the for-

[1]Yutaka YONETANI is a doctoral candicate in Doctoral Program in Quantitative Finance and Management, Graduate School of Systems and Information Engineering, University of Tsukuba, 1-1-1 Ten-noh-dai, Tsukuba, Ibaraki 305-8573, Japan. His e-mail address is yyonetan@sk.tsukuba.ac.jp.
Yuichiro KANAZAWA is Professor of Statistics in Department of Social Systems and Management, Graduate School of Systems and Information Engineering, University of Tsukuba. His e-mail address is kanazawa@sk.tsukuba.ac.jp.
Satoshi MYOJO is a research fellow at the 1st Theory-Oriented Research Group, National Institute of Science and Technology Policy (NISTEP), Ministry of Education, Culture, Sports, Science and Technology (MEXT), Chuo Godochosya 7 East 16th Floor, 3-2-2 Kasumigaseki, Chiyoda-ku, Tokyo 100-0013, Japan. His e-mail address is smyojo@nistep.go.jp.
Stephen John Turnbull is Associate Professor of Economics in Department of Social Systems and Management, Graduate School of Systems and Information Engineering, University of Tsukuba. His e-mail address is turnbull@sk.tsukuba.ac.jp.

This research is supported in part by the Grant-in-Aid for Scientific Research (C) 16510103, (C) 18530227, (C) 21530207 and (B)20310081 from the Japan Society for the Promotion of Science.

mer. Second, the method takes account for price endogeneity which reflects the fact that consumers' purchasing behavior is affected by prices set by suppliers based on expected consumers' responses. Note that sales volumes are obtained by aggregating the consumers' purchasing behaviors. Third, the method also takes account for consumer heterogeneity which indicates that individuals have different preferences. Additionally, the estimation for the proposed method is implemented in Bayesian framework. We can thus resort to the same Bayesian advantages as those similar Bayesian methods has (Yang et al., 2003; Romeo, 2007; Jiang et al., 2008; Musalem et al., 2009): We can conceptually construct a joint posterior distribution of parameters for a complex model given their prior distribution and data distribution; we can facilitate finite-sample inferences about not only the parameters but also their functions of interest without relying on asymptotics; and we can incorporate pre-existing information about the parameters in their prior.

The remaining of this paper is organized as follows. In Section 2, we will briefly review the method. In Section 3, we will show the example of using the method with data from the U.S. automobile market. Conclusions and discussions will be presented in Section 4.

2 Review of the method

2.1 Model specification

Demand Model

The simultaneous demand and supply model has two endogenous variables of sales volume and price. The demand model explains the former.

Suppose that there are J products in a differentiated product market where a consumer purchases one unit of a product with the highest utility in the course of observation; and that we observe a $J \times 1$ sales volume vector $\boldsymbol{v}^o = (v_1^o, \ldots, v_J^o)'$ and the overall market size $M = \sum_{j=0}^{J} v_j^o$ with v_0^o being the number of consumers choosing the outside good $j = 0$.

The consumer i's utility for product j is

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

where y_i and $\boldsymbol{\theta}_i = (\alpha_i, \boldsymbol{\beta}_i')'$ are his/her income and $Q \times 1$ coefficient vector respectively, p_j, $\boldsymbol{x}_j$ and ξ_j are product j's unit price, $1 \times (Q-1)$ observed characteristic vector and unobserved (by researchers) characteristic respectively, and ε_{ij} is a consumer-level sampling error term. Notice that different values for

$\boldsymbol{\theta}_i$ among consumers reflect consumer heterogeneity. For $j = 0$, we assume $p_0 = 0$, $\boldsymbol{x}_0 = \boldsymbol{0}$ and $\xi_0 = 0$.

Suppose that ε_{ij} is independent of the other terms and independently and identically Gumbel distributed across consumers and products in addition to that a consumer purchases one unit of a product with the highest utility. Then the consumer i's logit choice probability for product j,

$$s_{ij} = s_{ij}(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}, y_i, \boldsymbol{\theta}_i) = \frac{\exp\{\alpha_i \log(y_i - p_j) + \boldsymbol{x}_j \boldsymbol{\beta}_i + \xi_j\}}{\sum_{k=0}^{J} \exp\{\alpha_i \log(y_i - p_k) + \boldsymbol{x}_k \boldsymbol{\beta}_i + \xi_k\}}, \tag{1}$$

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

To explain the sales volume, we obtain the market share for product j as

$$s_j = s_j(\boldsymbol{p}) = s_j(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{I} \sum_{i=1}^{I} s_{ij} \tag{2}$$

with a sample of I consumers instead of averaging s_{ij} over the population, where $\boldsymbol{y} = (y_1, \dots, y_I)'$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \dots, \boldsymbol{\theta}_I)$. Let $\boldsymbol{s} = (s_1, \dots, s_J)'$ denote a $J \times 1$ market share vector. Let $\boldsymbol{v} = (v_1, \dots, v_J)'$ also denote a $J \times 1$ sales volume vector in the I consumers, for which we define $v_j = \text{int}(Iv_j^o/M + 0.5)$ for $j = 1, \dots, J$. Note $\text{int}(\cdot)$ is the integral part in the expression $(\cdot)$ and $v_0 = M - \sum_{j=1}^{J} v_j$.

Supply Model

The supply model explains the other endogenous variable of price. Suppose that there are F firms in an oligopolistic market for the J products with Bertrand competition; and that each firm f produces an exclusive subset of the J products and sets prices for its products to maximize its total profit

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

where c_j is a unit cost. Then we obtain the first order condition according to the Bertrand competition for $j = 1, \dots, J$ as

$$\boldsymbol{p} = -\left\{ \left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}} \right)' \right\}^{-1} \boldsymbol{s} + \boldsymbol{c}, \tag{3}$$

assuming the inverse above exists, where $\boldsymbol{c} = (c_1, \dots, c_J)'$ and $(\partial \boldsymbol{G}/\partial \boldsymbol{p}) = (\partial \boldsymbol{s}/\partial \boldsymbol{p}) * \boldsymbol{\delta}$ with the sign $*$ indicating the element-by-element multiplication of the matrices it connects and with the (j, k) element of $\boldsymbol{\delta}$ being 1 if the products j and k are produced by the same firm and 0 otherwise. In the first order condition, we model $\log c_j = \boldsymbol{z}_j\boldsymbol{\gamma} + \eta_j$, where $\boldsymbol{z}_j$ and η_j are product j's $1 \times S$ cost characteristic vector and unobserved cost characteristic respectively and $\boldsymbol{\gamma}$ is a $S \times 1$ coefficient vector. Note $\boldsymbol{Z} = (\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_J)'$ and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_J)'$. We obtain the pricing equation by substituting $\exp\{\boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\eta}\}$ for $\boldsymbol{c}$ in (3) and obtain

$$\log\left[\boldsymbol{p} + \left\{\left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}}\right)'\right\}^{-1} \boldsymbol{s}\right] = \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\eta} \tag{4}$$

with $\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{s}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{\delta}, \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\eta}, \boldsymbol{\gamma})$.

Simultaneous demand and supply model

From (2) and (4), the simultaneous demand and supply model is written as

$$\begin{aligned} &\boldsymbol{s} \mid \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{y}, \boldsymbol{\theta}, && (2) \\ &\boldsymbol{p} \mid \boldsymbol{s}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{\delta}, \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\eta}, \boldsymbol{\gamma}. && (4) \end{aligned}$$

Given the market size M, we can say that product j has its market share s_j is equivalent in saying its sales volume is v_j for $j = 1, \ldots, J$ in a sample of I consumers. Hence we rewrite the simultaneous demand and supply model as

$$\begin{aligned} &\boldsymbol{v} \mid \boldsymbol{p}, \boldsymbol{\xi}, \boldsymbol{\theta}, && (2)' \\ &\boldsymbol{p} \mid \boldsymbol{v}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\gamma}. && (4)' \end{aligned}$$

Note that we also removed exogenous observed $\boldsymbol{X}$, $\boldsymbol{Z}$ and $\boldsymbol{\delta}$ for notational simplicity; and removed $\boldsymbol{y}$ because we will obtain them from an income distribution and thus we can regard them as exogenous observed data.

2.2 Bayesian estimation

Priors and distributions of the endogenous variables

For unobserved product and cost characteristics, we assume $\xi_j \sim N(0, \sigma_d^2)$ and $\eta_j \sim N(0, \sigma_s^2)$ for $j = 1, \ldots, J$. We also assume $\boldsymbol{\theta}_i \sim MVN(\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ for $i = 1, \ldots, I$ with $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ to be estimated in the Bayesian hierarchical model. To obtain a joint posterior of the parameters $\boldsymbol{\theta}$, $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2, we first hypothesize conjugate priors as

$$\begin{aligned} &\bar{\boldsymbol{\theta}} \sim MVN(\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}, \boldsymbol{V}_{\bar{\boldsymbol{\theta}}}), \ \ \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \sim IW_{g_{\theta}}(\boldsymbol{G}_{\boldsymbol{\theta}}), \\ &\boldsymbol{\gamma} \sim MVN(\bar{\boldsymbol{\gamma}}, \boldsymbol{V}_{\boldsymbol{\gamma}}), \ \ \sigma_d^2 \sim IG_{g_d/2}(G_d/2), \ \ \sigma_s^2 \sim IG_{g_s/2}(G_s/2). \end{aligned}$$

We also use a multinomial distribution for the endogenous $\boldsymbol{v}$ with $\boldsymbol{s}$ as

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

Since there is no explicit solution for $\boldsymbol{p}$ in (4), we apply the variable transformation method with $\eta_j \sim N(0,\sigma_s^2)$ for $j = 1,\ldots,J$ and derive a distribution for the other endogenous $\boldsymbol{p}$ as

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

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

Posterior estimation

We derive the following joint posterior from the distributions so far.

$$\begin{aligned} f(\boldsymbol{\xi},\boldsymbol{\theta},\bar{\boldsymbol{\theta}},\boldsymbol{\Sigma_\theta},\boldsymbol{\gamma},\sigma_d^2,\sigma_s^2|\boldsymbol{v},\boldsymbol{p}) \;\propto\;& f(\boldsymbol{v}|\boldsymbol{p},\boldsymbol{\xi},\boldsymbol{\theta})f(\boldsymbol{p}|\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\gamma},\sigma_s^2)\\ &\times\left[\prod_{j=1}^{J}f(\xi_j|\sigma_d^2)\right]\left[\prod_{i=1}^{I}f(\boldsymbol{\theta}_i|\bar{\boldsymbol{\theta}},\boldsymbol{\Sigma_\theta})\right]\\ &\times f(\bar{\boldsymbol{\theta}})f(\boldsymbol{\Sigma_\theta})f(\boldsymbol{\gamma})f(\sigma_d^2)f(\sigma_s^2) \end{aligned} \tag{5}$$

To obtain the joint posterior of the parameters, we require to average (5) over $\boldsymbol{\xi}$ as

$$f(\boldsymbol{\theta},\bar{\boldsymbol{\theta}},\boldsymbol{\Sigma_\theta},\boldsymbol{\gamma},\sigma_d^2,\sigma_s^2|\boldsymbol{v},\boldsymbol{p}) = \int f(\boldsymbol{\xi},\boldsymbol{\theta},\bar{\boldsymbol{\theta}},\boldsymbol{\Sigma_\theta},\boldsymbol{\gamma},\sigma_d^2,\sigma_s^2|\boldsymbol{v},\boldsymbol{p})d\boldsymbol{\xi}. \tag{6}$$

However, since it is difficult to solve the integral in (6) analytically, we numerically obtain the joint posterior of the parameters by using the data augmentation (Tanner and Wong, 1987) in which we further apply the Gibbs sampler (Geman and Geman, 1984). Specifically, we form the MCMC algorithm in Appendix A. In the MCMC algorithm, we generate random draws of $\boldsymbol{\xi}$, $\boldsymbol{\theta}$, $\bar{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma_\theta}$, $\boldsymbol{\gamma}$, σ_d^2 and σ_s^2 from their conditional posteriors of

$$f(\boldsymbol{\xi}|\boldsymbol{\theta},\boldsymbol{\gamma},\sigma_d^2,\sigma_s^2,\boldsymbol{v},\boldsymbol{p}) \propto f(\boldsymbol{v},\boldsymbol{p}|\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\gamma},\sigma_s^2)\left[\prod_{j=1}^{J}f(\xi_j|\sigma_d^2)\right]$$
$$\propto s_0^{v_0}\cdots s_J^{v_J}$$

$$\times(\sigma_s^2)^{-\frac{J}{2}}\left\|\left(\frac{\partial\eta}{\partial p}\right)\right\|\exp\left[-\frac{1}{2\sigma_s^2}\sum_{j=1}^J\left[\log\left[p_j+\left\{\left(\frac{\partial G}{\partial p}\right)'\right\}^{-1}_{j\cdot}s\right]-z_j\gamma\right]^2\right]$$

$$\times(\sigma_d^2)^{-\frac{J}{2}}\exp\left(-\frac{1}{2}\sum_{j=1}^J\xi_j^2\right), \qquad (7)$$

$$f(\theta|\bar{\theta},\Sigma_\theta,\gamma,\sigma_s^2,\xi,v,p)\propto f(v,p|\xi,\theta,\gamma,\sigma_s^2)\left[\prod_{i=1}^I f(\theta_i|\bar{\theta},\Sigma_\theta)\right]$$

$$\propto s_0^{v_0}\cdots s_J^{v_J}$$

$$\times(\sigma_s^2)^{-\frac{J}{2}}\left\|\left(\frac{\partial\eta}{\partial p}\right)\right\|\exp\left[-\frac{1}{2\sigma_s^2}\sum_{j=1}^J\left[\log\left[p_j+\left\{\left(\frac{\partial G}{\partial p}\right)'\right\}^{-1}_{j\cdot}s\right]-z_j\gamma\right]^2\right]$$

$$\times|\Sigma_\theta|^{-\frac{I}{2}}\exp\left\{-\frac{1}{2}\sum_{i=1}^I(\theta_i-\bar{\theta})'\Sigma_\theta^{-1}(\theta_i-\bar{\theta})\right\}$$

$$\bar{\theta}|\theta,\Sigma_\theta\sim MVN((I\Sigma_\theta^{-1}+V_{\bar{\theta}}^{-1})^{-1}(I\Sigma_\theta^{-1}\nu+V_{\bar{\theta}}^{-1}\mu_{\bar{\theta}}),(I\Sigma_\theta^{-1}+V_{\bar{\theta}}^{-1})^{-1}),$$

$$\Sigma_\theta|\theta,\bar{\theta}\sim IW_{g_\theta+I}\left(\sum_{i=1}^I(\theta_i-\bar{\theta})(\theta_i-\bar{\theta})'+G_\theta\right),$$

$$\gamma|\theta,\sigma_s^2,\xi,p\sim MVN((\Sigma_{s*}^{-1}+V_\gamma^{-1})^{-1}(\mu_{\gamma*}+V_\gamma^{-1}\bar{\gamma}),(\Sigma_{s*}^{-1}+V_\gamma^{-1})^{-1}),$$

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

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

respectively, where

$$\nu=\frac{1}{I}\sum_{i=1}^I\theta_i,\quad \mu_{\gamma*}=\frac{1}{\sigma_s^2}\sum_{j=1}^J z_j'\left[\log\left[p_j+\left\{\left(\frac{\partial G}{\partial p}\right)'\right\}^{-1}_{j\cdot}s\right]\right],\quad \Sigma_{s*}^{-1}=\frac{1}{\sigma_s^2}\sum_{j=1}^J z_j'z_j.$$

Notice that the conditional posteriors of ξ and θ have nonstandard parametric forms. In the MCMC algorithm, we thus apply the Metropolis-Hastings algorithm of the third method in Chib and Greenberg (1995) to them: To generate ξ, we first generate proposal draws from $\prod_{j=1}^J f(\xi_j|\sigma_d^2)$ in (7) which is a mixture of J identical normal distributions. Then we accept the proposal draws by an acceptance probability of the ratio of $f(\boldsymbol{v},\boldsymbol{p}|\boldsymbol{\xi},\boldsymbol{\theta},\boldsymbol{\gamma},\sigma_s^2)$ with proposal and current $\boldsymbol{\xi}$. To generate $\boldsymbol{\theta}$, we follow the similar way.

2.3 Implementation issues and their remedies

Before our example of using Yonetani et al.'s (2010) method, we summarize implementation issues and remedies for the issues proposed in their paper. Their method can run into problems of nonpositive cost and computational zero likelihood. It can also encounter a problem of overestimations of $\boldsymbol{\Sigma_\theta}$, σ_d^2 and σ_s^2 while it can correctly estimate $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\gamma}$ when we use so-called diffuse priors for these parameters.

The nonpositive cost problem is induced by inappropriate values for the hyperparameters $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G_\theta}$, g_d and G_d and for $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$ and $\sigma_d^{2(0)}$ in **MCMC0** in the MCMC algorithm in Appendix A. The computational zero likelihood problem is induced by inappropriate values for $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$ and $\sigma_s^{2(0)}$. The overestimation problem is induced by inappropriate priors for $\boldsymbol{\Sigma_\theta}$, σ_d^2 and σ_s^2.

To avoid the three problems, we have to set appropriate values for the hyperparameters $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G_\theta}$, 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 $\sigma_s^{2(0)}$. We can easily find the inappropriate values generating the nonpositive cost and computational zero likelihood problems by re-setting their values and re-running the MCMC algorithm several times.

On the other hand, it is difficult to avoid inappropriate priors of $\boldsymbol{\Sigma_\theta}$, σ_d^2 and σ_s^2. If we have little information enough to set appropriate priors for them, the following three remedies were proposed by Yonetani et al. (2010). First, we can obtain informative priors of σ_d^2 and σ_s^2 based on $R_{\boldsymbol{\xi}^\bullet}$ and obtain those of $\boldsymbol{\Sigma_\theta}$ and σ_s^2 based on $R_{\boldsymbol{\theta}^\bullet}$ in the MCMC algorithm. Priors with extremely large or small $R_{\boldsymbol{\xi}^\bullet}$ or $R_{\boldsymbol{\theta}^\bullet}$ should be avoided because the MCMC algorithm with them could be inefficient. Second, we can also obtain an appropriate prior of $\boldsymbol{\Sigma_\theta}$ based on sign conditions for components of $\boldsymbol{\theta}_i$ for $i = 1, \ldots, I$ with correctly estimated $\bar{\boldsymbol{\theta}}$: We set values for components of $\boldsymbol{G_\theta}$ so that the corresponding components of $\boldsymbol{\theta}_i$ for $i = 1, \ldots, I$ have expected signs. Third, we can also obtain appropriate priors of σ_d^2 and σ_s^2 so that the orders of their theoretical prior means are smaller than the smallest orders of variances of observed product characteristics $p_j, x_{j1}, \ldots, x_{j,Q-1}$ and of observed cost characteristics $z_{j1}, \ldots, z_{jS}$ respectively. To obtain such priors, we use an assumption that we observe all of the product and cost characteristics that are largely influential on consumers' utilities and products' prices respectively.

3 Empirical study

As an example of using Yonetani et al.'s (2010) method, we bring the method to data from the U.S. automobile market. Specifically, we will first estimate model parameters, using the 1995 market data. Then we will compare predicted and observed market shares for the 1996 market, using the estimated posteriors. Since vehicle is a differenciated product, it is appropriate to use the assumption that a consumer chooses one unit of product with the highest utility to derive the logistic choice probability (1). We choose the years of 1995 and 1996 because of our use of an estimated 8-year total cost of ownership (TCO) used by Myojo (2007) and of consumer preference stability. We think it more reasonable to use the TCO instead of the manufacturer suggested retail price as a consumer's expense of the endogenous price for owing his/her vehicle. In the years of 1995 and 1996, consumer preferences would be stable between the introductions of minivan in the mid 1980s and of hybrid electronic vehicles in the late 1990s.

3.1 Data

Our observed data for M, $\boldsymbol{v}^o$, $\boldsymbol{p}$, $\boldsymbol{X}$, $\boldsymbol{\delta}$ and $\boldsymbol{Z}$ are obtained from several public sources. In what follows, we particularly explain how to obtain the 1995 data. The 1996 data are obtained in the similar way.

We assume the market size M to be

$$\begin{aligned} M &= \text{(the number of vehicles per household)} \times \text{(the number of households)} \\ &\quad /\text{(median age of vehicles on the road)}, \end{aligned}$$

where the first two on the right hand side are from *Consumer Expenditure Survey 1995* (henceforth *CEX 1995*) and the last one from *Ward's Motor Vehicle Facts & Figures 1999.* We obtain $\boldsymbol{v}^o$ from *Ward's Automotive Yearbook 1996.* We use the top 50 models in sales (40 are from U.S. manufacturers and 10 are from Japanese manufacturers). The 50 models occupy 70.48% and 37.27% of M and the whole sales for the 160 models for which data for all of the remaining observed variables are available.

For $\boldsymbol{p}$, we use the TCO during 1995 and 2002, of which the maintenance and repair cost calculation was proposed by Puripunyavanich et al. (2004). For $\boldsymbol{X}$, we use a measure of acceleration (horsepower/weight), size (length$\times$width), Japan and U.S. dummies indicating the country of origin of manufacturers, minivan, pickup truck and SUV dummies and a measure of safety (a dummy indicating whether dual air bags are available standard or optional). Data for $\boldsymbol{X}$ and $\boldsymbol{\delta}$ are from *Ward's Automotive Yearbook 1995.* For $\boldsymbol{Z}$, we use an intercept, mileage, reliability and a scale variable of the logarithm of $\boldsymbol{v}^o$ in addition to $\boldsymbol{X}$ except for U.S. dummy.[2] The measure of reliability is the predicted reliability with five-point scale from *Consumer Reports April Annual Auto Issue 1996.* Note that we exclude mileage and reliability from $\boldsymbol{X}$ because they are included to calculate TCO (See Puripunyavanich et al., 2004). Notice that $Q = 9$ and $S = 11$, of which Q will be changed to 8.

For $\boldsymbol{y}$, we use data on the website of *Integrated Public Use Microdata Series - Current Population Survey 1995.* Corresponding our estimated eight-year TCO, we multiply the "Total Household Incomes" by eight years. Then we randomly draw I = 1,000 households in 1995 with their eight-year "Total Household Incomes" greater than the largest TCO.

3.2 MCMC estimation with the 1995 data

Hyperparameter values we set are

$$\begin{aligned}
\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}} &= (\mu_{\bar{\alpha}}, \mu_{\bar{\beta}_{size}}, \mu_{\bar{\beta}_{safety}}, \mu_{\bar{\beta}_{minivan}}, \mu_{\bar{\beta}_{pickup}}, \mu_{\bar{\beta}_{SUV}}, \mu_{\bar{\beta}_{Japan}}, \mu_{\bar{\beta}_{U.S.}})' \\
&= (54.48, 2.88, 0.27, 0.57, 0.093, 1.39, -4.57, -4.67)', \\
\boldsymbol{V}_{\bar{\boldsymbol{\theta}}} &= \mathrm{diag}(V_{\bar{\alpha}}, V_{\bar{\beta}_{size}}, V_{\bar{\beta}_{safety}}, V_{\bar{\beta}_{minivan}}, V_{\bar{\beta}_{pickup}}, V_{\bar{\beta}_{SUV}}, V_{\bar{\beta}_{Japan}}, V_{\bar{\beta}_{U.S.}}) \\
&= \mathrm{diag}(4.26, 0.13, 0.022, 0.038, 0.037, 0.039, 0.20, 0.25), \\
g_{\boldsymbol{\theta}} &= 12, \\
\boldsymbol{G}_{\boldsymbol{\theta}} &= \mathrm{diag}(G_{\alpha}, G_{\beta_{size}}, G_{\beta_{safety}}, G_{\beta_{minivan}}, G_{\beta_{pickup}}, G_{\beta_{SUV}}, G_{\beta_{Japan}}, G_{\beta_{U.S.}}) \\
&= \mathrm{diag}(15, 0.99, 0.0091, 2.51, 2.62, 0.040, 0.11, 0.23), \\
\bar{\boldsymbol{\gamma}} &= (0, \dots, 0)', \quad \boldsymbol{V}_{\boldsymbol{\gamma}} = 10^2 \boldsymbol{E}_{11}, \quad g_d = 5, \quad G_d = 0.0003, \quad g_s = 5, \quad G_s = 0.03.
\end{aligned}$$

We set $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$ and the diagonal components of $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$ based on a pre-analytical procedure in Appendix B where we exclude acceleration from $\boldsymbol{X}$, that is, we reset $Q = 8$. Given $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$ and $g_{\boldsymbol{\theta}}$, we set $\boldsymbol{G}_{\boldsymbol{\theta}}$ except for $G_{\beta_{pickup}}$ so that the corresponding components of $\boldsymbol{\theta}_i$ for $i = 1, \dots, I$ have the same signs as those of $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$ have; and we set $G_{\beta_{pickup}}$ so that the variance of $\beta_{i,pickup}$ is equal to $V_{\bar{\beta}_{pickup}}$. Given g_d and g_s, we set G_d and G_s based on the assumption that we included all of the largely influential product and cost characteristics in $\boldsymbol{X}$ and $\boldsymbol{Z}$ respectively. The $g_{\boldsymbol{\theta}}$, g_d and g_s are the smallest values to define their theoretical prior variance-covariance matrix or variances. The $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, g_d and G_d avoid the nonpositive cost problem, and g_s and G_s in addition to $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, g_d and G_d avoid extremely large or small values for $R_{\boldsymbol{\xi}^*}$ and $R_{\boldsymbol{\theta}^*}$.

[2] We use a concept of cost shifter by which we use observed product characteristics as alternatives to actual observed cost characteristics.

We run three independent sequences for the MCMC algorithm in Appendix A, each of which has T = 50,000 iterations with a different set of initial parameter values. As for initial parameter values, we set $\boldsymbol{\beta}^{(0)} = \mathbf{0}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)} = \boldsymbol{E}_8$, $\sigma_d^{2(0)} = 10^{-10}$ and $\sigma_s^{2(0)} = 1$ for all of the three MCMC sequences. The $\bar{\alpha}^{(0)}$ and $\boldsymbol{\gamma}^{(0)}$ has three sets as the large, middle and small. One of the three sequences has only the large set and another one has only the small set. The remaining seqeuence has the middle set with a uniformly generated random value for each parameter with the upper and lower bounds corresponding to values in the forementioned large and small sets respectively. The large set has $\bar{\alpha}^{(0)} = 60$ and $\boldsymbol{\gamma}^{(0)} = (5, \ldots, 5)'$. The small set has $\bar{\alpha}^{(0)} = 50$ and $\boldsymbol{\gamma}^{(0)} = (-5, \ldots, -5)'$. Given $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$ and $\sigma_d^{2(0)}$, we generate $\boldsymbol{\xi}^{(0)} = (\xi_1^{(0)}, \ldots, \xi_{10}^{(0)})'$ from $N(0, \sigma_d^{2(0)})$ and $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, \ldots, \boldsymbol{\theta}_{1,000}^{(0)})$ from $MVN(\bar{\boldsymbol{\theta}}^{(0)}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)})$ in each sequence. These initial parameter values are designed to avoid not only the nonpositive cost problem but also the computational zero likelihood problem.

We summarize the results in Table 1, using the last halves of draws in the three MCMC sequences. We checked the convergence of the MCMC algorithm, using time-series plots for all of the parameters in Figures 1 through 3. The means of $R_{\boldsymbol{\xi}^*}$ and $R_{\boldsymbol{\theta}^*}$ are 87.64% and 68.26% respectively.

According to the 95% posterior intervals for $\bar{\boldsymbol{\theta}}$, consumers' utility was in average enhanced by size, dual air bags, minivans and SUVs while it was reduced by TCO. The negative signs for Japan and U.S. dummies were measured against the outside good with the highest market share of 62.73 among $j = 0, \ldots, 50$. Our results for $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ indicated there existed consumers' heterogeneity for all of the product characteristics. Specifically, the results indicated that some consumers prefer pickup trucks while others not, given the fact that the 95% posterior interval for $\bar{\beta}_{pickup}$ included zero. Our results for $\boldsymbol{\gamma}$ indicated that size, dual air bags took more cost; and minivans and SUVs took more cost while pickup trucks took less cost 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.

Table 1: Posterior means, standard deviations and quantiles (2.5%, 5%, 50%, 95% and 97.5%)

Parameter	Mean	Std.Dev.	2.5%	5%	50%	95%	97.5%
$\bar{\alpha}$	53.47	2.43	48.92	49.40	53.42	57.53	57.87
$\bar{\beta}_{size}$	2.81	0.20	2.42	2.49	2.80	3.15	3.22
$\bar{\beta}_{safety}$	0.25	0.11	0.028	0.047	0.27	0.41	0.42
$\bar{\beta}_{minivan}$	0.54	0.11	0.37	0.39	0.53	0.74	0.83
$\bar{\beta}_{pickup}$	0.076	0.13	−0.20	−0.15	0.088	0.27	0.29
$\bar{\beta}_{SUV}$	1.38	0.13	1.10	1.14	1.39	1.60	1.63
$\bar{\beta}_{Japan}$	−4.80	0.30	−5.45	−5.31	−4.78	−4.34	−4.28
$\bar{\beta}_{U.S.}$	−4.65	0.26	−5.18	−5.11	−4.64	−4.24	−4.17
σ^2_{α}	4.38	3.29	1.12	1.32	3.46	10.91	13.77
$\sigma^2_{\beta_{size}}$	0.27	0.28	0.074	0.084	0.19	0.67	1.22
$\sigma^2_{\beta_{safety}}$	0.0024	0.0019	0.00066	0.00075	0.0019	0.0061	0.0078
$\sigma^2_{\beta_{minivan}}$	0.011	0.0089	0.0030	0.0036	0.0085	0.025	0.031
$\sigma^2_{\beta_{pickup}}$	0.034	0.028	0.0088	0.010	0.025	0.086	0.11
$\sigma^2_{\beta_{SUV}}$	0.064	0.049	0.017	0.020	0.050	0.16	0.20
$\sigma^2_{\beta_{Japan}}$	0.58	0.34	0.20	0.23	0.48	1.26	1.48
$\sigma^2_{\beta_{U.S.}}$	0.88	1.04	0.20	0.23	0.55	2.71	4.47
$\gamma_{intercept}$	−1.25	0.77	−2.75	−2.50	−1.25	0.012	0.27
$\gamma_{hp/weight}$	0.016	0.032	−0.047	−0.036	0.016	0.069	0.079
γ_{size}	0.62	0.25	0.13	0.21	0.62	1.02	1.10
$\gamma_{mileage}$	−0.044	0.011	−0.065	−0.062	−0.044	−0.027	−0.023
$\gamma_{reliability}$	0.023	0.030	−0.036	−0.026	0.024	0.073	0.083
γ_{safety}	0.16	0.062	0.033	0.054	0.16	0.26	0.28
$\gamma_{minivan}$	0.16	0.097	−0.028	0.0041	0.16	0.32	0.36
γ_{pickup}	−0.20	0.10	−0.41	−0.37	−0.20	−0.031	0.0035
γ_{SUV}	0.32	0.13	0.069	0.11	0.32	0.53	0.57
γ_{Japan}	0.11	0.093	−0.074	−0.045	0.11	0.26	0.29
$\gamma_{\ln \boldsymbol{v}^o}$	−0.089	0.057	−0.20	−0.18	−0.089	0.0040	0.023
σ^2_d	0.000098	0.00011	0.000023	0.000027	0.000067	0.00026	0.00036
σ^2_s	0.025	0.0057	0.016	0.017	0.024	0.036	0.039

Figure 1: Time series plots for $\bar{\boldsymbol{\theta}}$.

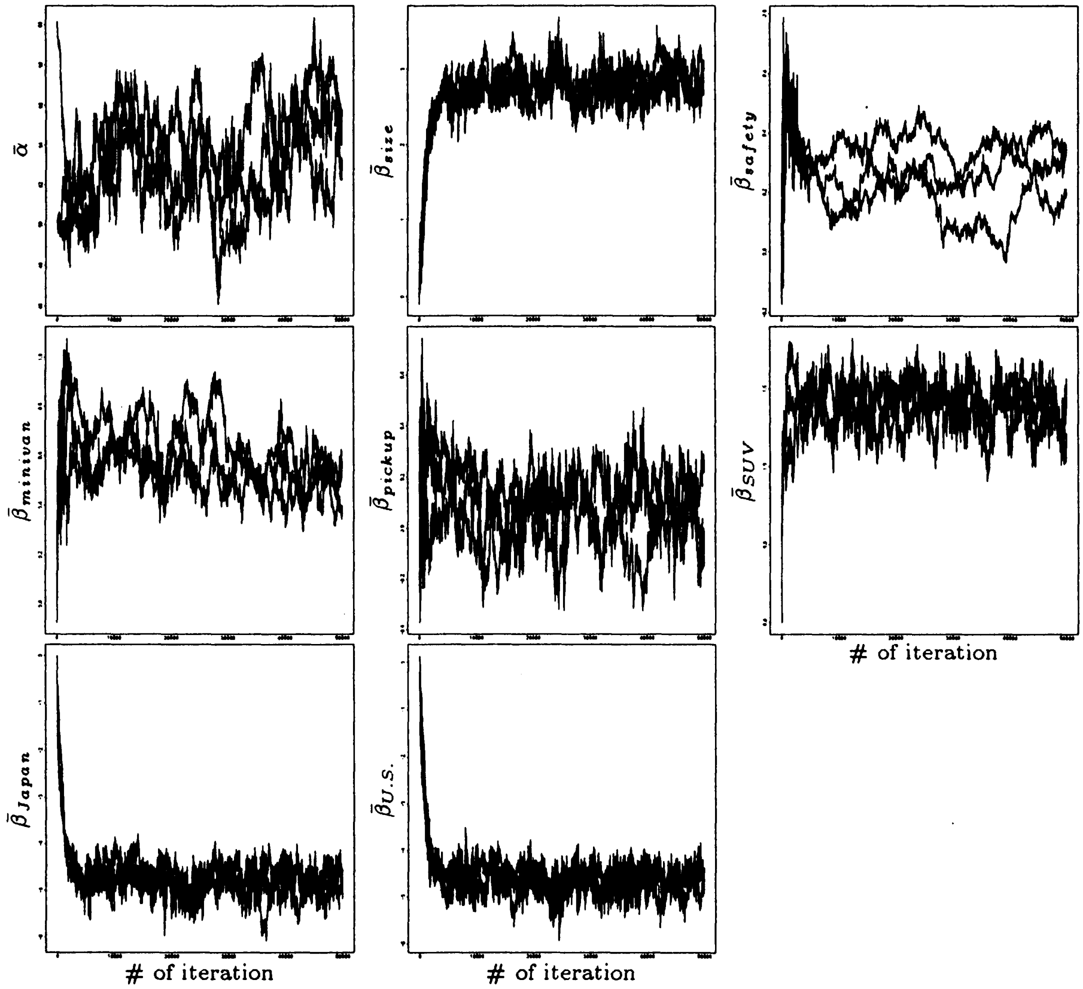


Figure 2: Time series plots for the diagonal components of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$.

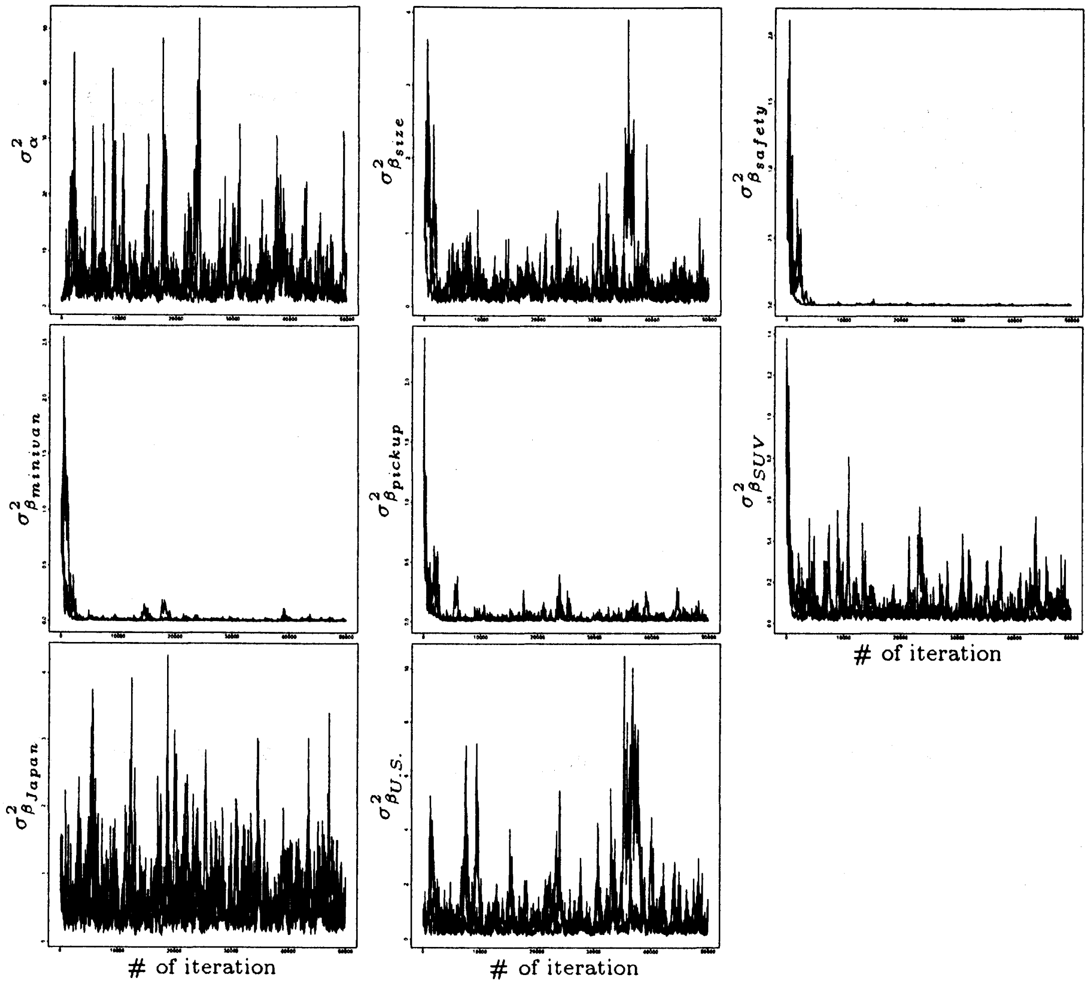
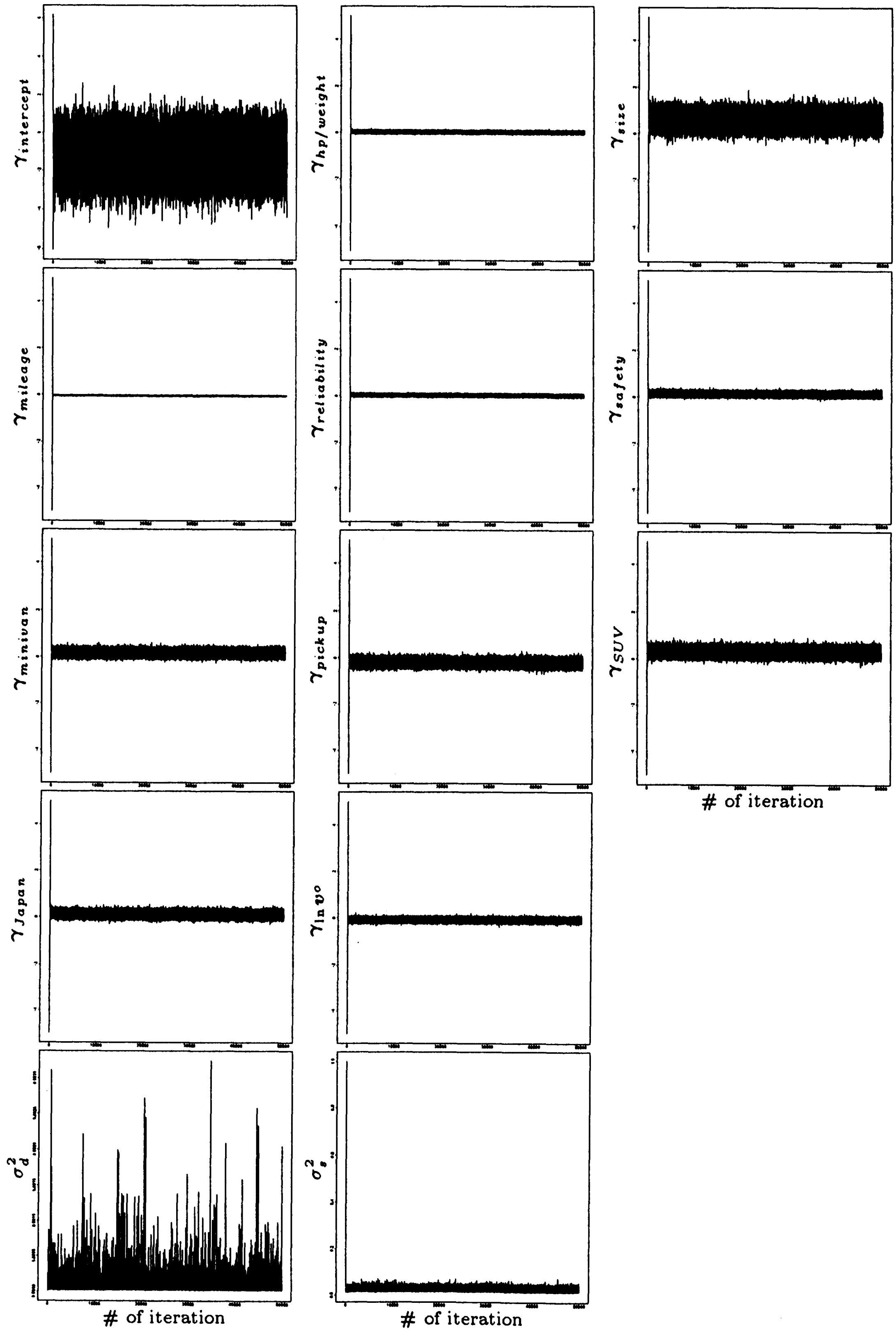


Figure 3: Time series plots for γ, σ_d^2 and σ_s^2.



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

We compare observed market shares for the top 50 models in sales in the 1996 market with their predicted market shares from the last halves of draws for $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ and observed data of their TCOs and $\boldsymbol{X}$. Note that the last halves of draws for $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ are posterior draws for them. Specifically, we randomly obtain 300 sets of draws for $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ and calculate the predicted market shares for $j = 0, \dots, 50$ in terms of each set of the obtained $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$. Our predicted market shares are the means of the 300 sets of the market shares for $j = 0, \dots, 50$.

The predicted and observed market shares for $j = 0, \dots, 50$ are presented in Figure 4. 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 70.70% while the mean of the absolute deviations was 0.0045.

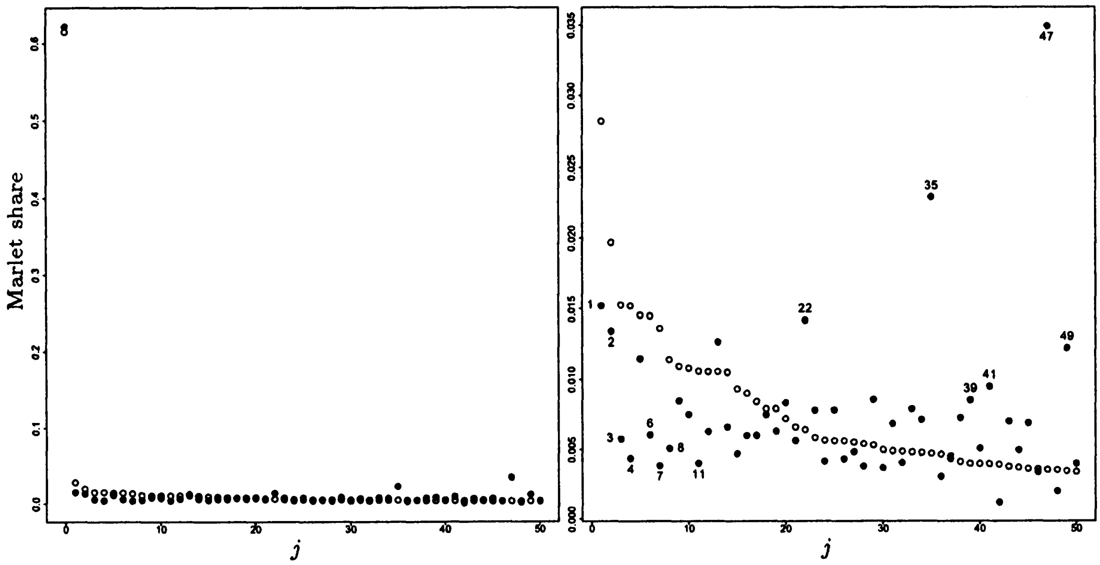
Except for the 14 largely over/underestimated models, our Bayesian estimates predicted the observed market shares well for the other 36 models. Note that the underestimated 8 models had been acheiving top ranks of sales at least in the past 5 years. Therefore, we could fail to capture the reputation 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 $\boldsymbol{\xi}$. We will discuss the limitation in Section 4.

4 Conclusion and discussion

In this paper, we reviewed a Bayesian estimation method for a simultaneous demand and supply model for market-level data which was proposed by Yonetani et al. (2010) and then showed an example of using it. In the example, we concluded that it could be difficult to predict observed market shares for some products due to our limitation of the posterior estimation method for $\boldsymbol{\xi}$. Specifically, we could fail to caputure an unobserved product characteristic of reputation for some popular vehicles by $\boldsymbol{\xi}$.

We have two ideas to improve it. The first idea is to estimate the components of $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{50})'$ individually to obtain each model-specific value more precisely. Note that we estimated $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{50})'$ all at once in **MCMC1** through **MCMC3** in the MCMC algorithm in Appendix A to reduce computational burden. However, it could be more difficult to estimate a model-specfic

Figure 4: 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
3. Ford Explorer
4. Ford Taurus
6. Honda Accord
7. Toyota Camry
8. Dodge Caravan
11. Jeep Grand Cherokee
22. GMC Sierra C/K
35. Chevrolet Tahoe
39. Plymouth Neon
41. Dodge Dakota
47. Chevrolet Suburban
49. Jeep Wrangler

value for ξ_j for $j = 1, \dots, 50$ efficiently when we estimated $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{50})'$ all at once. If we improved our MCMC algorithm to estimate $\boldsymbol{\xi} = (\xi_1, \dots, \xi_{50})'$ individually, we would increase computational burden when we use the 50 or more models.

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. However, we also have to pay attention to keeping appropriate values for $R_{\boldsymbol{\xi}^*}$ if we retain to estimate $\boldsymbol{\xi}$ all at once.

In future, we need more empirical studies in other differentiated product markets, using the proposed method. To predict consumers' purchasing behavior from market-level sales volume data is useful because it is more difficult or costs more to obtain consumer-level purchase incidence data. We believe that these empirical studies will be an important contribution to the literature of consumers' purchasing behavior.

A MCMC algorithm

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

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

For $t = 1, \dots,$

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

MCMC2 Calculate

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

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

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

MCMC5 Calculate

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

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

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

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

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

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

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

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

B A pre-analytical procedure

We explain how we obtain the values for $\mu_{\bar{\theta}}$ and the diagonal components of $V_{\bar{\theta}}$ in Section 3, using a pre-analytical MCMC. In the pre-analytical MCMC, we assume that there is no consumer heterogeneity, that is, $\theta_i = \bar{\theta}$ for $i = 1, \ldots, 1{,}000$. The pre-analytical MCMC includes the original **MCMC0** through

MCMC3, **MCMC9** through **MCMC12** and an additional random walk Metropolis-Hastings algorithm corresponding to the first method in Chib and Greenberg (1995) to generate $\bar{\boldsymbol{\theta}}$. There are two points to be noted. First, the pre-analytical MCMC no longer requires **MCMC4** through **MCMC8** to generate $\boldsymbol{\theta}$, $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$. Second, we use a proposal distribution of the multivariate normal distribution with the mean vector of the current $\bar{\boldsymbol{\theta}}^{(t-1)}$ and the variance-covariance matrix of diag$(1, 0.01, \ldots, 0.01)$ in the additional Metropolis-Hastings algorithm to generate $\bar{\boldsymbol{\theta}}$.

To obtain the values for $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$ and the diagonal components of $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, we will have two steps. We will first obtain tentative values for $\mu_{\bar{\alpha}}$ and $V_{\bar{\alpha}}$. Given the tentative $\mu_{\bar{\alpha}}$ and $V_{\bar{\alpha}}$, we will then obtain values of $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$ and the diagonal components of $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$.

In the first step, we set $Q = 1$ and run the pre-analytical MCMC with $T = 2{,}000$. For the hyperparameters, we set $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}} = \mu_{\bar{\alpha}} = 20$ and $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}} = V_{\bar{\alpha}} = 100$ to obtain a so-called diffuse prior for $\bar{\boldsymbol{\theta}} = \bar{\alpha}$; and set the same hyperparameter values for $\bar{\boldsymbol{\gamma}}$, $\boldsymbol{V}_{\boldsymbol{\gamma}}$, g_d, G_d, g_s and G_s as those in Section 3. Note that we no longer need $g_{\boldsymbol{\theta}}$ and $\boldsymbol{G}_{\boldsymbol{\theta}}$ for the prior of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$. The number of the pre-analytical MCMC sequences and the initial parameter values required except for $\bar{\alpha}^{(0)}$ are the same as those in Section 3. As for the initial parameter values for $\bar{\alpha}^{(0)}$, we set 45 and 65 for the two pre-analytical MCMC sequences with the large and small sets of the initial parameter values respectively; and we set a uniform random number from $U(45, 65)$ for $\bar{\alpha}^{(0)}$ for the remaining one pre-analytical MCMC sequence.

In the second step, we include $\boldsymbol{X}$ as well as $\boldsymbol{p}$ in the model and thus $Q = 9$. Then the number of iterations for the pre-analytical MCMC in the second step is 20,000. For the hyperparameters, we set[3]

$$\begin{aligned}
\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}} &= (\mu_{\bar{\alpha}}, \mu_{\bar{\beta}_{hp/weight}}, \mu_{\bar{\beta}_{size}}, \mu_{\bar{\beta}_{safety}}, \mu_{\bar{\beta}_{minivan}}, \mu_{\bar{\beta}_{pickup}}, \mu_{\bar{\beta}_{SUV}}, \mu_{\bar{\beta}_{Japan}}, \mu_{\bar{\beta}_{U.S.}})' \\
&= (56.15, 0, \ldots, 0)', \\
\boldsymbol{V}_{\bar{\boldsymbol{\theta}}} &= \mathrm{diag}(V_{\bar{\alpha}}, V_{\bar{\beta}_{hp/weight}}, V_{\bar{\beta}_{size}}, V_{\bar{\beta}_{safety}}, V_{\bar{\beta}_{minivan}}, V_{\bar{\beta}_{pickup}}, V_{\bar{\beta}_{SUV}}, V_{\bar{\beta}_{Japan}}, V_{\bar{\beta}_{U.S.}}) \\
&= \mathrm{diag}(4.33, 100, \ldots, 100)
\end{aligned}$$

for the prior of $\bar{\boldsymbol{\theta}}$; and set the same hyperparameter values for $\bar{\boldsymbol{\gamma}}$, $\boldsymbol{V}_{\boldsymbol{\gamma}}$, g_d, G_d, g_s and G_s as those in Section 3. Note that $\mu_{\bar{\alpha}} = 56.15$ and $V_{\bar{\alpha}} = 4.33$ are the posterior mean and variance of $\bar{\boldsymbol{\theta}}$ from the last halves of draws in the first pre-analytical MCMC sequences. The number of the second pre-analytical MCMC sequences and the initial parameter values required except for $\bar{\alpha}^{(0)}$ for

[3]The notation "*hp*" stands for horsepower.

the second pre-analytical MCMC are the same as those in Section 3. We use the same settings for the initial parameter values for $\bar{\alpha}^{(0)}$ for the second step as those in the first step.

The results indicated that the 95% posterior interval for $\bar{\beta}_{hp/weight}$ was below zero although it was expected to be above zero. We thus remove acceleration from $\boldsymbol{X}$ and re-run the pre-analytical MCMC with the same settings as those in the last pre-analytical MCMC. Notice $Q = 8$. The values for $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$ and the diagonal components of $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$ in Section 3 are the corresponding posterior means and variances of $\bar{\boldsymbol{\theta}}$ from the last halves of draws in the last pre-analytical MCMC sequences.

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.

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/1995/Standard/cusize.pdf`. 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.

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.

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

Jiang, R., P. Manchanda, P. E. Rossi. 2008. Bayesian analysis of random coefficient logit models using aggregate data. Working paper, University of Chicago.

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.

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.

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.

Yonetani, Y., Y. Kanazawa, S. Myojo, S. J. Turnbull. 2010. A bayesian simultaneous demand and supply model for aggregate data in a differentiated product market. Working paper, University of Tsukuba.