Duality-Based Bayesian Analysis of Residential Gas Demand under Decreasing Block Rate Pricing

Koji Miyawaki,[*] Yasuhiro Omori[†]

Abstract

This paper discusses a novel Bayesian estimation method for the residential gas demand function in Japan where the price per unit decreases as the demand exceeds certain thresholds. Such a price system is known as decreasing block rate pricing. The demand function under decreasing block rate pricing is derived by using the well-known discrete/continuous choice approach. However, because of the nonconvex budget set, the conventional approach imposes highly nonlinear constraints on the model parameters, thus making the maximization of the likelihood function under such constraints difficult to implement. To overcome this difficulty, we first apply the duality relationship in consumer theory, and approximate the conditional expenditure in order to linearize these nonlinear constraints. Then, we adopt a Bayesian approach with the Markov chain Monte Carlo simulation in order to estimate the model parameters under linear constraints. Our proposed method is illustrated by a numerical example and is adopted to analyze the demand for residential gas in Japan.

JEL classification: C11, D12, Q41

keyword: Block rate pricing, Discrete/continuous choice approach, Duality, Bayesian analysis, Markov chain Monte Carlo

1 Introduction

Energy resources are often supplied under block rate pricing, in which case the price per unit changes as the consumption exceeds certain thresholds. Further, when the unit price decreases as the consumption increases, such a system is called decreasing block rate pricing. The price system followed by the gas services in Japan is an example of this decreasing block rate pricing. On the other hand, when the unit price increases as the consumption increases, such a system is called increasing block rate pricing. The water and electricity services in Japan follow this increasing block rate pricing. Under block rate pricing systems, consumers need to choose both price and consumption simultaneously. The discrete/continuous (D/C) choice approach commonly deals with such a simultaneous decision based on microeconomic theory.

Under increasing block rate pricing, this D/C choice approach is directly applicable to derive demand functions that are used to conduct model parameter estimation because its budget set is convex (see Hewitt and Hanemann 1995; Olmstead, Hanemann, and Stavins 2007; Reiss and White 2005; Miyawaki,

[*]Graduate School of Economics, University of Tokyo (E-mail: *ee57008@mail.ecc.u-tokyo.ac.jp*).

[†]Faculty of Economics, University of Tokyo (E-mail: *omori@e.u-tokyo.ac.jp*).

Omori, and Hibiki 2006). In contrast, under decreasing block rate pricing, its budget set becomes nonconvex, as will be seen in the next section, and prevents the direct application of this D/C choice approach.

The D/C choice approach with decreasing block rate pricing requires consumers to compare utility levels conditional on every block choice in order to decide their optimal block. The resulting demand function inherits this comparison as nonlinear constraints that are mostly a set of nonlinear functions of prices and income, and the maximum likelihood estimation of such demand functions becomes computationally prohibitive as the number of block increases. In fact, empirical analyses have been limited to the case where there are only two decreasing blocks (see, e.g., Hausman 1980; Burtless and Moffitt 1985; de Jong 1990).

At a practical level, however, gas services are supplied under three or more blocks in Japan. Thus, to overcome the difficulty in fitting the D/C choice approach, we first apply the idea of duality in consumer theory, and approximate the conditional expenditure in order to linearize constraints. Then, we take a Bayesian approach with the Markov chain Monte Carlo (MCMC) simulation in order to estimate the model parameters under linear constraints.

The advantages of our duality-based approach are as follows. First, because nonlinear constraints are all approximated by linear constraints, it is straightforward to derive the demand function under multiple-block decreasing block rate pricing as a multinomial extension of the Type II Tobit model. Second, as a result of the approximation with duality, the model is relatively insensitive to the consumer preference change and specific consumer preferences. As a result, provided the demand function conditional on the block choice is linear in its parameters, we have relatively robust parameter estimates of the demand function regardless of consumer preferences. Third, in contrast to the nonparametric approach proposed by Blomquist and Newey (2002), our approach is parametric, and thus, the model parameters have economic implications in terms of the price and income elasticities.

The estimation of demand functions under decreasing block rate pricing plays an important role in policy making. Theoretically, among block rate pricing systems, decreasing block rate pricing can attain the second-best optimality (see, e.g., Chapter 7 of Train 1991). Thus, we can discuss whether the current price system is desirable in terms of economic efficiency. Furthermore, from the viewpoint of environmental economics, it is substantial for policy makers to estimate the residential gas demand function and its price elasticity to discuss the second-best taxation on the monopolistic or oligopolistic gas market (see, e.g., Chapter 6 of Baumol and Oates 1988 and Chapter 5 of Xepapadeas 1997).

This paper is organized as follows. Section 2 derives the demand function under decreasing block rate pricing. In Section 3, we derive the likelihood function, posterior distribution, and Gibbs sampler for the Bayesian analysis of the D/C choice model based on duality. Then, we apply our proposed method to the simulated data in Section 4 and conduct an empirical analysis using Japanese residential gas demand data in Section 5. Section 6 concludes the paper.

2 Demand Function under Decreasing Block Rate Pricing

2.1 Model Setting

We first describe the decreasing block rate pricing system for a good, the demand of which is denoted by Y. The other goods are treated as a numeraire commodity, and their price is normalized as one. In the

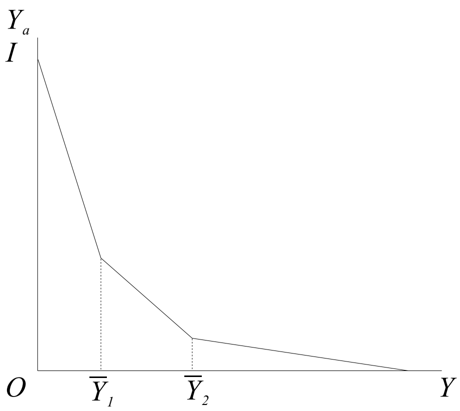

Figure 1: Budget Constraint: The Three-Block Case.

block rate pricing system, the consumption is divided into K blocks. Each block's upper limit is called the threshold, and the k-th threshold is denoted by $\bar{Y}_k$ $(k = 0,\ldots,K)$. In this paper, we set $\bar{Y}_0 \equiv 0$ and $\bar{Y}_K \equiv \infty$ for convenience. Associated with each thresholds, there are unit prices denoted by P_k, which is constant provided the consumption is more than $\bar{Y}_{k-1}$ but less than or equal to $\bar{Y}_k$ $(k = 1,\ldots,K)$. The price system follows decreasing block rate pricing when its price per unit decreases as k increases, that is, $P_k > P_{k+1}$ for $k = 1,\ldots,K-1$. Finally, there is a basic connection charge, FC, which is treated as a fixed cost.

Let Y_a and I be the expenditure for the other goods except Y and the total income, respectively, and let $U(Y,Y_a)$ denote the well-defined utility function. Then, we organize the consumer's utility maximization problem of the two goods as below.

$$V = \max_{Y,Y_a} U(Y,Y_a) \quad \text{subject to } Y_a + c(Y) \leq I, \tag{1}$$

where $c(Y)$ is the cost function for consuming Y and is given by

$$c(Y) = FC + P_k(Y - \bar{Y}_{k-1}) + \sum_{j=1}^{k-1} P_j(\bar{Y}_j - \bar{Y}_{j-1}), \quad \text{if } \bar{Y}_{k-1} < Y \leq \bar{Y}_k. \tag{2}$$

Figure 1 illustrates the budget constraint under three-block decreasing block rate pricing. As we can see from Figure 1, the budget constraint becomes piecewise linear in the consumption Y; furthermore, because of the decreasing block rate pricing system, consumers face a nonconvex budget set.

In order to derive the demand function, we define the conditional demands, Y_k, and the corresponding conditional indirect utility functions, V_k, for each block $(k = 1,\ldots,K)$. This pair of conditional values (Y_k,V_k) denote the maximizer and solution, respectively, to the k-th conditional utility maximization problem given below.

$$V_k = \max_{Y,Y_a} U(Y,Y_a) \quad \text{subject to } Y_a + P_kY \leq Q_k, \tag{3}$$

where Q_k is the virtual income defined as $Q_k = I - FC - \sum_{j=1}^{k-1}(P_j - P_{j+1})\bar{Y}_j$. We assume Q_K to be positive such that this conditional problem can be well defined.[1]

[1]When $Q_{K'+1} \leq 0$, we cut down the model to K'-block decreasing block rate pricing.

Further, this paper focuses on the following linear structure on the conditional demand Y_k.

$$\ln Y_k = \beta_1 \ln P_k + \beta_2 \ln Q_k \tag{4}$$

because it is one of the most popular forms in the demand analysis. In this log-linear model, β_1 and β_2 represent the price and income elasticity, respectively, conditional on the block choice. In order to avoid tedious notations, we set $(y, y_k, \bar{y}_k, p_k, q_k) = (\ln Y, \ln Y_k, \ln \bar{Y}_k, \ln P_k, \ln Q_k)$. Then, the k-th conditional demand is rewritten as

$$y_k = \beta_1 p_k + \beta_2 q_k \equiv \boldsymbol{x}_k' \boldsymbol{\beta}, \tag{5}$$

where $\boldsymbol{x}_k = (p_k, q_k)'$ and $\boldsymbol{\beta} = (\beta_1, \beta_2)'$. Finally, we have the demand function with the D/C choice approach.

$$y = y_k, \quad \text{if } V_k = \max_j V_j, \tag{6}$$

(see, e.g., Section 2 of Moffitt 1986 for the discussion on the two-block rate pricing model).

We mention three points relating with this demand model. First, the model implicitly assumes that there is only one optimal block. Under decreasing block rate pricing, the nonconvex budget set can yield multiple solutions. However, as we will discuss in Section 3, such a case can be excluded in our statistical modeling.

Second, the model would be sensitive to the underlying consumer preference. In other words, even when consumer preferences change infinitesimally, thereby resulting in a new but similar indifference curve, the new optimal block would neither remain the same as before nor change in the neighborhood of the previously optimal block. While this is consistent with the rational consumer, such extreme block change is impractical. As can be shown in Subsection 2.3, this paper introduces an approximation of consumer preferences into the model, which leads to a less sensitive model with the preference change than the original one. Theoretical modification, however, would remain a challenge for a future research.

Third, this demand model is not concerned with the effect of the supply structure because in most cases, goods or services under block rate pricing are provided by regionally monopolistic companies, and such companies are obliged to supply as much of the goods and services as the consumer needs. It would be possible to introduce the supply system into this model by utilizing the so-called disequilibrium model (see, e.g., Chapter 10 of Maddala 1983), which is also an issue for future studies.

2.2 Conventional Approach

In general, the structural demand analysis assumes a certain consumer preference behind the utility maximization problem and requires that the demand functions be consistent with this preference. For this purpose, there are three major methods to determine the functional form of the demand functions.

- Specify the direct utility function to derive its demand function.
- Specify the indirect utility to derive its demand function.
- Specify the demand function and recover its indirect utility.

The third method is often selected for the block rate pricing problem because of its flexibility in fitting the demand function to the data (see, e.g., Hausman 1985). However, when we consider more than two-block decreasing block rate pricing, this approach reveals its limitation. Under the assumption of the log-linear conditional demand function, Roy's identity recovers the corresponding conditional indirect utility, which is given by

$$V_k = -\frac{P_k^{1+\beta_1}}{1+\beta_1} + \frac{Q_k^{1-\beta_2}}{1-\beta_2}, \quad k = 1, \ldots, K. \tag{7}$$

The conventional approach directly plugs these nonlinear conditional indirect utilities into the demand function (6) and estimates its model parameters. It is, however, difficult to evaluate the parameters' space satisfying $V_k = \max_j V_j$ when $K > 2$. Thus, the next subsection explains an improved duality-based approach for demand functions under decreasing block rate pricing.

Remark 1. Even when $K = 2$, the conventional approach takes much more computational time in estimating the model parameters than the duality-based approach. Bayesian estimation with the conventional approach was conducted with the same set of parameters and data as described in Section 4. Further, we impose some conditions on the data set to reduce the computational burden: the logarithm of the price and virtual income for the first block is negative, such that the difference of the indirect utilities between the first and second block is monotonically decreasing in terms of β_j ($j = 1, 2$) given β_k ($k \neq j$, $k = 1, 2$). Then, the conventional approach takes more than seven times as much time as does the duality approach does and requires more samples to converge to their posterior distribution.

2.3 Duality Approach

This subsection proposes the duality-based approach for the estimation of the demand functions under decreasing block rate pricing. First, suppose that the k-th block is optimal, that is, $V_k = \max_j V_j$. Next, we define the j-th conditional expenditure E_j ($j = 1, \ldots, K$) as the solution to the conditional expenditure minimization problem given below.

$$E_j = \min_{Y, Y_a} Y_a + P_j Y \quad \text{subject to } U(Y, Y_a) \leq V, \tag{8}$$

where V is the solution to the original utility maximization problem (1). This Problem (8) is the dual problem to Problem (3).

Thanks to duality in consumer theory, there is a relationship between conditional expenditures and indirect utilities: $E_j = Q_j$ if and only if $V_j = V$, and $E_j > Q_j$ if and only if $V_j < V$.[2] Figure 2 illustrates two cases when the second block is both optimal and suboptimal under three-block decreasing block rate pricing. Because the D/C choice model assumes that $V = V_k = \max_j V_j$, we have the relationship between conditional indirect utilities and conditional expenditures, which is stated as below.

$$V_k = \max_j V_j \iff \begin{cases} E_k = Q_k, \\ E_j > Q_j, \quad \text{for } j \neq k. \end{cases} \tag{9}$$

It is still difficult to jointly estimate the parameters of conditional expenditure and demand function, and thus, we approximate E_j by Taylor expansion around P_k evaluated at P_j. The Envelope Theorem in consumer theory yields

$$E_j \approx Q_k + Y_k (P_j - P_k). \tag{10}$$

[2]See Figure 2.8 on p.38 of Deaton and Muellbauer (1980) for a general relationship between the indirect utility and expenditure function.

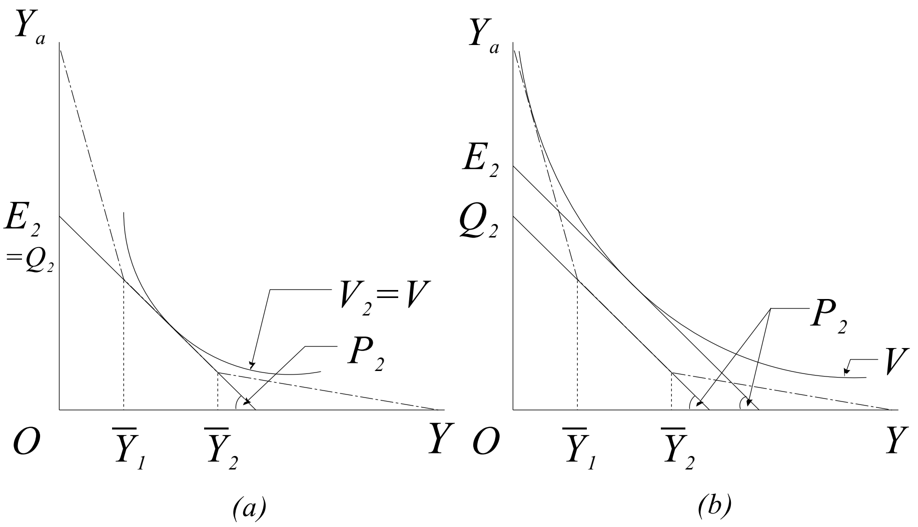

Figure 2: The Three-Block Case; the Second Block is Optimal (left) and Suboptimal (right).

A proof for this approximation is given below.

Proof of Eq. (10). Suppose that the k-th block is optimal. Let $E(P)$ be the solution to the minimization problem below.

$$E(P) = \min_{Y, Y_a} Y_a + PY \quad \text{subject to } U(Y, Y_a) \leq V. \tag{11}$$

It is straightforward that $E(P_j) = E_j$. Then, we approximate $E(P)$ by Taylor expansion around P_k.

$$E(P) = E_k + \left.\frac{\partial E(P)}{\partial P}\right|_{P=P_k} (P - P_k) + R_1 = E_k + Y_k (P - P_k) + L_1, \tag{12}$$

where L_1 is the first-order Lagrange reminder. The second equality holds from the Envelope Theorem with respect to the expenditure function.[3] Evaluation at P_j yields

$$E(P_j) = E_j \approx Q_k + Y_k (P_j - P_k). \tag{13}$$

We use the optimality at the k-th block, which implies that $E_k = Q_k$. $\square$

There are two additional comments on this approximation. First, the approximation of $E_j - E_k$, $Q_k + Y_k(P_j - P_k) - E_k$, can be interpreted as that of the compensating variation. Suppose Y is supplied with a uniform price P_k and the consumer attains the indirect utility level V. Then, the price changes from P_k to P_j. In this setting, we can define the expenditure minimization problems for P_k and P_j as being identical form to Eq. (8), and E_j becomes the expenditure for a consumer to maintain the indirect utility level at V, which was achieved under the previous price P_k. Then, by definition, the $E_j - E_k$ amount becomes the compensating variation.

Second, in the study of welfare economics, we sometimes find that the expenditure function is approximated by Taylor expansion up to the second order (see, e.g., Eq. 2 of Irvine and Sims 1998). In our case, however, the parameter estimation becomes difficult with such a second-order approximation.

[3] See Proposition 3.G.1 on p.68 of Mas-Colell, Whinston, and Green (1995) for details regarding the Envelope Theorem with respect to the expenditure function.

Thus, this paper considers the first-order approximation with log-linear conditional demand. Section 4 will evaluate this approximation with simulated data.

Approximation Eq. (10) transforms Condition (9) into

$$\begin{cases} E_k = Q_k, & \\ E_j > Q_j, & \text{for } j \neq k, \end{cases} \iff \begin{cases} Y_k > H_j, & \text{if } j < k, \\ Y_k < H_j, & \text{if } j > k, \end{cases} \tag{14}$$

where $H_j = (Q_j - Q_k)/(P_j - P_k)$. Therefore, the demand function under decreasing block rate pricing is given by

$$Y = Y_k, \quad \text{if} \begin{cases} \bar{Y}_{k-1} < Y_k < \bar{Y}_k, & \\ Y_k > H_j, & \text{for } j < k, \\ Y_k < H_j, & \text{for } j > k, \end{cases} \tag{15}$$

where the additional interval condition $\bar{Y}_{k-1} < Y_k < \bar{Y}_k$ is needed to guarantee that the conditional demand is restricted in its corresponding block. Finally, under the log-linear conditional demand assumption,

$$y = y_k, \quad \text{if} \begin{cases} \bar{y}_{k-1} < y_k < \bar{y}_k, & \\ y_k > h_j, & \text{for } j < k, \\ y_k < h_j, & \text{for } j > k, \end{cases} \tag{16}$$

where $y_k = \boldsymbol{x}_k' \boldsymbol{\beta}$ and $h_j = \ln H_j$.

The duality-based approach has the following three advantages. First, in contrast to the conventional approach, nonlinear constraints are all approximated by linear constraints such that it is straightforward to extend the demand function under multiple-block decreasing block rate pricing. Further, the demand function becomes a multinomial extension of the Type II Tobit model. The parameters of the Type II Tobit model can be estimated using the Bayesian approach with the MCMC simulation (see the next section for details). Second, the model is, to some extent, free of the consumer preference change and specific consumer preferences in exchange for this approximation. Consumer preferences that produce nonlinear constraints are approximately included in linear constraints. Thus, the model is relatively independent of consumer preferences, and we have robust parameter estimates of the demand function. Third, in recent years, the nonparametric approach has been proposed by Blomquist and Newey (2002) that is robust with distributional misspecification. Our approach, in contrast, is parametric and has an advantage over their nonparametric approach because model parameters have economic implications in terms of the price and income elasticities.

3 Bayesian Analysis of Duality-Based Demand Functions

3.1 Posterior Distribution

We assume the log conditional demand y_{ik} for each observation i, and it is observed with two additive disturbances: consumer heterogeneity w_i^* and measurement error u_i ($i = 1, \dots, n$).

Consumer heterogeneity is an unobserved variable, which is assumed to have a linear structure, as follows.

$$y_i^* = y_{ik} + w_i^*, \quad w_i^* = \boldsymbol{z}_i' \boldsymbol{\delta} + v_i, \quad v_i \sim \text{i.i.d. } N(0, \sigma_v^2), \tag{17}$$

where $\boldsymbol{z}_i$ and $\boldsymbol{\delta}$ represent a $d \times 1$ explanatory vector and its coefficient vector, respectively, and v_i is a random variable that is independently and identically distributed as the normal distribution with mean 0 and variance σ_v^2. This paper names y_i^* the unobserved demand in contrast to the observed demand y_i.

The role of this heterogeneity w_i^* is explained below. First, it represents the consumer's characteristics. Further, the continuous random variable v_i included in this heterogeneity excludes multiple solutions to the original utility maximization problem, Problem (1), or more precisely, assigns a zero probability for these multiple solutions. Thus, it is sufficient to consider the single-solution case as we referred to in Subsection 2.1.

By introducing consumer heterogeneity, we can interpret the D/C choice model as one where heterogeneity is the fundamental factor in deciding the optimal consumption. In Eq. (16), each block's conditional demands decide the optimal block and demand. After heterogeneity is included in this model, it is heterogeneity that controls the consumer's optimal choice. More precisely, we replace three conditions in Eq. (16) with another equivalent condition using the heterogeneity interval R_{ik} ($k = 1, \dots, K_i$).

$$\begin{cases} \bar{y}_{k-1} < y_i^* < \bar{y}_k, & \\ y_i^* > h_j, & \text{for } j < k, \\ y_i^* < h_j, & \text{for } j > k, \end{cases}$$

$$\iff w_i^* \in R_{ik} = \left\{ \max\left(\bar{y}_{i,k-1}, \max_{j<k} h_{ij}\right) - y_{ik}, \min\left(\bar{y}_{ik}, \min_{j>k} h_{ij}\right) - y_{ik} \right\}, \tag{18}$$

where $h_{ij} = \ln H_{ij} = \ln\{(Q_{ij} - Q_{ik})/(P_{ij} - P_{ik})\}$.

In addition to the heterogeneity, we assume that the demand is observed with the following measurement error.

$$y_i = y_i^* + u_i, \quad u_i \sim \text{i.i.d. } N(0, \sigma_u^2). \tag{19}$$

This error term has two purposes. First, as noted at the end of Section 3 of Moffitt (1986), u_i gives positive probabilities on regions $(RU_{ik}, RL_{i,k+1})$ for $k = 1, \dots, K_i - 1$, which are often nonempty depending on the data. We use (RL_{ik}, RU_{ik}) as the lower and upper bounds of the heterogeneity interval R_{ik} defined above. Second, u_i allows the consumer to select the different block from the one which is indicated by the unobserved demand, y_i^*; that is, while y_i is observed in the l-th block, y_i^* is in the k-th block ($k \neq l$). Because the block choice based on y_i^* is free of measurement error and is ideal, it is natural to allow such a difference between the observed and unobserved block choice.

Finally, we have the model of the demand function under decreasing block rate pricing.

$$\begin{aligned} & y_{ik} = \boldsymbol{x}_{ik}' \boldsymbol{\beta}, \quad k = 1, \dots, K_i, \\ & w_i^* = \boldsymbol{z}_i' \boldsymbol{\delta} + v_i, \quad v_i \sim \text{i.i.d. } N(0, \sigma_v^2), \\ & r_i^* = k, \quad \text{if } w_i^* \in R_{ik} \text{ and } k = 1, \dots, K_i, \\ & y_i = y_i^* + u_i = y_{ir_i^*} + w_i^* + u_i, \quad u_i \sim \text{i.i.d. } N(0, \sigma_u^2), \end{aligned} \tag{20}$$

where r_i^* is a discrete latent variable indicating the block that is selected by observation i depending on the heterogeneity. As stated in the previous subsection, the model falls into a multinomial extension of the Type II Tobit model. The mechanism of this model is summarized as follows: consumer heterogeneity, which is unobserved, determines the optimal block, and hence, the unobserved demand. Given this unobserved demand, we can observe the demand with some measurement error.

Next, we derive the augmented likelihood function for observation i. Because there are two unobserved components, heterogeneity w_i^* and state r_i^*, the likelihood function is augmented with these latent variables and is given by

$$
\begin{aligned}
& f\left(y_i, r_i^*, w_i^* \mid \boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2\right) \\
& \propto (\sigma_u \sigma_v)^{-1} \exp\left[-\frac{1}{2}\left\{\sigma_u^{-2}(y_i - y_i^*)^2 + \sigma_v^{-2}\left(w_i^* - \boldsymbol{z}_i' \boldsymbol{\delta}\right)^2\right\}\right] I\left(w_i^* \in R_{ir_i^*}\right) \prod_{k=1}^{K_i-1} I\left(RU_{ik} \leq RL_{i,k+1}\right),
\end{aligned}
\tag{21}
$$

where $I(A)$ is the indicator function taking the value 1 if A is true, and 0 otherwise.

The last truncation

$$
\prod_{k=1}^{K_i-1} I(RU_{ik} \leq RL_{i,k+1}), \tag{22}
$$

is termed the separability condition, which guarantees a disjoint parameter space of w_i^*, that is, a disjoint heterogeneity interval. When we estimate the demand function under increasing block rate pricing, the separability condition becomes

$$
\prod_{k=1}^{K_i-1} I(y_{i,k+1} \leq y_{ik}), \tag{23}
$$

which is more restrictive than Eq. (22): because $\min(\bar{y}_{ik}, \min_{j>k} h_{ij}) \leq \max(\bar{y}_{ik}, \max_{j<k+1} h_{ij})$, $y_{i,k+1} \leq y_{ik}$ is the sufficient condition for $RU_{ik} \leq RL_{i,k+1}$. Such a difference in above two separability conditions between increasing and decreasing block rate pricing cases comes from the comparison of indirect utilities among blocks that is additional to the demand function under decreasing block rate pricing. In spite of its different restriction from Eq. (23), the role of the separability condition for the decreasing block rate pricing case, Eq. (22), is the same as the one for the increasing block rate pricing case. (see Miyawaki et al. 2006 for a detailed discussion of this condition on the estimation of the demand function under increasing block rate pricing).

Finally, let $\pi(\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2)$ denote the prior density, and we have the posterior distribution of the demand function under decreasing block rate pricing as follows.

$$
\begin{aligned}
\pi\left(\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2, \boldsymbol{r}^*, \boldsymbol{w}^* \mid \boldsymbol{y}\right) \propto \pi\left(\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2\right) & \\
(\sigma_u \sigma_v)^{-n} \exp\left[-\frac{1}{2}\left\{\sigma_u^{-2}(\boldsymbol{y} - \boldsymbol{y}^*)'(\boldsymbol{y} - \boldsymbol{y}^*) + \sigma_v^{-2}(\boldsymbol{w}^* - \boldsymbol{Z}\boldsymbol{\delta})'(\boldsymbol{w}^* - \boldsymbol{Z}\boldsymbol{\delta})\right\}\right] & \\
\prod_{i=1}^{n} I\left(w_i^* \in R_{ir_i^*}\right) \prod_{k=1}^{K_i-1} I\left(RU_{ik} \leq RL_{i,k+1}\right), &
\end{aligned}
\tag{24}
$$

where $\boldsymbol{y} = (y_1, y_2, \ldots, y_n)'$, $\boldsymbol{y}^* = (y_1^*, y_2^*, \ldots, y_n^*)'$, $\boldsymbol{r}^* = (r_1^*, r_2^*, \ldots r_n^*)'$, $\boldsymbol{w}^* = (w_1^*, w_2^*, \ldots w_n^*)'$ and $\boldsymbol{Z} = (\boldsymbol{z}_1, \boldsymbol{z}_2, \ldots, \boldsymbol{z}_n)'$.

Remark 2. The alternative functional form of the conditional demands is the following linear function.

$$Y_{ik} = \beta_1 P_{ik} + \beta_2 Q_{ik}. \tag{25}$$

When this linear function is used, Taylor expansion of the conditional expenditure function up to the third order yields the exact equation.

$$E_{ij} = Q_{ik} + Y_{ik}(P_{ij} - P_{ik}) + \frac{1}{2}(\beta_1 + \beta_2 Y_{ik})(P_{ij} - P_{ik})^2 + \frac{1}{6}\beta_1\beta_2(P_{ij} - P_{ik})^3. \tag{26}$$

Further, instead of the interval condition and separability condition, we need the positivity condition given below.

$$Y_{ik} + w_i^* \geq 0 \quad \text{for } k = 1, \ldots, K_i,\ i = 1, \ldots, n. \tag{27}$$

The posterior distribution remains identical to Eq. (24) when we redefine the heterogeneity interval R_{ik} in accordance with Eqs. (26) and (27).

3.2 Gibbs Sampler

We assume the proper prior densities on model parameters $(\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2)$ such that $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ conditional on σ_u^2 and σ_v^2 follow multivariate normal distributions, and σ_u and σ_v follow inverse gamma distributions. More precisely,

$$\begin{aligned} &\boldsymbol{\beta} \mid \sigma_u^2 \sim N_2\left(\boldsymbol{\mu}_{\boldsymbol{\beta}_j,0}, \sigma_u^2 \boldsymbol{\Sigma}_{\boldsymbol{\beta},0}\right), \quad \boldsymbol{\delta} \mid \sigma_v^2 \sim N_d\left(\boldsymbol{\mu}_{\boldsymbol{\delta},0}, \sigma_v^2 \boldsymbol{\Sigma}_{\boldsymbol{\delta},0}\right), \\ &\sigma_u^2 \sim IG\left(\frac{n_{u,0}}{2}, \frac{S_{u,0}}{2}\right), \quad \sigma_v^2 \sim IG\left(\frac{n_{v,0}}{2}, \frac{S_{v,0}}{2}\right), \end{aligned} \tag{28}$$

where $\boldsymbol{\mu}_{\boldsymbol{\beta}_j,0}$ is a 2×1 known vector; $\boldsymbol{\Sigma}_{\boldsymbol{\beta},0} = \text{diag}(\sigma^2_{\beta_1,0}, \sigma^2_{\beta_1,0})$ is a known diagonal 2×2 covariance matrix; p$\boldsymbol{\mu}_{\boldsymbol{\delta},0}$ is a $d \times 1$ known vector; $\boldsymbol{\Sigma}_{\boldsymbol{\delta},0}$ is a known $d \times d$ covariance matrix; and $n_{v,0}$, $S_{v,0}$, $n_{u,0}$, and $S_{u,0}$ are some known positive constants. The subscript to the normal distribution indicates its dimension.

Then, we can implement the MCMC simulation by the Gibbs sampler in the following seven steps.

Algorithm 1: MCMC algorithm for the cross-section model

Step 1. Initialize $\boldsymbol{\beta}, \boldsymbol{\delta}, \boldsymbol{r}^*, \boldsymbol{w}^*, \sigma_u^2$, and σ_v^2.

Step 2. Generate β_1 given $\beta_2, \boldsymbol{r}^*, \boldsymbol{w}^*, \sigma_u^2$.

Step 3. Generate β_2 given $\beta_1, \boldsymbol{r}^*, \boldsymbol{w}^*, \sigma_u^2$.

Step 4. Generate $(\sigma_v^2, \boldsymbol{\delta})$ given $\boldsymbol{w}^*$.

- (a) Generate σ_v^2 given $\boldsymbol{w}^*$.
- (b) Generate $\boldsymbol{\delta}$ given $\boldsymbol{w}^*, \sigma_v^2$.

Step 5. Generate (r_i^*, w_i^*) given $\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2$ for $i = 1, \ldots, n$.

- (a) Generate r_i^* given $\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2$.
- (b) Generate w_i^* given $\boldsymbol{\beta}, \boldsymbol{\delta}, r_i^*, \sigma_u^2, \sigma_v^2$.

Step 6. Generate σ_u^2 given $\boldsymbol{\beta}, \boldsymbol{r}^*, \boldsymbol{w}^*$.

Step 7. Go to Step 2.

In Step 4, the blocking technique in $(\sigma_v^2, \boldsymbol{\delta})$ is used to accelerate the convergence of MCMC samples to their posterior distribution. On the other hand, in Step 5, we again use the blocking technique to avoid the relationship where w_i^* determines r_i^* as well as to obtain the convergence acceleration for MCMC samples to their posterior distribution. All full conditional distributions are standard distributions, and are provided in the next subsections following each steps of **Algorithm 1**.

In order to explain full conditionals, we introduce two notations and one assumption only for this Section 3.2. First, $TN_A(\mu, \sigma^2)$ denotes a truncated normal distribution with mean μ and variance σ^2 and truncated region A. Second, we define a new nonnegative variable $\bar{y}_{ik}^+$ $(k = 1, \ldots, K_i - 1)$ related to thresholds.

$$\bar{y}_{ik}^+ = \max\left(\bar{y}_{ik}, \max_{l<k+1} h_{il}\right) - \min\left(\bar{y}_{ik}, \min_{l>k} h_{il}\right). \tag{29}$$

Thirdly, without loss of generality, we assume p_{iK_i} and q_{iK_i} for all i to be strictly positive.

3.2.1 Step 2. Generate β_1 given $\beta_2, \boldsymbol{r}^*, \boldsymbol{w}^*, \sigma_u^2$.

Draw β_1 from the truncated normal distribution $TN_{R_1}(\mu_{\beta_1,1}, \sigma_u^2\sigma_{\beta_1,1}^2)$, where

$$\begin{aligned}
\sigma_{\beta_1,1}^{-2} &= \sigma_{\beta_1,0}^{-2} + \sum_{i=1}^{n} \left(p_{ir_i^*}\right)^2, \\
\mu_{\beta_1,1} &= \sigma_{\beta_1,1}^2 \left[\sigma_{\beta_1,0}^{-2}\mu_{\beta_1,0} + \sum_{i=1}^{n} p_{ir_i^*}\left(y_i - \beta_2 q_{ir_i^*} - w_i^*\right)\right], \\
R_1 &= \left\{\max_{i,k}\left(\frac{-\beta_2\left(q_{i,k+1} - q_{ik}\right) + \bar{y}_{ik}^+}{p_{i,k+1} - p_{ik}}, RL_{1i}\right), \min_i\left(\infty, RU_{1i}\right)\right\}.
\end{aligned} \tag{30}$$

The RL_{1i} and RU_{1i} denote the lower and upper intervals of R_{1i}, respectively, which is defined below.

$$R_{1i} = \left\{\frac{\max\left(\bar{y}_{i,r_i^*-1}, \max_{l<r_i^*} h_{il}\right) - \beta_2 q_{ir_i^*} - w_i^*}{p_{ir_i^*}}, \frac{\min\left(\bar{y}_{ir_i^*}, \min_{l>r_i^*} h_{il}\right) - \beta_2 q_{ir_i^*} - w_i^*}{p_{ir_i^*}}\right\}. \tag{31}$$

3.2.2 Step 3. Generate β_2 given $\beta_1, \boldsymbol{r}^*, \boldsymbol{w}^*, \sigma_u^2$.

Draw β_2 from the truncated normal distribution $TN_{R_2}(\mu_{\beta_2,1}, \sigma_u^2\sigma_{\beta_2,1}^2)$, where

$$\begin{aligned}
\sigma_{\beta_2,1}^{-2} &= \sigma_{\beta_2,0}^{-2} + \sum_{i=1}^{n} \left(q_{ir_i^*}\right)^2, \\
\mu_{\beta_2,1} &= \sigma_{\beta_2,1}^2 \left[\sigma_{\beta_2,0}^{-2}\mu_{\beta_2,0} + \sum_{i=1}^{n} q_{ir_i^*}\left(y_i - \beta_1 p_{ir_i^*} - w_i^*\right)\right], \\
R_2 &= \left\{\max_{i,k}\left(\frac{-\beta_1\left(p_{i,k+1} - p_{ik}\right) + \bar{y}_{ik}^+}{q_{i,k+1} - q_{ik}}, RL_{2i}\right), \min_i\left(\infty, RU_{2i}\right)\right\}.
\end{aligned} \tag{32}$$

The RL_{2i} and RU_{2i} are the lower and upper intervals of R_{2i}, respectively, which is defined below.

$$R_{2i} = \left\{\frac{\max\left(\bar{y}_{i,r_i^*-1}, \max_{l<r_i^*} h_{il}\right) - \beta_1 p_{ir_i^*} - w_i^*}{q_{ir_i^*}}, \frac{\min\left(\bar{y}_{ir_i^*}, \min_{l>r_i^*} h_{il}\right) - \beta_1 p_{ir_i^*} - w_i^*}{q_{ir_i^*}}\right\}. \tag{33}$$

Remark 3. The full conditional distributions utilized in Steps 2 and 3 are both interchangeable with each other, when we replace the regression coefficients (β_j, $j = 1,2$), regressors (the price and virtual income), and hyperparameters of prior densities.

3.2.3 Step 4. Generate $\left(\sigma_v^2, \boldsymbol{\delta}\right)$ given $\boldsymbol{w}^*$.

The blocking strategy is adopted to generate samples of $(\sigma_v^2, \boldsymbol{\delta})$. Draw σ_v^2 from $IG(\frac{n_{v,1}}{2}, \frac{S_{v,1}}{2})$, and $\boldsymbol{\delta}$ from $N(\boldsymbol{\mu}_{\boldsymbol{\delta},1}, \sigma_v^2 \boldsymbol{\Sigma}_{\boldsymbol{\delta},1})$, where $n_{v,1} = n_{v,0} + n$,

$$\begin{aligned}
&\boldsymbol{\Sigma}_{\boldsymbol{\delta},1}^{-1} = \boldsymbol{\Sigma}_{\boldsymbol{\delta},0}^{-1} + \boldsymbol{Z}'\boldsymbol{Z}, \\
&\boldsymbol{\mu}_{\boldsymbol{\delta},1} = \boldsymbol{\Sigma}_{\boldsymbol{\delta},1}\left(\boldsymbol{\Sigma}_{\boldsymbol{\delta},0}^{-1}\boldsymbol{\mu}_{\boldsymbol{\delta},0} + \boldsymbol{Z}'\boldsymbol{w}^*\right), \\
&S_{v,1} = S_{v,0} + \boldsymbol{\mu}_{\boldsymbol{\delta},0}'\boldsymbol{\Sigma}_{\boldsymbol{\delta},0}^{-1}\boldsymbol{\mu}_{\boldsymbol{\delta},0} + \boldsymbol{w}^{*\prime}\boldsymbol{w}^* - \boldsymbol{\mu}_{\boldsymbol{\delta},1}'\boldsymbol{\Sigma}_{\boldsymbol{\delta},1}^{-1}\boldsymbol{\mu}_{\boldsymbol{\delta},1}.
\end{aligned} \tag{34}$$

3.2.4 Step 5. Generate (r_i^*, w_i^*) given $\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2$ for $i = 1, \ldots, n$.

By marginalizing over w_i^*, we have the full conditional distribution for r_i^* as the multinomial distribution, which is given by

$$\pi\left(r_i^* = r \mid \boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2\right) \propto \tau\left[\Phi\left\{\tau^{-1}\left(RU_{ir} - \theta_{ir}\right)\right\} - \Phi\left\{\tau^{-1}\left(RL_{ir} - \theta_{ir}\right)\right\}\right] \exp\left(-\frac{m_{ir}}{2}\right), \tag{35}$$

for $s = 1, \ldots, K_i$ where $\tau^2 = (\sigma_u^{-2} + \sigma_v^{-2})^{-1}$. The pair (m_{ir}, θ_{ir}) are given by

$$(m_{ir}, \theta_{ir}) = \left(\frac{(\sigma_u\sigma_v)^{-2}\left(y_i - y_{ir} - \boldsymbol{z}_i'\boldsymbol{\delta}\right)^2}{\sigma_u^{-2} + \sigma_v^{-2}}, \frac{\sigma_u^{-2}\left(y_i - y_{ir}\right) + \sigma_v^{-2}\boldsymbol{z}_i'\boldsymbol{\delta}}{\sigma_u^{-2} + \sigma_v^{-2}}\right). \tag{36}$$

Given $r_i^* = r$, it is straightforward to derive the full conditional distribution for w_i^*, which is the truncated normal distribution $TN_{R_{ir}}(\theta_{ir}, \tau^2)$.

3.2.5 Step 6. Generate σ_u^2 given $\boldsymbol{\beta}, \boldsymbol{r}^*, \boldsymbol{w}^*$.

The full conditional distribution for σ_u^2 is $IG(\frac{n_{u,1}}{2}, \frac{S_{u,1}}{2})$, where $n_{u,1} = n_{u,0} + 2 + n$ and

$$S_{u,1} = S_{u,0} + \left(\boldsymbol{\beta} - \boldsymbol{\mu}_{\boldsymbol{\beta},0}\right)'\boldsymbol{\Sigma}_{\boldsymbol{\beta},0}^{-1}\left(\boldsymbol{\beta} - \boldsymbol{\mu}_{\boldsymbol{\beta},0}\right) + \left(\boldsymbol{y} - \boldsymbol{y}^*\right)'\left(\boldsymbol{y} - \boldsymbol{y}^*\right). \tag{37}$$

3.3 Predictive Distribution

As analyzed in Section 8 of Reiss and White (2005), the estimated demand function can be utilized to predict future demand according to the change in the price system. The ordinary Bayesian tool for prediction is the predictive distribution, namely, a conditional distribution of future dependent variables given the presently observed data and model, and marginalized over the model parameters. This subsection describes the predictive distribution of our D/C choice model, and its sampling algorithm based on the Gibbs sampler derived in the previous subsection.

Let $\boldsymbol{\zeta} = (\boldsymbol{\beta}, \boldsymbol{\delta}, \sigma_u^2, \sigma_v^2)$, $\boldsymbol{\Psi}^* = (\boldsymbol{r}^*, \boldsymbol{w}^*)$, and let $(\tilde{\boldsymbol{y}}, \tilde{\boldsymbol{\Psi}}^*)$ be the predicted values of demand and latent variables, respectively. Then, the predictive distribution is given by

$$\begin{aligned} f(\tilde{\boldsymbol{y}} \mid \boldsymbol{y}) &= \int f\left(\tilde{\boldsymbol{y}} \mid \boldsymbol{y}, \tilde{\boldsymbol{\Psi}}^*, \boldsymbol{\zeta}, \boldsymbol{\Psi}^*\right) g\left(\tilde{\boldsymbol{\Psi}}^* \mid \boldsymbol{y}, \boldsymbol{\zeta}, \boldsymbol{\Psi}^*\right) \pi(\boldsymbol{\zeta}, \boldsymbol{\Psi}^* \mid \boldsymbol{y}) \, d\tilde{\boldsymbol{\Psi}}^* d\boldsymbol{\zeta} d\boldsymbol{\Psi}^* \\ &= \int f\left(\tilde{\boldsymbol{y}} \mid \tilde{\boldsymbol{\Psi}}^*, \boldsymbol{\beta}, \sigma_u^2\right) g\left(\tilde{\boldsymbol{\Psi}}^* \mid \boldsymbol{y}, \boldsymbol{\zeta}\right) \pi(\boldsymbol{\zeta}, \boldsymbol{\Psi}^* \mid \boldsymbol{y}) \, d\tilde{\boldsymbol{\Psi}}^* d\boldsymbol{\zeta} d\boldsymbol{\Psi}^*, \end{aligned} \tag{38}$$

where

$$f\left(\tilde{\boldsymbol{y}} \mid \tilde{\boldsymbol{\Psi}}^*, \boldsymbol{\beta}, \sigma_u^2\right) = \prod_{i=1}^{n} N\left(y_{i\tilde{r}_i^*} + \tilde{w}_i^*, \sigma_u^2\right), \tag{39}$$

$$g\left(\tilde{\boldsymbol{\Psi}}^* \mid \boldsymbol{y}, \boldsymbol{\zeta}\right) \propto \prod_{i=1}^{n} \exp\left[-\frac{1}{2}\left\{\sigma_u^{-2}\left(y_i - y_{i\tilde{r}_i^*} - \tilde{w}_i^*\right)^2 + \sigma_v^{-2}\left(\tilde{w}_i^* - \boldsymbol{z}_i^* \boldsymbol{\delta}\right)^2\right\}\right] I\left(\tilde{w}_i^* \in R_{i\tilde{r}_i^*}\right), \tag{40}$$

and $\pi(\boldsymbol{\zeta}, \boldsymbol{\Psi}^* \mid \boldsymbol{y})$ is the posterior distribution, Eq. (24). The first equality of Eq. (38) is the definition of the predictive distribution, and the second follows from the D/C choice model. We point out that Eqs. (39) and (40) are evaluated at a new price system, and the latter conditional distribution, Eq. (40), is proportional to the product of the full conditional distribution of (r_i^*, w_i^*) used in Step 5 of **Algorithm 1** over all observations.

In order to obtain samples from this predictive distribution, Section 9 of Chib (2001) suggests the method of composition that makes use of the MCMC samples previously drawn by the Gibbs sampler described in the previous subsection. The method of composition in our case is implemented in the following four steps.

Algorithm 2: sampling from predictive distribution
Step 1. Pick $\boldsymbol{\zeta}^{(j)} = (\boldsymbol{\beta}^{(j)}, \boldsymbol{\delta}^{(j)}, \sigma_u^{2(j)}, \sigma_v^{2(j)})$, the j-th MCMC sample drawn from $\pi(\boldsymbol{\zeta}, \boldsymbol{\Psi}^* \mid \boldsymbol{y})$.
Step 2. Generate $(\tilde{r}_i^*, \tilde{w}_i^*)$ given $\boldsymbol{\zeta}^{(j)}$ for $i = 1, \ldots, n$.

(a) Generate $\tilde{r}_i^*$ given $\boldsymbol{\zeta}^{(j)}$.
(b) Generate $\tilde{w}_i^*$ given $\boldsymbol{\zeta}^{(j)}$ and $\tilde{r}_i^*$.

Step 3. Generate $\tilde{y}_i$ given $\tilde{r}_i^*$, $\tilde{w}_i^*$, $\boldsymbol{\beta}^{(j)}$, $\sigma_u^{2(j)}$ for $i = 1, \ldots, n$.
Step 4. Set $j \to j+1$ and go to Step 1.

To generate $(\tilde{r}_i^*, \tilde{w}_i^*)$ and $\tilde{y}_i$, we use the same distributions used in Step 5 of **Algorithm 1** and normal distribution $N(y_{i\tilde{r}_i^*} + \tilde{w}_i^*, \sigma_u^2)$, respectively.

4 Numerical Example Using Simulated Data

This section presents a numerical example to illustrate our estimation procedure using the simulated data and evaluates the accuracy of our approximation.

We generate 100 observations under two-block decreasing block rate pricing. For a single threshold, we set $\bar{Y}_{i1} = 3.5$ for all observations. As for price, the second block's price is distributed as $|N(0.9, 0.3^2)|$, the absolute value of a normal random number with mean 0.9 and variance 0.3^2. The price of the first block is generated by adding $|N(0.3, 0.2^2)|$ to that of the second block. The basic connection charge is

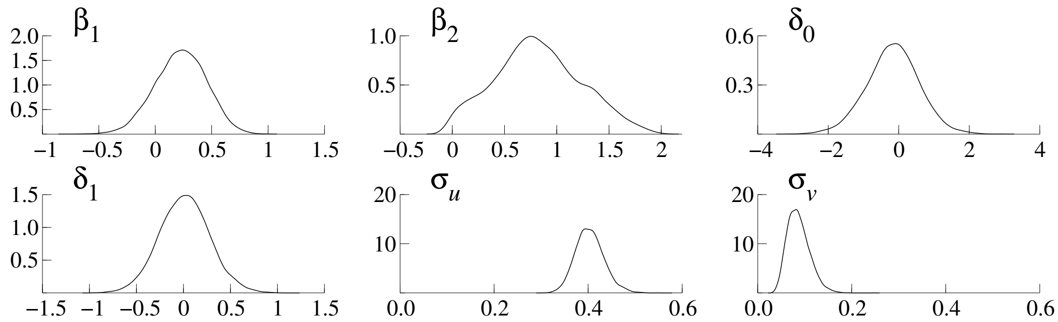

Figure 3: Estimated Marginal Posterior Distributions with Simulated Data.

Table 1: Results with Simulated Data*

Parameter	True	Mean	SD	95% interval		$p>0$	INEF	CD
β_1	$-.1$	.23	.23	[$-$ .23	.65]	.84	4.7	.254
β_2	.5	.82	.41	[.052	1.65]	1.00	176.3	.166
δ_0 (constant)	.1	$-$ .16	.75	[$-$1.68	1.29]	.42	6.3	.079
δ_1	.1	.013	.27	[$-$ .51	.55]	.52	36.8	.243
σ_u (measurement error)	.4	.40	.030	[.35	.47]	—	1.0	.107
σ_v (heterogeneity)	.1	.087	.024	[.049	.14]	—	1.8	.640

* "$p>0$", "INEF", and "CD" denote the marginal posterior probability above zero, inefficiency factor, and convergence diagnostic, respectively.

considered to be zero for simplicity. We generate the income I_i and explanatory variable z_{i2} for heterogeneity, using $I_i \sim |N(3.0, 0.3^2)|$ and $\boldsymbol{z}_i' = (z_{i1}, z_{i2}) = \left(1.0, |N(3.0, 0.1^2)|\right)$, respectively.

The prior densities are assumed to be as follows.

$$\begin{aligned} &\boldsymbol{\beta} \mid \sigma_u^2 \sim N_2\left(\mathbf{0}, 100\sigma_u^2 \boldsymbol{I}\right), \quad \boldsymbol{\delta} \mid \sigma_v^2 \sim N_2\left(\mathbf{0}, 100\sigma_v^2 \boldsymbol{I}\right), \\ &\sigma_u^2 \sim IG(0.01, 0.01), \quad \sigma_v^2 \sim IG(0.01, 0.01). \end{aligned} \tag{41}$$

After deleting 3×10^4 samples as the burn-in period, the subsequent 3×10^5 MCMC samples are drawn by **Algorithm 1**, and every 30-th sample is selected to obtain 10^4 samples, which are used for Bayesian inferences. The results are found in Figure 3 and Table 1. Table 1 reports the true parameter value, estimated posterior mean, posterior standard deviation, 95% credible interval, marginal posterior probability above zero, inefficiency factor, and convergence diagnostic.

The inefficiency factor is an index defined as $1 + 2\sum_{j=1}^{\infty} \rho(j)$, where $\rho(j)$ is the sample autocorrelation at lag j, and indicates the estimated loss of MCMC samples as compared to independent ones (see Section 3.2 of Chib 2001). When the inefficiency factor is close to one, sampling is almost as efficient as independent draw. The convergence diagnostic, on the other hand, is the p-value of the test statistic that tests the equality of two means of the first n_A and last n_B draws (see Subsection 3.2 of Geweke 1992). If the null hypothesis that two means are equal is rejected, we conclude that there is a significant evidence against the convergence of MCMC samples to their posterior distribution. To calculate the convergence

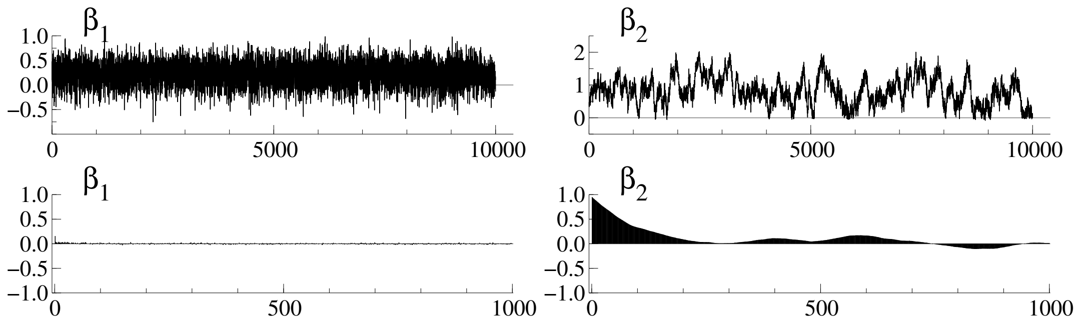

Figure 4: Sample Paths (upper) and Sample Autocorrelation Functions (lower) for β_1 and β_2.

diagnostic, we take the first 10% for n_A and the last 50% for n_B of MCMC samples, respectively, as suggested by Geweke (1992).

From Figure 3 and Table 1, we can observe that our estimation method effectively estimates the true values because true parameter values are included around the modes of marginal posterior distributions. There are two other features in our MCMC estimation. First, the MCMC samples of β_2 converge slowly to its marginal posterior distribution. Figure 4 shows the sample paths and sample autocorrelation functions for β_1 and β_2. This figure reveals that β_2's sample path is more correlated and its sample autocorrelation function decays slower than those of β_1. Furthermore, the inefficiency factor for β_2 is much larger than those of the other parameters. Thus, we conclude that β_2 slowly converges to its marginal posterior distribution. A similar tendency was also found in the estimation of the demand function under increasing block rate pricing (see Table 1 of Miyawaki et al. 2006). One possible reason for this slow convergence is the use of the virtual income. Further analysis of the virtual income to accelerate the convergence of the MCMC chain to its posterior distribution would be necessary in future work.

Second, two variance parameters (σ_u^2, σ_v^2) are identified in this numerical example. As explained in the previous section, when, for example, heterogeneity intervals are separated from each other, that is, when $(RU_{ik}, RL_{i,k+1})$ $(k = 1, \ldots, K_i - 1)$ is nonempty, u_i has its own information from the data so that variance parameters can be identified. In some cases, however, the data seem to fail in providing sufficient information for u_i in order to identify these two variance parameters, and the estimates of these variance parameters are strongly influenced by their prior densities. Thus, we should pay careful attention to whether the data set includes information on the variance parameters to be identified with each other.

At the end of this numerical example, we evaluate the accuracy of the expenditure function approximation, Eq. (10). Hausman (1981) derived the exact expenditure function recovered from the log-linear demand (see Eq. 22). Using Hausman (1981)'s equation, we compute the approximation loss of $E_{ij} - E_{ik}$ $(j \neq k,\ j = 1, 2, \ldots, K_i$ and $i = 1, 2, \ldots, 100)$ where the k-th block is the true selected block calculated from the true heterogeneity. As noted in Subsection 2.3, the difference $E_{ij} - E_{ik}$ is virtually interpreted as being the compensating variation in the hypothetical uniform price system so that the approximation error of this difference is regarded as that of the compensating variation.

Table 2: Gas Service in Japan

Gas type	No. of companies	Rate admission	Consumption percentage*
natural gas	around 200	regulated	19%
LP gas	over $20,000$	not regulated	13%

* Percentage in total energy use measured by Joule in 2004.

The percentage of this approximation loss for observation i is given by

$$\frac{(E_{ij}-E_{ik})-\left\{Q_{ik}+Y_i^*(P_{ij}-P_{ik})-E_{ik}\right\}}{Q_{ik}+Y_i^*(P_{ij}-P_{ik})-E_{ik}}\times 100=\frac{E_{ij}-\left\{Q_{ik}+Y_i^*(P_{ij}-P_{ik})\right\}}{Y_i^*(P_{ij}-P_{ik})}\times 100, \tag{42}$$

where E_{ij} is the exact expenditure derived by Hausman (1981) and $Y_i^*=\exp(y_i^*)=\exp(y_{ik}+w_i^*)$. We use the optimality condition at the k-th block that $E_{ik}=Q_{ik}$. With $u_i=v_i=0$ for simplicity, our approximation loss is calculated as around -7.55 percent on average with a standard deviation of 2.17 percent. Irvine and Sims (1998), on the other hand, provide another numerical example with the linear demand function in their Subsection II-B, and their error of the compensating variation with Taylor expansion up to the second order is calculated as -2.67 percent. Thus, we expect that the estimation with our approximated duality approach yields the relatively acceptable precision as compared to that with the approximated compensating variation used in welfare economics.

5 Empirical Analysis

5.1 Data Description

First, we briefly describe the gas service in Japan. The residential gas service is supplied by regionally monopolistic companies and is categorized into two by type of gas: natural gas and liquefied petroleum (LP) gas. While natural gas is supplied through pipe to each customer, LP gas is stored in a gas cylinder and is delivered to each consumer by truck. Thus, because of the large fixed cost to construct gas pipes, natural gas services are limited to city areas. With regard to price, the price per one cubic meter of natural gas is likely to be higher than that of LP gas, because a calorie of natural gas per one cubic meter is higher than that of LP gas. The other characteristics of companies dealing in each gas type are summarized in Table 2.

Next, the data set of residential gas demand in Japan is explained for the empirical analysis in the next subsection. This paper uses aggregate data collected at the capitals of each Japanese prefecture for the year 1999. The number of observations is 49. For the dependent variable y_i, we select the average gas charge of one household in one month taken from the Family Income and Expenditure Survey (FIES) reported by the Ministry of Internal Affairs and Communications (MIAC). Gas charge is reported in the unit of yen, so that we transform the gas charge into the amount with the unit of cubic meter by applying the corresponding price tables, which are explained in the next paragraph.

The attributes of explanatory variables are found in Table 3. For the price and income, we choose them in the following manner. We pick up each prefecture's price table for residential gas use from the

Table 3: Explanatory Variables used in Gas Demand Function

Variable	Coefficient	Data
price	β_1	price for residential use (¥$10^3/\mathrm{m}^3$)
income		living expenditure (¥10^3)
fixed cost		basic connection charge (¥10^3)
variables for w_i^*	δ_1	average no. of members (persons)
	δ_2	average no. of earners (persons)
	δ_3	average age of household head (10^2 years old)
	δ_4	average floor space ($10\mathrm{m}^3$)
	δ_5	average temperature (°C)

Gas Industry Manual published by the Japan Gas Association. Because it is often observed that there exit several gas companies in one prefecture, we select the price table used by the gas company that includes the prefecture's capital as its supply area. Then, all price tables follow decreasing block rate pricing with two or three blocks in 1999.[4] For the income variable, data taken from the FIES on the average living expenditure of one household in one month is used.

As regards the explanatory variables for heterogeneity, five variables found in Table 3 are selected. These variables are selected such that they reflect the household characteristics and environments of each prefecture. The average number of members and earners and the age of the household head are taken from the FIES, the average floor space is taken from the Housing and Land Survey reported by the MIAC, and the average temperature is reported online by the Japanese Meteorological Agency. Because the Housing and Land Survey is reported every five years, we use 1998 data instead.

5.2 Estimation of the Japanese Residential Gas Demand Function

The same prior densities as the ones used in the previous numerical example are assumed.

$$\begin{aligned} &\boldsymbol{\beta} \mid \sigma_u^2 \sim N_2\left(\mathbf{0}, 100\sigma_u^2\boldsymbol{I}\right), \quad \boldsymbol{\delta} \mid \sigma_v^2 \sim N_6\left(\mathbf{0}, 100\sigma_v^2\boldsymbol{I}\right), \\ &\sigma_u^2 \sim IG(0.01, 0.01), \quad \sigma_v^2 \sim IG(0.01, 0.01). \end{aligned} \tag{43}$$

Then, with the log-linear conditional demand model, we estimate the Japanese residential gas demand function by applying **Algorithm 1**. After discarding 4×10^5 samples, we draw 10^6 samples, and reduce them to 10^4 samples by picking up every 100-th value. The results are found in Figure 5 and Table 4, and five points are set forth with respect to these results.

First, the income elasticity β_2 is estimated to be 0.54 in its posterior mean and is positive because its 95% credible interval does not include 0. Economic theory expects the income elasticity be positive in most cases, which is consistent with our results. In contrast to the income elasticity, the price elasticity includes 0 in its 95% credible interval, and we can conclude that the price has no effect on gas demand.

We observe that the marginal posterior density of β_1 has a sharp peak at 0. One possible reason for

[4]Only five companies used two-block pricing in this year.

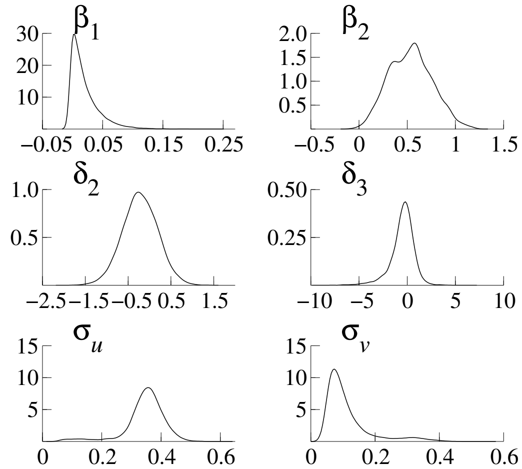

Figure 5: Estimated Marginal Posterior Distributions.

Table 4: Gas Demand Function*

Parameter	Mean	SD	95% interval		$p>0$	INEF	CD
β_1 (price)	.020	.024	[− .004	.084]	.85	2.50	.134
β_2 (income)	.54	.22	[.13	.98]	1.00	185.10	.016
δ_0 (constant)	− .15	1.07	[−2.48	1.98]	.44	42.25	.018
δ_1 (no. of members)	.13	.41	[− .65	.95]	.62	34.31	.016
δ_2 (no. of earners)	− .22	.41	[−1.04	.59]	.30	3.77	.097
δ_3 (age of household head)	− .50	1.24	[−3.67	1.50]	.35	9.81	.046
δ_4 (floor space)	.057	.035	[− .013	.13]	.95	.69	.174
δ_5 (temperature)	.033	.027	[− .021	.085]	.90	1.22	.783
σ_u (measurement error)	.35	.070	[.12	.45]	—	2.99	.420
σ_v (heterogeneity)	.11	.072	[.046	.34]	—	5.82	.206

* "$p>0$", "INEF", and "CD" denote the marginal posterior probability above zero, inefficiency factor, and convergence diagnostic, respectively.

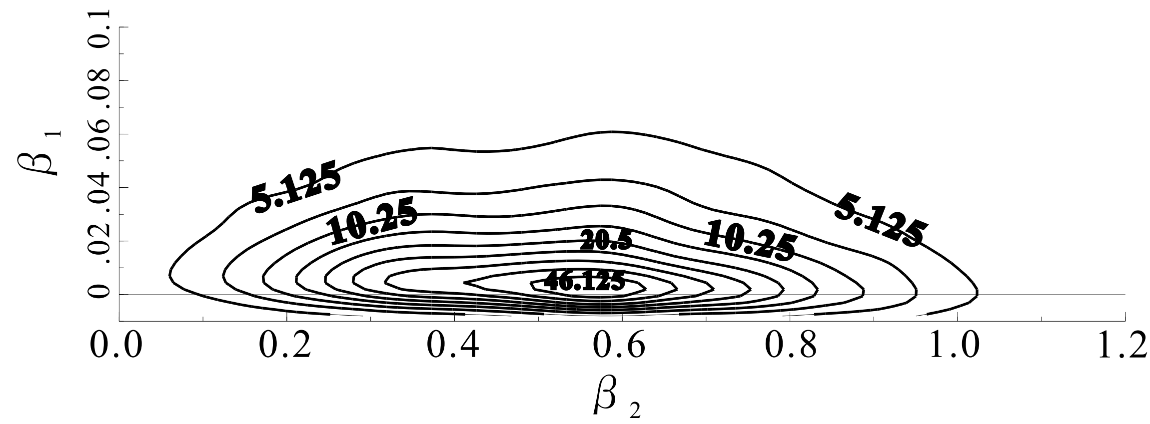

Figure 6: Contour Plot of Join Density (β_1, β_2).

this peak is the separability condition, which is a set of linear constraints on $\boldsymbol{\beta}$'s parameter space. The β_1's marginal posterior density shows such form probably because our gas data set imposes stronger constraints on β_1 than on β_2 through the separability condition (see also Figure 6). Later in this section, the regression model without block choice, which is inevitably a model without the separability condition, is estimated. The estimation results of this unrestrained model demonstrate how this condition influences the elasticity parameters, while they suffer from the model misspecification bias in the sense that block choice is ignored.

Second, as shown in the numerical example described in the previous section, the MCMC samples of β_2 converge slowly to its posterior distribution. The inefficiency factor of β_2 is the highest among those of other parameters. Figure 6 draws a contour plot of the joint density of (β_1, β_2). From this figure, we can see that this joint density has a very steep slope in the south, which causes the slow convergence of the MCMC chain to its posterior distribution. The improvement of such a deficiency in our algorithm would constitute a topic for future research.

Third, no variable for heterogeneity has a positive or negative effect on gas demand in terms of its 95% credible interval. The probability of the marginal posterior distribution above 0, however, indicates that the average floor space (δ_4) has a positive effect on gas demand with the probability close to 0.95. (Table 4 reports rounded values, and this probability with three decimal places is 0.946.) In Japan, one of the main types of residential gas uses is a hot water supply for bathrooms. It is possible that the average floor space would become a proxy for this hot water supply such that the average floor space has a positive effect on gas demand.

Fourth, we briefly discuss the model misspecification problem. One misspecified model is a regression model with a known block choice and is given by

$$\begin{aligned} r_i^* &= k, \quad \text{if } \bar{y}_{i,k-1} < y_i \leq \bar{y}_{ik}, \\ y_i &= \tilde{\boldsymbol{x}}_{ir_i^*}' \tilde{\boldsymbol{\beta}} + \tilde{u}_i, \quad \tilde{u}_i \sim \text{i.i.d. } N(0, \tilde{\sigma}_u^2), \end{aligned} \tag{44}$$

where $\tilde{\boldsymbol{x}}_{ik} = (p_{ik}, q_{ik}, \boldsymbol{z}_i')'$ and $\tilde{\boldsymbol{\beta}} = (\beta_1, \beta_2, \boldsymbol{\delta}')'$. In this model, the optimal block that the consumer chooses is assumed to be observable, and the latent variable r_i^* becomes no longer unobservable. With the prior

Table 5: Gas Demand Function with Regression Model*

Parameter	Mean	SD	95% interval		$p>0$
$\tilde{\beta}_1$ (price)	$-.90$	.052	[-1.01	$-.80$]	.00
$\tilde{\beta}_2$ (income)	.46	.17	[.13	.81]	1.00
$\tilde{\beta}_3$ (constant)	-1.69	1.02	[-3.68	.33]	.05
$\tilde{\beta}_4$ (no. of members)	.30	.18	[$-.052$	.65]	.95
$\tilde{\beta}_5$ (no. of earners)	$-.065$	.18	[$-.42$	.29]	.36
$\tilde{\beta}_6$ (age of household head)	$-.85$	.93	[-2.67	.97]	.18
$\tilde{\beta}_7$ (floor space)	.009	.013	[$-.017$	.035]	.75
$\tilde{\beta}_8$ (temperature)	.011	.011	[$-.009$	.032]	.86
$\tilde{\sigma}_u$ (error)	.13	.014	[.11	.16]	—

* "$p>0$" denotes the marginal posterior probability above zero.

Table 6: Summary Statistics of Actual Log Demand in 1999

Variable	Mean	SD	Minimum	Maximum
Actual Log Demand, y_i, ($\log \mathrm{m}^3$)	3.84	.38	2.77	4.53

densities,

$$\tilde{\boldsymbol{\beta}} \mid \tilde{\sigma}_u^2 \sim N_8\left(\mathbf{0}, 100\tilde{\sigma}_u^2 \boldsymbol{I}\right), \quad \tilde{\sigma}_u^2 \sim IG(0.01, 0.01), \tag{45}$$

the estimation is conducted by generating 10^4 Monte Carlo samples. The results are found in Table 5. The model misspecification problem particularly affects the price elasticity, which is estimated to be -0.90 with respect to the posterior mean, in contrast to its D/C choice model estimate of 0.020.

Finally, we point out that the electricity and gas would be considered to be substitutable. These two energy resources are the first and second largest in residential use. When the substitution effect is considered, these results would differ.

5.3 Analysis of the Effect of a Change in the Price System

This subsection analyzes the effect of a change in the price system on the demand for residential gas in Japan using the predictive distribution derived in Section 3.3 and its sampling algorithm, **Algorithm 2**. For a prospective price system, we choose the uniform price system in order to evaluate how the block rate pricing affects the demand. The unit price, which is the only one, is selected as the price of the block where consumption is actually made in 1999. While there is no threshold, we use the same basic connection charge with the one in 1999. The logarithm of residential gas demand in 1999, y_i, is summarized in Table 6.

Along **Algorithm 2**, we draw 10^6 samples of predictive demand for each observation, and reduce them to 10^4 samples by picking up every 100-th value. Figure 7 draws the predictive distribution of average of log demand, $\bar{y}_p = n^{-1}\sum_{i=1}^n \tilde{y}_i$. Summary statistics of this predictive distribution are as follows: mean, 3.84m^3; standard deviation, 0.07m^3; and 95% credible interval, (3.70m^3, 3.98m^3). Because the mode of the predictive distribution is almost around the actual average of log demand in 1999, $\bar{y}_{99} =$

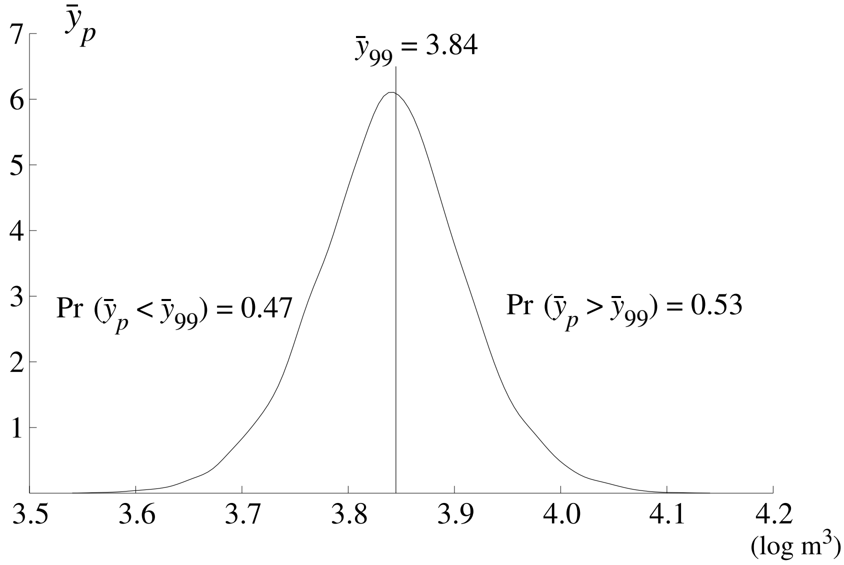

Figure 7: The Predictive Distribution of Average of Log Demand under Uniform Price System.

3.84m^3, and the probability that the average log demand will increase, $\Pr(\bar{y}_{99} < \bar{y}_p)$, is calculated to be 0.53, we can conclude that our prospective uniform price system would have a small effect on the average of log demand.

For a further investigation, Figure 8 draws boxplots of each observation's predictive distribution. Each observation is sorted in ascending order of its actual log demand in 1999, and then, its predictive distribution is summarized as the boxplot. In Figure 8, horizontal axis represents the order of each boxplot ascending from the smallest to the largest actual log demand in 1999 and vertical axis measures the values of log demands. As of the boxplot, each box represents the range between the lower and upper quantiles, and the lower and upper whiskers indicate the 5-th and 95-th percentiles, respectively. In addition to these boxplots, corresponding actual log demands, y_i $(i = 1, \ldots, n)$, are also plotted over these boxplots by the solid line.

This figure shows that most observation suffers small influence on its log demand from our prospective uniform price system because its interquantile range includes the actual log demand. Observations who consume close to the largest or smallest amounts in 1999, however, are affected by this price system change, that is, they are predicted to decrease or increase their consumption, respectively. Therefore, under the D/C choice model, it is concluded that the uniform price system would tend to yield similar residential gas demands with each other compared to the block rate pricing.

6 Concluding Remarks

This paper proposed a duality-based analysis of the demand function under decreasing block rate pricing and conducted an empirical analysis of the Japanese residential gas demand function.

This approach is simple enough to extend not only to multiple-block rate pricing but also to a multivariate extension of the demand function under block rate pricing. One example of the latter model is

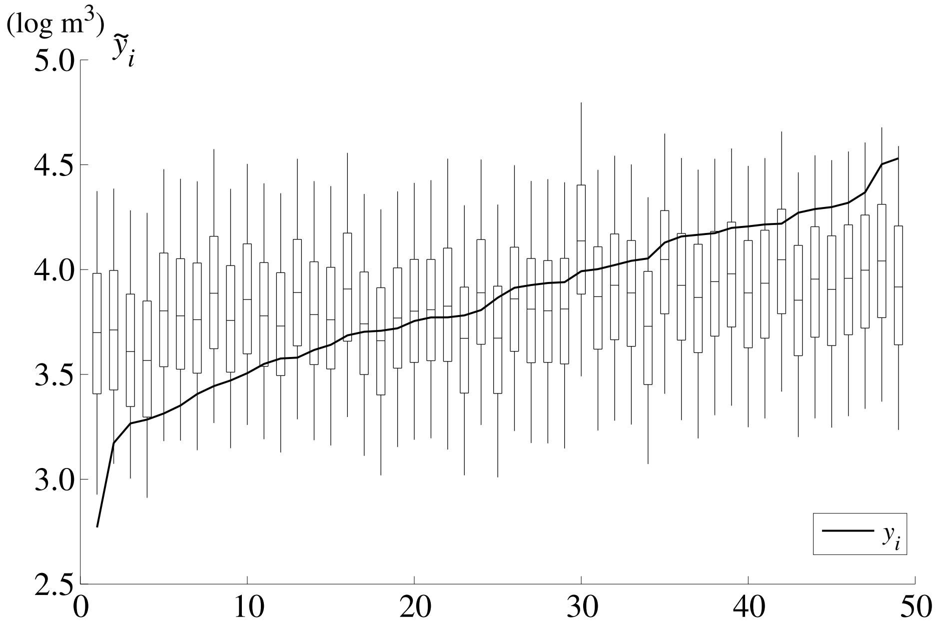

Figure 8: Boxplot of Predictive Distributions.

the energy demand function. Electricity and gas services are supplied under increasing and decreasing block rate pricing in Japan, respectively, and they are both substitutable goods as pointed out at the end of Subsection 5.2. The previous literature, however, does not explicitly take block choice endogeneity into consideration (see, e.g., Dubin and McFadden 1984, Baker, Blundell, and Micklewright 1989 and Lee and Singh 1994). Thus, it is natural to extend the discrete/continuous choice approach to a multivariate model, so that we can analyze this substitution effect by simultaneously dealing with block choice.

Finally, we point out an alternative estimation method proposed by Blomquist and Newey (2002). Their approach is a nonparametric one and is therefore free of distributional misspecification. Furthermore, they does not include any approximation in consumer preferences. Thus, it is possible that their method would yield more accurate estimates of the price and income elasticities than ours. A comparison between parametric and nonparametric estimation is a topic for future analysis.

Acknowledgement

This work is supported by the Grants-in-Aid for Scientific Research 18330039 from the Japanese Ministry of Education, Science, Sports, Culture and Technology and by the Japan Society for the Promotion of Science Fellowship. The authors thank Professor H. K. van Dijk (Erasmus Universiteit Rotterdam), Professor A. C. Harvey (University of Cambridge), Professor M. Fukushige (Osaka University), K. Kakamu (Chiba University), and K. Fukumoto (Gakushuin University) for their helpful comments. The computational results were calculated by using Ox version 4.1a for Linux (see Doornik 2002).

References

Baker, P., R. Blundell, and J. Micklewright (1989). Modelling household energy expenditures using micro data. *The Economic Journal 99*, 720–738.

Baumol, W. J. and W. E. Oates (1988). *The Theory of Environmental Policy* (2nd ed.). Cambridge: Cambridge University Press.

Blomquist, S. and W. Newey (2002). Nonparametric estimation with nonlinear budget sets. *Econometrica 70*(6), 2455–2480.

Burtless, G. and R. A. Moffitt (1985). The joint choice of retirement age and postretirement hours of work. *Journal of Labor Economics 3*(2), 207–236.

Chib, S. (2001). Markov chain Monte Carlo methods: Computation and inference. In J. J. Heckman and E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5, Chapter 57, pp. 3569–3649. Amsterdam: North-Holland.

de Jong, G. C. (1990). An indirect utility model of car ownership and private car use. *European Economic Review* (34), 971–985.

Deaton, A. and J. Muellbauer (1980). *Economics and Consumer Behavior*. Cambridge: Cambridge University Press.

Doornik, J. A. (2002). *Object-Oriented Matrix Programming Using Ox* (3rd ed.). London: Timberlake Consultants Press and Oxford.

Dubin, J. A. and D. L. McFadden (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica 52*(2), 345–362.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith (Eds.), *Bayesian Statistics 4*, pp. 169–193. Oxford: Oxford University Press.

Hausman, J. A. (1980). The effect of wages, taxes, and fixed costs on women's labor force participation. *Journal of Public Economics 14*, 161–194.

Hausman, J. A. (1981). Exact consumer's surplus and deadweight loss. *American Economic Review 71*(4), 662–676.

Hausman, J. A. (1985). The econometrics of nonlinear budget sets. *Econometrica 53*(6), 1255–1282.

Hewitt, J. A. and W. M. Hanemann (1995). A discrete/continuous choice approach to residential water demand under block rate pricing. *Land Economics 71*, 173–192.

Irvine, I. J. and W. A. Sims (1998). Measuring consumer surplus with unknown Hicksian demands. *American Economic Review 88*(1), 314–322.

Lee, R.-S. and N. Singh (1994). Patterns in residential gas and electricity consumption: An econometric analysis. *Journal of Business and Economic Statistics 12*(2), 233–241.

Maddala, G. S. (1983). *Limited-dependent and Qualitative Variables in Econometrics*. Number 3 in Econometric Society Monographs. Cambridge, New York: Cambridge University Press.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microecnomic Theory*. New York: Oxford University Press.

Miyawaki, K., Y. Omori, and A. Hibiki (2006). Bayesian estimation of demand functions under block rate pricing. University of Tokyo CIRJE Discussion Paper Series CIRJE-F-424.

Moffitt, R. (1986). The econometrics of piecewise-linear budget constraint. *Journal of Business and Economic Statistics 4*(3), 317–328.

Olmstead, S. M., W. M. Hanemann, and R. N. Stavins (2007). Water demand under alternative price structures. *Journal of Environmental Economics and Management*. forthcoming.

Reiss, P. C. and M. W. White (2005). Household electricity demand, revisited. *Review of Economic Studies 72*, 853–883.

Train, K. E. (1991). *Optimal Regulation: the Economic Theory of Natural Monopoly*. Cambridge, Mass.: MIT Press.

Xepapadeas, A. (1997). *Advanced Principles in Environmental Policy*. New Horizons in Environmental Economics. Cheltenham, UK: Edward Elgar.