数理解析研究所講究録
第 1621 巻 2009 年 42-64

A simulation study of Bayesian estimation with diffuse priors on simultaneous demand and supply with market-level data

Yutaka YONETANI[1] and Yuichiro KANAZAWA[2]

1 Introduction

Suppose we wish to investigate what motivates consumers to purchase a certain good over others offered in a market. Marketers and economists usually frame these purchasing behaviors in terms of consumers' maximizing their utilities. For some goods, notably agricultural products such as corn, soybeans and wheat, the only differentiating characteristic is often price. On the other hand, many industrial durable goods such as automobiles have many differentiating characteristics. We call the market of these goods a differentiated product market. As a consumer, your utility is higher for products with a lot of desirable product characteristics, but you are expected to pay a premium for such characteristics. This can be incorporated into utility with the price coefficient having a negative sign while other characteristics coefficients taking positive signs. Analysis, however, can improve if we

[1]Yutaka YONETANI is a doctoral candidate at the 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.

[2]Yuichiro KANAZAWA is Professor of Statistics at the 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 kanazawa@sk.tsukuba.ac.jp. This research is supported in part by the Grant-in-Aid for Scientific Research (C)(2)16510103, (B)19330081 and (B)20310081 from the Japan Society for the Promotion of Science.

incorporate suppliers into the equation. Modern marketing and economic demand analyses, therefore, often model a demand side as well as a supply side simultaneously. This is sometimes called by marketers and economists as price endogeneity.

Consumers are in general very heterogenous in terms of income, education, ethnicity, other attributes as well as tastes. As a result, their utilities vary widely and this variabilities are transmitted to differing purchasing patterns or differing utility coefficients. This is often referred by marketers and economists as consumer heterogeneity. We have to account for the price endogeneity as well as consumer heterogeneity when we model consumers' purchasing behaviors in a differentiated product markets.

In some markets, we have access to a detailed individual purchasing history from, for instance, POS (point-of-sale) scanning data. In other markets—the market of differentiated products being the one—only products' market shares and possibly overall market sizes are available. We call the former consumer-level data while the latter is usually classified as market level-data.

Yonetani et al. (2007) proposed a Bayesian simultaneous demand and supply model with consumers' heterogeneity for market-level data. Then Yonetani et al. (2008) examined the validity of the same model through a simulation study only with non-diffuse priors and the small number of parameters. Additionally, we sometimes encounter problems such as non-positive product cost and very long time to convergence in their model.

The purpose of this paper is two-fold. First, we examine causes of the problems in Yonetani et al's (2007) model. Second, we implement simulations for their model with diffuse priors and the large number of parameters to find out its applicability and to give practical suggestions. This paper is organized as follows. In Sections 2 and 3, we briefly review Yonetani et al.'s (2007) model and estimation method respectively (See Yonetani et al. (2008) for more specific explanations). Section 4 examines the problems. Section 5 contains the simulation study. Summaries are presented in Section 6.

2 Model specification

2.1 Demand Model

We assume that there are J products in a market of a differentiated durable product where a consumer purchases one unit of a product. Let us 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 $j = 0$ being the outside good.

Each consumer i has his/her utility for product j as

$$u_{ij} = u_{ij}(p_j, \boldsymbol{x}_j, \xi_j, y_i, \boldsymbol{\theta}_i, \varepsilon_{ij}) = \alpha_i \log(y_i - p_j) + \boldsymbol{x}_j \boldsymbol{\beta}_i + \xi_j + \varepsilon_{ij}, \qquad (2.1)$$

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. For $j = 0$, we assume $p_0 = 0$, $\boldsymbol{x}_0 = \boldsymbol{0}$ and $\xi_0 = 0$.

In (2.1), we assume that ε_{ij} is independent of the other terms and independently and identically Gumbel (type I extreme value) distributed across consumers and products. Then we derive a consumer i's logit choice probability for product j as

$$s_{ij} = s_{ij}(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}, y_i, \boldsymbol{\theta}_i) = \frac{\exp\{\alpha_i \log(y_i - p_j) + \boldsymbol{x}_j \boldsymbol{\beta}_i + \xi_j\}}{\sum_{k=0}^{J} \exp\{\alpha_i \log(y_i - p_k) + \boldsymbol{x}_k \boldsymbol{\beta}_i + \xi_k\}}, \qquad (2.2)$$

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

The market share of product j in I sample consumers is

$$s_j = s_j(\boldsymbol{p}) = s_j(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{y}, \boldsymbol{\theta}) = \frac{1}{I}\sum_{i=1}^{I} s_{ij}, \tag{2.3}$$

where $\boldsymbol{y} = (y_1, \ldots, y_I)'$ and $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$. We denote $\boldsymbol{s}$ as a $J \times 1$ market share vector for product $j = 1, \ldots, J$:

$$\boldsymbol{s} = \boldsymbol{s}(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{y}, \boldsymbol{\theta}) = (s_1, \ldots, s_J)'. \tag{2.4}$$

We also denote $\boldsymbol{v} = (v_1, \ldots, v_J)'$ as a $J \times 1$ sales volume vector for product $j = 1, \ldots, J$ in the I consumers where we define

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

Note that int$(\cdot)$ is the integral part in the expression $(\cdot)$. The number of consumers for $j = 0$ in the I consumers is thus $v_0 = I - \sum_{j=1}^{J} v_j$.

2.2 Supply Model

We assume that fixed F firms are in an oligopolistic market of the J products with Bertrand competition. We also assume that each firm f produces a subset of the J products and sets prices for its products to maximize its total profit

$$\Pi_f = \sum_{j \in f} M s_j(\boldsymbol{p})(p_j - c_j), \tag{2.5}$$

where c_j is a unit cost. The Bertrand competition leads to the first order condition for $j = 1, \ldots, J$ from (2.5) as

$$\boldsymbol{p} = -\left\{\left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}}\right)'\right\}^{-1} \boldsymbol{s} + \boldsymbol{c}, \tag{2.6}$$

assuming the inverse above exists. Note $\boldsymbol{c} = (c_1, \ldots, c_J)'$ and $(\partial \boldsymbol{G}/\partial \boldsymbol{p}) = (\partial \boldsymbol{s}/\partial \boldsymbol{p}) * \boldsymbol{\delta}$ where the sign $*$ represents the element-by-element multiplication of the matrices it connects and the (j, k) element δ_{jk} of $\boldsymbol{\delta}$ is 1 if the products j and k are produced by the same firm and 0 otherwise.[3] As for the cost c_j, we assume

$$\log c_j = \boldsymbol{z}_j \boldsymbol{\gamma} + \eta_j, \tag{2.7}$$

where $\boldsymbol{z}_j$ and η_j are product j's $1 \times S$ cost shifter vector and unobserved cost respectively and $\boldsymbol{\gamma}$ is a $S \times 1$ coefficient vector.

Let us denote $\boldsymbol{Z} = (\boldsymbol{z}'_1, \ldots, \boldsymbol{z}'_J)'$ and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_J)'$. Substituting $\mathbf{exp}\{\boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\eta}\}$ for $\boldsymbol{c}$ in (2.6), we obtain the pricing equation

$$\mathbf{log}\left[\boldsymbol{p} + \left\{\left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}}\right)'\right\}^{-1} \boldsymbol{s}\right] = \boldsymbol{Z}\boldsymbol{\gamma} + \boldsymbol{\eta}. \tag{2.8}$$

We can also write $\boldsymbol{p}$ as

$$\boldsymbol{p} = \boldsymbol{p}(\boldsymbol{s}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{\delta}, \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\eta}, \boldsymbol{\gamma}). \tag{2.9}$$

3 Bayesian Estimation

3.1 Parameters and their prior distributions

Given the overall market size M, product j's market share s_j and sales volume v_j are the one-to-one correspondence for $j = 1, \ldots, J$. Therefore, we can rewrite the simultaneous demand and supply model from (2.4) and (2.9) as

$$\boldsymbol{v} | \boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{y}, \boldsymbol{\theta}, \qquad (2.4)'$$

$$\boldsymbol{p} | \boldsymbol{v}, \boldsymbol{X}, \boldsymbol{\xi}, \boldsymbol{\delta}, \boldsymbol{y}, \boldsymbol{\theta}, \boldsymbol{Z}, \boldsymbol{\eta}, \boldsymbol{\gamma}. \qquad (2.9)'$$

[3] The elements of $(\partial \boldsymbol{s}/\partial \boldsymbol{p})$ and $(\partial \boldsymbol{G}/\partial \boldsymbol{p})$ are specified in Yonetani et al. (2008).

In terms of unobserved product and cost characteristics $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, we assume

$$\boldsymbol{\xi}\,|\boldsymbol{\Sigma}_d \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_d), \tag{3.1}$$

$$\boldsymbol{\eta}\,|\boldsymbol{\Sigma}_s \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_s). \tag{3.2}$$

These assumptions extend the simultaneous demand and supply model as

$$\begin{aligned} &\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}, \boldsymbol{\theta}, && (2.4)' \\ &\boldsymbol{p}|\boldsymbol{v}, \boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\eta}, \boldsymbol{\gamma}, && (2.9)' \\ &\boldsymbol{\xi}|\boldsymbol{\Sigma}_d, && (3.1) \\ &\boldsymbol{\eta}|\boldsymbol{\Sigma}_s. && (3.2) \end{aligned}$$

Note that the exogenous $\boldsymbol{X}$, $\boldsymbol{y}$, $\boldsymbol{\delta}$ and $\boldsymbol{Z}$ are left out from (2.4)′ and (2.9)′ for notational simplicity.

We next hypothesize prior distributions for the parameters $\boldsymbol{\theta}$, $\boldsymbol{\gamma}$, $\boldsymbol{\Sigma}_d$ and $\boldsymbol{\Sigma}_s$. As for $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$, we introduce a hierarchical structure where the prior of $\boldsymbol{\theta}_i$ for $i = 1, \ldots, I$ is

$$\boldsymbol{\theta}_i|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \sim MVN(\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}) \tag{3.3}$$

and $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ are also treated as parameters with the priors[4]

$$\bar{\boldsymbol{\theta}} \sim MVN(\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}, \boldsymbol{V}_{\bar{\boldsymbol{\theta}}}), \;\; \boldsymbol{\Sigma}_{\boldsymbol{\theta}} \sim IW_{g_{\theta}}(\boldsymbol{G}_{\boldsymbol{\theta}}). \tag{3.4}$$

As for the remaining parameters, we assume

$$\boldsymbol{\gamma} \sim MVN(\bar{\boldsymbol{\gamma}}, \boldsymbol{V}_{\boldsymbol{\gamma}}), \;\; \boldsymbol{\Sigma}_d \sim IW_{g_d}(\boldsymbol{G}_d), \;\; \boldsymbol{\Sigma}_s \sim IW_{g_s}(\boldsymbol{G}_s). \tag{3.5}$$

---

[4]The Bayesian hierarchical estimation can complement the lack of information about $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$ of the I consumers. It can also take into account some posterior uncertainty for $(\boldsymbol{\theta}_{I+1}, \boldsymbol{\theta}_{I+2}, \ldots)$ of additional sample consumers.

3.2 Distributions of endogenous observed data

With $\boldsymbol{s} = (s_1, \ldots, s_J)'$ in (2.4), we obtain a multinomial distribution for $\boldsymbol{v} = (v_1, \ldots, v_J)'$ as

$$f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}, \boldsymbol{\theta}) = \frac{I!}{v_0! \cdots v_J!} s_0^{v_0} \cdots s_J^{v_J}. \tag{3.6}$$

Since the pricing equation (2.8) is implicit in $\boldsymbol{p}$, we solve it with respect to $\boldsymbol{\eta}$ and then apply the variable transformation formula with $\boldsymbol{\eta} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_s)$ in (3.2) to obtain the distribution of $\boldsymbol{p}$,[5]

$$\begin{aligned} &f(\boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s) \\ &= (2\pi)^{-\frac{J}{2}} |\boldsymbol{\Sigma}_s|^{-\frac{1}{2}} \left\| \left(\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{p}} \right) \right\| \\ &\times \exp\left[-\frac{1}{2} \left[\log \left[\boldsymbol{p} + \left\{ \left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}} \right)' \right\}^{-1} \boldsymbol{s} \right] - \boldsymbol{Z\gamma} \right]' \boldsymbol{\Sigma}_s^{-1} \left[\log \left[\boldsymbol{p} + \left\{ \left(\frac{\partial \boldsymbol{G}}{\partial \boldsymbol{p}} \right)' \right\}^{-1} \boldsymbol{s} \right] - \boldsymbol{Z\gamma} \right] \right]. \end{aligned} \tag{3.7}$$

3.3 The joint posterior of the parameters

The distributions so far lead to

$$\begin{aligned} f(\boldsymbol{\xi}, \boldsymbol{\theta}, \bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_d, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s|\boldsymbol{v}, \boldsymbol{p}) \propto & f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}, \boldsymbol{\theta}) f(\boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s) \\ & \times f(\boldsymbol{\xi}|\boldsymbol{\Sigma}_d) \left[\prod_{i=1}^{I} f(\boldsymbol{\theta}_i|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}) \right] \\ & \times f(\bar{\boldsymbol{\theta}}) f(\boldsymbol{\Sigma}_{\boldsymbol{\theta}}) f(\boldsymbol{\Sigma}_d) f(\boldsymbol{\gamma}) f(\boldsymbol{\Sigma}_s) \end{aligned}$$

from which we obtain the joint posterior of the parameters as

$$f(\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_d, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s|\boldsymbol{v}, \boldsymbol{p}) = \int f(\boldsymbol{\xi}, \boldsymbol{\theta}, \bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_d, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s|\boldsymbol{v}, \boldsymbol{p}) d\boldsymbol{\xi}. \tag{3.8}$$

Since it is difficult to solve the integral in (3.8) analytically, we numerically obtain the joint posterior as follows. First, we apply the data augmentation

[5]The elements of $(\partial \boldsymbol{\eta}/\partial \boldsymbol{p})$ are specified in Yonetani et al. (2008).

technique (Tanner & Wong, 1987) to the equation (3.8). Let us denote $\boldsymbol{\psi} = (\boldsymbol{\theta}, \bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_d, \gamma, \boldsymbol{\Sigma}_s)$. The equation (3.8) can be rewritten so that the joint posteior $f(\boldsymbol{\psi}|\boldsymbol{v},\boldsymbol{p})$ appears on both sides as

$$f(\boldsymbol{\psi}|\boldsymbol{v},\boldsymbol{p}) = \int f(\boldsymbol{\psi}|\boldsymbol{\xi},\boldsymbol{v},\boldsymbol{p}) f(\boldsymbol{\xi}|\boldsymbol{v},\boldsymbol{p}) d\boldsymbol{\xi} \tag{3.9}$$

$$= \int f(\boldsymbol{\psi}|\boldsymbol{\xi},\boldsymbol{v},\boldsymbol{p}) \left[\int f(\boldsymbol{\xi}|\boldsymbol{\psi},\boldsymbol{v},\boldsymbol{p}) f(\boldsymbol{\psi}|\boldsymbol{v},\boldsymbol{p}) d\boldsymbol{\psi} \right] d\boldsymbol{\xi}. \tag{3.10}$$

The equation (3.10) suggests an iterative process:

Step A In the brackets, we generate $\boldsymbol{\psi}_l$ from $f(\boldsymbol{\psi}|\boldsymbol{v},\boldsymbol{p})$ and then generate $\boldsymbol{\xi}_l$ from $f(\boldsymbol{\xi}|\boldsymbol{\psi}_l,\boldsymbol{v},\boldsymbol{p})$ to obtain $\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_L$.[6]

Step B We calculate a Monte Carlo estimator of $f(\boldsymbol{\psi}|\boldsymbol{v},\boldsymbol{p})$ as $\sum_{l=1}^{L} f(\boldsymbol{\psi}|\boldsymbol{\xi}_l,\boldsymbol{v},\boldsymbol{p})/L$ from which we generate $\boldsymbol{\psi}_l$ in **Step A**.

Second, we set $L = 1$. Then we no longer need **Step B** and rewrite **Step A** as

Step A In the brackets, we generate $\boldsymbol{\psi}$ from $f(\boldsymbol{\psi}|\boldsymbol{\xi},\boldsymbol{v},\boldsymbol{p})$ and then generate $\boldsymbol{\xi}$ from $f(\boldsymbol{\xi}|\boldsymbol{\psi},\boldsymbol{v},\boldsymbol{p})$.

We apply the Gibbs sampler to a nonstandard parametric $f(\boldsymbol{\psi}|\boldsymbol{\xi},\boldsymbol{v},\boldsymbol{p})$. In the Gibbs sampler, we further apply the Metropolis-Hastings algorithm to the conditional posterior of $\boldsymbol{\theta}$ which has a nonstandard parametric form.[7][8]

[6]In other words, we apply the composition method to the integral in the brackets in (3.10) to generate $\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_L$ from $f(\boldsymbol{\xi}|\boldsymbol{v},\boldsymbol{p})$ in (3.9).

[7]Note that we further apply the Gibbs sampler to the conditional posterior of $\boldsymbol{\theta}$ and then apply the Metropolis-Hastings algorithm to the conditional posterior of $\boldsymbol{\theta}_i$ for $i = 1,\ldots,I$ in the MCMC algorithm in Yonetani et al. (2008). In this paper, we directly apply the Metropolis-Hastings algorithm to the conditional posterior of $\boldsymbol{\theta}$ to reduce the computation time.

[8]As the Metropolis-Hastings algorithm, we employ the third method in Chibs & Greenberg (1995).

We can generate draws of the other parameters from their standard parametric posteriors. On the other hand, we also apply the Metropolis-Hastings algorithm to a conditional posterior $f(\boldsymbol{\xi}|\boldsymbol{\psi}, \boldsymbol{v}, \boldsymbol{p})$ which also has a nonstandard parametric form. The resulting MCMC algorithm is in Appendix A. We also list the posteriors from which we generate the draws in Appendix B.

4 On the MCMC problems

To start the MCMC algorithm, we have to set initial parameter values and hyperparameter values in **MCMC0** in the MCMC algorithm in Appendix A. We find that inappropriate choices for some of these values prevent the MCMC algorithm from proceeding.

The first type of problem is induced by inappropriate $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\boldsymbol{\xi}^*$ and $\boldsymbol{\theta}^*$ generating nonpositive values for some components of cost $\boldsymbol{c}$ in the density of $\boldsymbol{p}$ in (3.7). This problem can occur in **MCMC2** and **MCMC5**. When this problem occurs, we have to stop the MCMC algorithm because whatever a firm produces takes cost.

The second type of problem is induced by an inappropriate set of $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\gamma^{(0)}$ and $\boldsymbol{\Sigma}_s^{(0)}$ generating the likelihood $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)}, \gamma^{(0)}, \boldsymbol{\Sigma}_s^{(0)}) = 0$ computationally. Even when this problem occurs, we can proceed with the MCMC algorithm. Once this problem occurs, however, the MCMC algorithm can continues to hover on the range of the computational $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \gamma, \boldsymbol{\Sigma}_s) = 0$ for a while before it finds a combination of values for $\boldsymbol{\xi}$, $\boldsymbol{\theta}$, γ and $\boldsymbol{\Sigma}_s$ generating computational $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \gamma, \boldsymbol{\Sigma}_s) > 0$. Since the set of true parameter values must be on the range with $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \gamma, \boldsymbol{\Sigma}_s) > 0$, this hovering can be a waste of time. In the following, we elaborate the mechanisms of these problems.

4.1 Nonpositive cost problem

Given price $\boldsymbol{p}$, the range of values $\{(\partial \boldsymbol{G}/\partial \boldsymbol{p})'\}^{-1}\boldsymbol{s}$ can take have to be restricted to make cost $\boldsymbol{c} = \boldsymbol{p} + \{(\partial \boldsymbol{G}/\partial \boldsymbol{p})'\}^{-1}\boldsymbol{s}$ positive. Hence, the MCMC algorithm must be able to find values for $\boldsymbol{\theta}$ and $\boldsymbol{\xi}$ so that $\boldsymbol{p} > \{(\partial \boldsymbol{G}/\partial \boldsymbol{p})'\}^{-1}\boldsymbol{s}$ while it attains convergence. There are four cases under which some components of cost $\boldsymbol{c}$ are nonpositive.

The first case occurs in **MCMC2** in the first iteration $t = 1$ where we calculate the density of $\boldsymbol{p}$ with $\boldsymbol{\xi}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$ to obtain $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)}, \gamma^{(0)}, \boldsymbol{\Sigma}_s^{(0)})$. The second case also occurs in **MCMC2** for $t = 1$ where we calculate $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}^*, \boldsymbol{\theta}^{(0)}, \gamma^{(0)}, \boldsymbol{\Sigma}_s^{(0)})$ with $\boldsymbol{\xi}^*$ from $MVN(\boldsymbol{0}, \boldsymbol{\Sigma}_d^{(0)})$ in **MCMC1** and $\boldsymbol{\theta}^{(0)}$. The third case takes place in **MCMC2** for $t = 2, \ldots$ where we calculate $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}^*, \boldsymbol{\theta}^{(t-1)}, \gamma^{(t-1)}, \boldsymbol{\Sigma}_s^{(t-1)})$ with $\boldsymbol{\xi}^*$ from $MVN(\boldsymbol{0}, \boldsymbol{\Sigma}_d^{(t-1)})$ in **MCMC1** given $\boldsymbol{\theta}^{(t-1)}$. The fourth case arises in **MCMC5** for $t = 1, \ldots$ where we calculate $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}^{(t)}, \boldsymbol{\theta}^*, \gamma^{(t-1)}, \boldsymbol{\Sigma}_s^{(t-1)})$ with $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_I^*)$ from $MVN(\bar{\boldsymbol{\theta}}^{(t-1)}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})$ in **MCMC4** given $\boldsymbol{\xi}^{(t)}$.

To avoid the nonpositive cost problem, we should set not only appropriate $\boldsymbol{\xi}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$ but also appropriate $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$ and $\boldsymbol{\Sigma}_d^{(0)}$ and $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, g_d and $\boldsymbol{G}_d$ in **MCMC0** because of the following reasons. We know that $\boldsymbol{\xi}^*$ depends on $\boldsymbol{\Sigma}_d^{(t-1)}$ for $t = 1, \ldots$ in **MCMC1**. For $t = 1$, we can alter $\boldsymbol{\Sigma}_d^{(0)}$ in **MCMC0**. For $t = 2, \ldots$, the range of values $\boldsymbol{\Sigma}_d^{(t-1)}$ can take in its poseterior in (B.3) is determined by g_d and $\boldsymbol{G}_d$ in its prior in (3.5) whose values can be also altered in **MCMC0**. We also know that $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_I^*)$ depend on $\bar{\boldsymbol{\theta}}^{(t-1)}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)}$ for $t = 1, \ldots$ in **MCMC4**. For $t = 1$, we can alter $\bar{\boldsymbol{\theta}}^{(0)}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$ in **MCMC0**. For $t = 2, \ldots$, the ranges of values $\bar{\boldsymbol{\theta}}^{(t-1)}$ and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)}$ can take in their conditional posteriors in (B.1) and (B.2) are determined by $\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}}$ and $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$ in the prior of $\bar{\boldsymbol{\theta}}$ and $g_{\boldsymbol{\theta}}$ and $\boldsymbol{G}_{\boldsymbol{\theta}}$ in the prior of

Σ_{θ} respectively in (3.4) whose values can be also altered in **MCMC0**.

4.2 Computational zero likelihood problem

Computationally, the likelihood $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s)$ has a narrow range of $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s) > 0$ and a wide range of $f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s) = 0$. The likelihood with $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\boldsymbol{\gamma}^{(0)}$ and $\boldsymbol{\Sigma}_s^{(0)}$ is written as

$$f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)}, \boldsymbol{\gamma}^{(0)}, \boldsymbol{\Sigma}_s^{(0)}) = f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)}) f(\boldsymbol{p}|\boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)}, \boldsymbol{\gamma}^{(0)}, \boldsymbol{\Sigma}_s^{(0)}).$$

This computational problem arises from either $f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)}) = 0$ or $f(\boldsymbol{p}|\boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)}, \boldsymbol{\gamma}^{(0)}, \boldsymbol{\Sigma}_s^{(0)}) = 0$ as explained below or both.

As for $f(\boldsymbol{p}|\boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)}, \boldsymbol{\gamma}^{(0)}, \boldsymbol{\Sigma}_s^{(0)})$ from (3.7), if we calculate it under an inappropriately small $\boldsymbol{\Sigma}_s^{(0)}$ relative to $\mathbf{log}[\boldsymbol{p} + \{(\boldsymbol{\partial G}/\boldsymbol{\partial p})'\}^{-1}\boldsymbol{s}] - \boldsymbol{Z}\boldsymbol{\gamma}^{(0)}$ which depends on inappropriate $\boldsymbol{\xi}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$ as well as $\boldsymbol{\gamma}^{(0)}$ given $\boldsymbol{y}$, $\boldsymbol{p}$, $\boldsymbol{X}$ and $\boldsymbol{Z}$, then it can be zero computationally. The $f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)})$ from (3.6) also becomes zero computationally with inappropriate $\boldsymbol{\xi}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$ generating extremely small $s_j(\boldsymbol{p}, \boldsymbol{X}, \boldsymbol{\xi}^{(0)}, \boldsymbol{y}, \boldsymbol{\theta}^{(0)})$ which in turn generates extremely small Is_j relative to the corresponding v_j. This is because $f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}^{(0)}, \boldsymbol{\theta}^{(0)})$ involves s_j and v_j in the form of $s_j^{v_j}$.

We next describe how inappropriate $\boldsymbol{\xi}^{(0)}$ and $\boldsymbol{\theta}^{(0)}$ make s_j extremely small, using a consumer i's representative utility for product j with them,

$$\begin{aligned} \alpha_i^{(0)} \log(y_i - p_j) + \boldsymbol{x}_j \boldsymbol{\beta}_i^{(0)} + \xi_j^{(0)} =& \alpha_i^{(0)} \log(y_i - p_j) + \beta_{i1}^{(0)} x_{j1} + \cdots \\ & + \beta_{iq}^{(0)} x_{jq} + \cdots + \beta_{i(Q-1)}^{(0)} x_{j(Q-1)} + \xi_j^{(0)}, \end{aligned} \tag{4.1}$$

in s_{ij} in (2.2) which is used to calculate s_j in (2.3). Given $\boldsymbol{y}$, $\boldsymbol{p}$ and $\boldsymbol{X}$, there are three cases.

The first case occurs when $\alpha_i^{(0)}$ is large so that the influence of $\alpha_i^{(0)}\log(y_i - p_j)$ on (4.1) is large relative to the influences of the remaining terms. This leads s_0 very large relative to $s_1, \ldots, s_J$ because the representative utility for $j = 0$ practically depends only on $\alpha_i^{(0)} \log y_i$ with $p_0 = 0$, $\boldsymbol{x}_0 = \mathbf{0}$ and $\xi_0 = 0$ which is larger than $\alpha_i^{(0)} \log(y_i - p_j)$ for $j = 1, \ldots, J$. When this happens, $s_1, \ldots, s_J$ can be practically zero.

The second case takes place when $\beta_{iq}^{(0)}$ is large so that the influence of $\beta_{iq}^{(0)} x_{jq}$ on (4.1) is large relative to the influences of the remaining terms. This makes s_j with the highest x_{jq} among $x_{1q}, \ldots, x_{Jq}$ very large relative to $s_0, \ldots, s_{j-1}, s_{j+1}, \ldots, s_J$ and so some of them can be practically zero.

The third case arises when the influence of $\xi_j^{(0)}$ on (4.1) is large relative to the influences of the remaining terms. This makes s_j with the highest $\xi_j^{(0)}$ among $\boldsymbol{\xi}^{(0)} = (\xi_1^{(0)}, \ldots, \xi_J^{(0)})'$ very large relative to $s_0, \ldots, s_{j-1}, s_{j+1}, \ldots, s_J$ and so some of them can be practically zero.

Note we have not encountered the computational zero likelihood problem so far with $\boldsymbol{\xi}^{(t)}$, $\boldsymbol{\theta}^{(t)}$, $\gamma^{(t)}$ and $\boldsymbol{\Sigma}_s^{(t)}$ for $t = 1, \ldots$. So it is important to have an appropriate set of $\boldsymbol{\xi}^{(0)}$, $\boldsymbol{\theta}^{(0)}$, $\gamma^{(0)}$ and $\boldsymbol{\Sigma}_s^{(0)}$ in **MCMC0**.

5 Simulation study

In this section, we obtain implications of our model from its simulation study where we test if the model can recover true parameter values with simulated data and diffuse priors. In subsection 5.1, we explain the simulation design. Subsection 5.2 explains how we set true parameter values and exogenous and endogenous variables. Subsection 5.3 impliments the MCMC algorithm according to the design and summarizes the results. As it turns out, the posterior standard deviation of each component of $\bar{\boldsymbol{\beta}}$ is large relative to that of $\bar{\alpha}$ but tends to be smaller as the number of consumers increases. We examine how many consumers we need to obtain reliable results for the components of $\bar{\boldsymbol{\beta}}$ in subsection 5.4 in the most complex case in our design. We also find that the simulations overestimate $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$, $\boldsymbol{\Sigma}_d$ and $\boldsymbol{\Sigma}_s$. Subsection 5.5 examines the causes of the overestimations.

5.1 Simulation design

We assume an oligopolistic market of a durable product where a consumer purchases one unit of a product. We set the overall market size $M = 500$, 1000 or 2000 and then use all the M consumers as the sample consumers ($M = I$). The market offers $J = 5$, 10 or 25 products. The number $J = 5$ implies that the market is highly oligopolistic. The number $J = 25$ comes from our upcoming empirical study of the U.S. automobile market in 1996 where the sales of the top 25 cars occupy about 51.3% of the total sales. We set $j = 0$ for the outside good.

On the demand side, a consumer i has his/her own utility u_{ij} for product j in (2.1). On the supply side, the pricing equation of firms is in (2.8).

For each combination of I and J, we change the number of products one firm produces, that of observed product characteristics $\boldsymbol{x}_j$ in (2.1), that of cost shifters $\boldsymbol{z}_j$ in (2.7) and the degree to which $\boldsymbol{x}_j$ overlaps $\boldsymbol{z}_j$. The details are as follows.

The number of products one firm produces

When $J = 5$, each firm produces one product. When $J = 10$, it produces either one or two products. This means that there are ten or five firms respectively. When $J = 25$, it produces either one or five products. Again this corresponds to either 25 or five firms respectively. Note that the markets are highly oligopolistic when $J = 10$ and $J = 25$ with each firm producing multiple products as well as when $J = 5$.

The numbers of observed product characteristics and cost shifters

We consider three cases where both are 1, 5 or 10. Since Q—the length of the vector $\boldsymbol{\theta}_i$—includes price, these choices make Q and S as $(Q, S) = (2, 1)$, $(6, 5)$ or $(11, 10)$ respectively. Note that $Q = 2$ implies that researchers can observe only one differentiating product characteristic other than price because, for example, products in the market are homogeneous; and $Q = 11$ comes from the past empirical studies in the U.S. automobile market (Berry, Levinsohn & Pakes, 1995; Sudhir, 1999; Myojo, 2006). Given J, we only consider cases where S is less than J in the pricing equation (2.8): When $J = 5$, $(Q, S) = (2, 1)$; when $J = 10$, $(Q, S) = (2, 1)$ or $(6, 5)$; and when $J = 25$, $(Q, S) = (2, 1)$, $(6, 5)$ or $(11, 10)$.

The degree to which observed product characteristics overlap cost shifters

For each case of (Q, S), we further consider three cases: Independence where $\boldsymbol{x}_j$ and $\boldsymbol{z}_j$ are separate; overlap where they completely overlap; and partial overlap. When $(Q, S) = (2, 1)$, we have either independece with $\boldsymbol{z}_j = z_{j1}$ or overlap with $\boldsymbol{z}_j = \boldsymbol{x}_j$. Note that partial overlap is impossible when $(Q, S) = (2, 1)$. When $(Q, S) = (6, 5)$, $\boldsymbol{z}_j = (z_{j1}, \ldots, z_{j5})$ defines independence, $\boldsymbol{z}_j = \boldsymbol{x}_j$ defines overlap, and $\boldsymbol{z}_j = (x_{j1}, \ldots, x_{j4}, z_{j5})$ defines partial overlap. When $(Q, S) = (11, 10)$, $\boldsymbol{z}_j = (z_{j1}, \ldots, z_{j10})$ defines independence, $\boldsymbol{z}_j = \boldsymbol{x}_j$ defines overlap, and $\boldsymbol{z}_j = (x_{j1}, \ldots, x_{j8}, z_{j9}, z_{j10})$ defines partial overlap.

5.2 True parameter values and exogenous and endogenous variables

We obtain positive $y_1, \ldots, y_M$ randomly from the log normal distribution with mean 1 and standard deviation 0.1. We also obtain values for $x_{j1}, \ldots, x_{j(Q-1)}$ and $z_{j1}, \ldots, z_{jS}$ for $j = 1, \ldots, J$ randomly from $N(0, 0.1)$. As for the outside good $j = 0$, we set $p_0 = 0$, $\boldsymbol{x}_0 = \mathbf{0}$ and $\xi_0 = 0$. We also set

$$\bar{\boldsymbol{\theta}} = (\bar{\alpha}, \bar{\boldsymbol{\beta}}')' = (2, 2, \ldots, 2)', \ \boldsymbol{\Sigma}_{\boldsymbol{\theta}} = 10^{-1}\boldsymbol{E}_Q,$$
$$\boldsymbol{\gamma} = (1, \ldots, 1)', \ \boldsymbol{\Sigma}_d = 10^{-4}\boldsymbol{E}_J, \ \boldsymbol{\Sigma}_s = 10^{-4}\boldsymbol{E}_J$$

where $\boldsymbol{E}_Q$ and $\boldsymbol{E}_J$ are the $Q \times Q$ and $J \times J$ identity matrices respectively. We then generate $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_M)$ randomly from $MVN(\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$ in (3.3). We also generate $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_J)'$ and $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_J)'$ from $MVN(\mathbf{0}, \boldsymbol{\Sigma}_d)$ in (3.1) and $MVN(\mathbf{0}, \boldsymbol{\Sigma}_s)$ in (3.2) respectively. We determine $\boldsymbol{v}^o$ and $\boldsymbol{p}$ endogenously in the demand with (2.3) and supply with (2.8), using the Newton-Raphson method.

5.3 MCMC with diffuse priors

For each case in subsection 5.1, we run three independent MCMC sequences each of which has 10,000 iterations with a different set of initial parameter values. Based on the implications in Section 4, we set initial parameter values and hyperparameter values at **MCMC0**. To use relatively diffuse priors in (3.4) and (3.5), we set the hyperparameter values as

$$\boldsymbol{\mu}_{\bar{\boldsymbol{\theta}}} = (20, 0, \ldots, 0)', \ \boldsymbol{V}_{\bar{\boldsymbol{\theta}}} = 10^2\boldsymbol{E}_Q, \ g_{\boldsymbol{\theta}} = Q + 4, \ \boldsymbol{G}_{\boldsymbol{\theta}} = 3\boldsymbol{E}_Q, \ \bar{\boldsymbol{\gamma}} = (0, \ldots, 0)',$$
$$\boldsymbol{V}_{\boldsymbol{\gamma}} = 10^2\boldsymbol{E}_S, \ g_d = J + 4, \ \boldsymbol{G}_d = 3 \times 10^{-2}\boldsymbol{E}_J, \ g_s = J + 4, \ \boldsymbol{G}_s = 3 \times 10^{-2}\boldsymbol{E}_J.$$

We next set the initial parameter values as

$$\bar{\boldsymbol{\theta}}^{(0)} = (2.5, \ldots, 2.5)', \ \ (3, \ldots, 3)' \text{ and } (3.5, \ldots, 3.5)',$$

$$\gamma^{(0)} = (-5, \ldots, -5)', \ (0, \ldots, 0)' \text{ and } (5, \ldots, 5)',$$

respectively for each sequence, fixed $\mathbf{\Sigma}_{\boldsymbol{\theta}}^{(0)}$, $\mathbf{\Sigma}_d^{(0)}$ and $\mathbf{\Sigma}_s^{(0)}$:

$$\mathbf{\Sigma}_{\boldsymbol{\theta}}^{(0)} = \boldsymbol{E}_Q, \ \mathbf{\Sigma}_d^{(0)} = \boldsymbol{E}_J, \ \mathbf{\Sigma}_s^{(0)} = \boldsymbol{E}_J,$$

and $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, \ldots, \boldsymbol{\theta}_I^{(0)}) \sim MVN(\bar{\boldsymbol{\theta}}^{(0)}, \mathbf{\Sigma}_{\boldsymbol{\theta}}^{(0)})$ and $\boldsymbol{\xi} \sim MVN(\mathbf{0}, \mathbf{\Sigma}_d^{(0)})$.

We inspect a time-series plot of the draws for each parameter from the three sequences to assess the convergence of the MCMC. Given the last halves of the three sequences, we also check if the 95% posterior inverval of each parameter includes its true value.

We are confident that the components of $\bar{\boldsymbol{\theta}} = (\bar{\alpha}, \bar{\boldsymbol{\beta}}')'$ and γ converge to their true values in almost all the cases. The posterior standard deviation of each component of $\bar{\boldsymbol{\theta}} = (\bar{\alpha}, \bar{\boldsymbol{\beta}}')'$ tends to be smaller as I and J increase while that of each component of γ becomes smaller only as J increases but is not affected by the increasing I. This is because $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$ depend on $\bar{\boldsymbol{\theta}}$ and appear in the utility (2.1) on the demand side as well as in the pricing equation (2.8) on the supply side, while γ appears only in (2.8).

We are somewhat concerned about the following three facts. First, the posterior standard deviation of each component of $\bar{\boldsymbol{\beta}}$ is large relative to that of $\bar{\alpha}$. In subsetion 5.4, we examine how many consumers we need to obtain a reliable result for each component of $\bar{\boldsymbol{\beta}}$ in the most complex case in our simulation design. Second, some of the 95% posterior intervals of the components of $\bar{\boldsymbol{\beta}}$ and γ include 0 as well as their true values. Third, two 95% posterior intervals of the components of $\bar{\boldsymbol{\beta}}$ do not include their true values.

The problem we encountered is that the diagonal components of $\mathbf{\Sigma}_{\boldsymbol{\theta}}$, $\mathbf{\Sigma}_d$ and $\mathbf{\Sigma}_s$ are everestimated as far as their time-series plots and summary statistics are concerned. We explore reasons as to these phenomena in subsection 5.5.

5.4 The number of consumers for a reliable $\bar{\boldsymbol{\beta}}$

We examine how many consumers I we need to obtain more reliable estimates for the components of $\bar{\boldsymbol{\beta}}$. We consider a case of $M = 10000$, $J = 25$, $Q = 11$ and $S = 10$ with each firm producing five products in the partial overlap cost shifter case, which is the most complex case in our simulation design. The other settings for simulated data are the same as those in subsection 5.2. We use ten sets of consumers of $I = 500$, 1000 thorugh 9000 increment by $1,000$ drawn randomly from the original $10,000$ comsumers and all the $10,000$ consumers.

We run ten MCMC sequences each of which has $10,000$ iterations for each set of consumers. We set hyperparameter values and initial parameter values in the same way as that in subsection 5.3 except for $\bar{\boldsymbol{\theta}}^{(0)}$, $\gamma^{(0)}$ and $\boldsymbol{\Sigma}_d^{(0)}$. As for $\bar{\boldsymbol{\theta}}^{(0)}$, $\gamma^{(0)}$ and $\boldsymbol{\Sigma}_d^{(0)}$, we set

$$\bar{\boldsymbol{\theta}}^{(0)}=(2.05,\ldots,2.05)',\ldots,(2.5,\ldots,2.5)' \text{ increment by } 0.05,$$

$$\gamma^{(0)}=(-5,\ldots,-5)',\ldots,(5,\ldots,5)' \text{ except for } (1,\ldots,1)' \text{ increment by } 1,$$

$$\boldsymbol{\Sigma}_d^{(0)}=10^{-1}\boldsymbol{E}_J,$$

for each sequence based on the implications in Section 4.

As I increases, the amount of the reductions of the posterior standard deviation of each component of $\bar{\boldsymbol{\theta}}$ including $\bar{\boldsymbol{\beta}}$ decreases. When $I \geq 4000$, the fluctuations of the components of $\bar{\boldsymbol{\beta}}$ do not seem to improve noticeably based on their time-series plots and posterior standard deviations. Therefore, we can use $I = 5,000$ consumers to obtain reliable results for the components of $\bar{\beta}$ when $J = 25$, $Q = 11$, $S = 10$ with each firm producing five products in the partial overlap cost shifter case.

5.5 On overestimating Σ_{θ}, Σ_d and Σ_s

On overestimating Σ_{θ}

We found wrong $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$ induce the overestimated Σ_{θ} from the following three nested experiments to estimate Σ_{θ}. First has only **MCMC8** with the true $\bar{\boldsymbol{\theta}}$ and $\boldsymbol{\theta}$. Second is the Gibbs sampler with **MCMC7** and **MCMC8** with the true $\boldsymbol{\theta}$. Third is the Gibbs sampler with **MCMC4** through **MCMC8**. Note that hyperparameter values are the same as those in subsection 5.3 and initial parameter values for each experiment are far from thier true values. Although we can recover true Σ_{θ} in the first and second experiments, we can not in the third experiment. This implies that **MCMC4** thorugh **MCMC6** incorrectly estimate $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_I)$ which in turn induce the overestimated Σ_{θ}. The **MCMC4** thorugh **MCMC6** are the Metropolis-Hastings algorithm generating draws of $\boldsymbol{\theta}$ where we accept a proposal draw for $\boldsymbol{\theta}$ with an acceptance probability from the likelihood ratio in each iteration. This can not work well. We need to examine the likelihood ratio with simulated data.

On overestimating Σ_d and Σ_s

The overestimated Σ_d and Σ_s are induced by the large influence of each diffuse prior on its posterior with the small number of observations (one course of observation). If the number of observations was large enough, the influence of each diffuse prior on its corresponding posterior would be small.

6 Summary

In this paper, we reviewed Yonetani et al.'s (2007) Bayesian simultaneous demand and supply model with market-level data. We also summarized the nonpositive cost problem and computational zero likelihood problem which prevented the MCMC algorithm from proceeding. They imply that we can not always set any diffuse priors and initial parameter values for the MCMC algorithm. We also performed a simulation study with diffuse priors which could avoid the problems above.

In the simulation study, the means of consumers' coefficients and the coefficients for cost shifters were correctly estimated for almost all of the various cases we considered. The posterior standard deviations of the means of consumers' coefficients for observed product characteristics were large when the number of consumers is small. From the additional simulation study, we found that $5,000$ consumers could be used to obtain reliable estimates for them.

On the other hand, the variance-covariance matrices of consumers' coefficients and unobserved product and cost characteristics were overestimated. The variance-covariance matrix of consumers' coefficients was overestimated because of incorrectly estimated consumers' coefficients while the variance-covariance matrices of unobserved product and cost characteristics were overestimated because of the small number of observations.

In future, we need the following three studies. First, we examine the likelihood ratio in the Metropolis-Hastings algorithm generating the incorrect consumers' coefficients with simulated data to overcome their overestimated variance-covariance matrix. Second, we implement additional simulation studies with panel data which has the large number of observations to overcome the overestimated variance-covariance matrices of unobserved product and cost characteristics. Third, based on the fact that more informative priors can estimate all of the parameters correctly from Yonetani et al. (2008), we develop a pre-analytical process to obtain such priors.

A MCMC algorithm

MCMC0 Set $\mu_{\bar{\boldsymbol{\theta}}}$, $\boldsymbol{V}_{\bar{\boldsymbol{\theta}}}$, $g_{\boldsymbol{\theta}}$, $\boldsymbol{G}_{\boldsymbol{\theta}}$, g_d, $\boldsymbol{G}_d$, $\bar{\gamma}$, $\boldsymbol{V}_{\gamma}$, g_s and $\boldsymbol{G}_s$ and $\boldsymbol{\theta}^{(0)}$, $\bar{\boldsymbol{\theta}}^{(0)}$, $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(0)}$, $\gamma^{(0)}$, $\boldsymbol{\Sigma}_s^{(0)}$ $\boldsymbol{\Sigma}_d^{(0)}$ and $\boldsymbol{\xi}^{(0)}$.

For $t = 1, \ldots,$

MCMC1 Generate a proposal $\boldsymbol{\xi}^*$ from $MVN(\mathbf{0}, \boldsymbol{\Sigma}_d^{(t-1)})$.

MCMC2 Calculate

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

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 proposal $\boldsymbol{\theta}^* = (\boldsymbol{\theta}_1^*, \ldots, \boldsymbol{\theta}_I^*)$ randomly from $MVN(\bar{\boldsymbol{\theta}}^{(t-1)}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}^{(t-1)})$.

MCMC5 Calculate

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

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

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

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

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

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

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

MCMC12 If random draws from **MCMC6**, **MCMC7**, **MCMC8**, **MCMC9**, **MCMC10** and **MCMC11** stabilize, then stop the iteration. Otherwise increase t by one and return to **MCMC1**.

B Posteriors in MCMC

We obtain the (conditional) posteriors in the MCMC as follows.

$$
\begin{aligned}
&f(\boldsymbol{\xi}\,|\boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_d, \boldsymbol{\Sigma}_s, \boldsymbol{v}, \boldsymbol{p}) \propto f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s) f(\boldsymbol{\xi}|\boldsymbol{\Sigma}_d) \\
&= f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}, \boldsymbol{\theta}) f(\boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s) f(\boldsymbol{\xi}|\boldsymbol{\Sigma}_d) \\
&\propto s_0^{v_0} \cdots s_J^{v_J} \cdot \\
&\quad \times |\boldsymbol{\Sigma}_s|^{-\frac{1}{2}} \left\| \left(\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{p}} \right) \right\| \\
&\quad \times \exp\left[-\frac{1}{2} \left[\log \left[\boldsymbol{p} + \left\{ \left(\frac{\partial G}{\partial \boldsymbol{p}} \right)' \right\}^{-1} \boldsymbol{s} \right] - Z\boldsymbol{\gamma} \right]' \boldsymbol{\Sigma}_s^{-1} \left[\log \left[\boldsymbol{p} + \left\{ \left(\frac{\partial G}{\partial \boldsymbol{p}} \right)' \right\}^{-1} \boldsymbol{s} \right] - Z\boldsymbol{\gamma} \right] \right] \\
&\quad \times |\boldsymbol{\Sigma}_d|^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \boldsymbol{\xi}' \boldsymbol{\Sigma}_d^{-1} \boldsymbol{\xi} \right),
\end{aligned}
$$

$$
\begin{aligned}
&f(\boldsymbol{\theta}|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s, \boldsymbol{\xi}, \boldsymbol{v}, \boldsymbol{p}) \propto f(\boldsymbol{v}, \boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s) \left[\prod_{i=1}^{I} f(\boldsymbol{\theta}_i|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}) \right] \\
&= f(\boldsymbol{v}|\boldsymbol{p}, \boldsymbol{\xi}, \boldsymbol{\theta}) f(\boldsymbol{p}|\boldsymbol{\xi}, \boldsymbol{\theta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}_s) \left[\prod_{i=1}^{I} f(\boldsymbol{\theta}_i|\bar{\boldsymbol{\theta}}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}}) \right] \\
&\propto s_0^{v_0} \cdots s_J^{v_J}
\end{aligned}
$$

$$
\begin{aligned}
&\times |\boldsymbol{\Sigma}_s|^{-\frac{1}{2}} \left\| \left(\frac{\partial \eta}{\partial p} \right) \right\| \\
&\times \exp \left[-\frac{1}{2} \left[\log \left[p + \left\{ \left(\frac{\partial G}{\partial p} \right)' \right\}^{-1} s \right] - Z\gamma \right]' \boldsymbol{\Sigma}_s^{-1} \left[\log \left[p + \left\{ \left(\frac{\partial G}{\partial p} \right)' \right\}^{-1} s \right] - Z\gamma \right] \right] \\
&\times \prod_{i=1}^{I} \left[|\boldsymbol{\Sigma}_{\theta}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} (\theta_i - \bar{\theta})' \boldsymbol{\Sigma}_{\theta}^{-1} (\theta_i - \bar{\theta}) \right\} \right],
\end{aligned}
$$

$$
\bar{\theta} \,|\theta, \boldsymbol{\Sigma}_{\theta} \sim N \left(\left(I\boldsymbol{\Sigma}_{\theta}^{-1} + V_{\bar{\theta}}^{-1} \right)^{-1} \left(I\boldsymbol{\Sigma}_{\theta}^{-1}\nu + V_{\bar{\theta}}^{-1}\mu_{\bar{\theta}} \right), \left(I\boldsymbol{\Sigma}_{\theta}^{-1} + V_{\bar{\theta}}^{-1} \right)^{-1} \right) \tag{B.1}
$$

where $\nu = \frac{1}{I} \sum_{i=1}^{I} \theta_i,$

$$
\boldsymbol{\Sigma}_{\theta} \,|\theta, \bar{\theta} \sim IW_{g_{\theta}+I} \left(\sum_{i=1}^{I} (\theta_i - \bar{\theta})\, (\theta_i - \bar{\theta})' + G_{\theta} \right), \tag{B.2}
$$

$$
\gamma \,|\theta, \boldsymbol{\Sigma}_s, \xi, p \sim N \left(\left(\boldsymbol{\Sigma}_{s*}^{-1} + V_{\bar{\gamma}}^{-1} \right)^{-1} \left(\mu + V_{\bar{\gamma}}^{-1}\bar{\gamma} \right), \left(\boldsymbol{\Sigma}_{s*}^{-1} + V_{\bar{\gamma}}^{-1} \right)^{-1} \right),
$$

where $\mu = Z'\boldsymbol{\Sigma}_s^{-1} \left[\log \left[p + \left\{ \left(\frac{\partial G}{\partial p} \right)' \right\}^{-1} s \right] \right]$ and $\boldsymbol{\Sigma}_{s*}^{-1} = Z'\boldsymbol{\Sigma}_s^{-1} Z,$

$$
\begin{aligned}
&\boldsymbol{\Sigma}_s \,|\theta, \gamma, \xi, p \\
&\sim IW_{g_s+1} \left(\left(\log \left[p + \left\{ \left(\frac{\partial G}{\partial p} \right)' \right\}^{-1} s \right] - Z\gamma \right) \left(\log \left[p + \left\{ \left(\frac{\partial G}{\partial p} \right)' \right\}^{-1} s \right] - Z\gamma \right)' + G_s \right),
\end{aligned}
$$

$$
\boldsymbol{\Sigma}_d \,|\xi \sim IW_{g_d+1} \left(\xi\xi' + G_d \right). \tag{B.3}
$$

Reference

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

Chibs, S. & Greenberg, E., (1995). “Understanding the Metropolis-Hastings algorithm,” *American Statistical Association*, Vol.49(4), 327-335.

Myojo, S. “A study on the U.S. consumers’ automobile preferences,” Doctoral Dissertation, Doctoral Program in Social Systems & Management, Graduate School of Systems & Information Engineering, University of Tsukuba.

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

Tanner, M. A. & Wong, W. H. (1987). "The calculation of posterior distribution by data augmentation," *Journal of the American Statistical Association*, Vol.82, 528-540.

Yonetani, Y., Kanazawa, Y. & Myojo, S. (2007). "Bayesian analysis of simultaneous demand and supply with market-level data - U.S. auto market," 2007 INFORMS Marketing Science Conference.

Yonetani, Y., Kanazawa, Y. & Myojo, S. (2008). "A simulation study on Bayesian simultaneous demand and supply model with market-level data," *RIMS Kôkyûroku*, Vol.1603, 50-72, Research Institute for Mathematical Sciences, Kyoto Univeristy.