

**Research Unit for Statistical
and Empirical Analysis in Social Sciences (Hi-Stat)**

**Panel Data Analysis of
Japanese Residential Water Demand Using
a Discrete/Continuous Choice Approach**

Koji Miyawaki
Yasuhiro Omori
Akira Hibiki

March 2010

Panel Data Analysis of Japanese Residential Water Demand Using a Discrete/Continuous Choice Approach

Koji Miyawaki^a Yasuhiro Omori^b * Akira Hibiki^c

February 17, 2010

^a National Institute for Environmental Studies, Ibaraki 305-8506, Japan (miyawaki.koji@nies.go.jp).

^b Faculty of Economics, University of Tokyo, Tokyo 113-0033, Japan (omori@e.u-tokyo.ac.jp).

Phone:+81-3-5841-5516. Fax:+81-3-5841-5521.

^c National Institute for Environmental Studies, Ibaraki 305-8506, Japan (hibiki@nies.go.jp).

Abstract

Block rate pricing is often applied to income taxation, telecommunication services, and brand marketing in addition to its best-known application in public utility services. Under block rate pricing, consumers face piecewise-linear budget constraints. A discrete/continuous choice approach is usually used to account for piecewise-linear budget constraints for demand and price endogeneity. A recent study proposed a methodology to incorporate a separability condition that previous studies ignore, by implementing a Markov chain Monte Carlo simulation based on a hierarchical Bayesian approach. To extend this approach to panel data, our study proposes a Bayesian hierarchical model incorporating the individual effect. The random coefficients model result shows that the price and income elasticities are estimated to be negative and positive, respectively, and

*Corresponding author: Tel:+81-3-5841-5516, E-mail:omori@e.u-tokyo.ac.jp

the coefficients of the number of members and the number of rooms per household are estimated to be positive. Furthermore, the AR(1) error component model suggests that the Japanese residential water demand does not have serial correlation.

Key words: Block rate pricing, Bayesian analysis, Panel data, Residential water demand.

JEL classification: C11, C23, C24, Q25.

1 Introduction

Block rate pricing usually has been applied to services in public utility sectors such as water, gas, and electricity.¹ However, block rate pricing is becoming common in areas such as local and wireless telephone services and brand marketing. Under block rate pricing, unit price changes with quantity consumed. When unit price increases with quantity consumed, shown in Figure 1, such a price schedule is called the increasing block rate pricing. When unit price decreases with quantity consumed, it is called the decreasing block rate pricing. Then, under block pricing consumers maximize utility by selecting the unit price and the consumption amount. This circumstance leads to a utility-maximization problem under a piecewise-linear budget constraint.

As surveyed by Olmstead (2009), there are two types of estimation approaches that deal with this problem: reduced-form approaches, such as instrumental variables, and structural approaches. The structural approach solves a consumer's utility maximization problem in two steps. A consumer first decides appropriate consumption given each block's price, and then selects the block that maximizes consumer utility. This is also called a discrete/continuous choice approach because the block selection is discrete while the amount consumed is continuous. Its important feature is that the derived model explicitly addresses the relationship between the block choice and the amount consumed under block rate pricing.

As discussed in Olmstead (2009), the reduced-form approaches can incorporate only

¹The other example where the same rate structure is applied is a progressive tax rate in income tax systems.

limited aspects of the piecewise-linear budget constraint, while the structural approaches account just for the particular implications of piecewise-linear budget constraints for demand as well as price endogeneity. Thus the latter approaches have two main advantages over the former ones: (1) the structural approaches can produce unbiased and consistent estimates of parameters of the price and the income² and (2) they are consistent with utility theory. Despite these advantages, most previous studies employ reduced-form approaches. Structural approaches are rare in demand analysis.³ This is because the discrete/continuous choice approach had been applied only to the simplified block rate price structure—for example, the number of blocks is fixed at two.

Pint (1999); Rietveld et al. (2000); Olmstead, Hanemann, and Stavins (2007); Olmstead (2009) considered multiple-block pricing—for example, the number of blocks varies from two to four. Miyawaki, Omori, and Hibiki (2010) proposed a methodology to solve the problem of multi-tier block rate pricing by implementing a Markov chain Monte Carlo (MCMC) simulation based on a hierarchical Bayesian approach. Their method also showed that model parameters are subject to the separability condition, or set of linear inequality constraints. Despite its importance, previous literature generally ignores the condition because the parameter region becomes tightly restrained, making numerical maximization of the likelihood function difficult (Pint, 1999; Rietveld et al., 2000; Olmstead et al., 2007; Olmstead, 2009).

To extend Miyawaki et al. (2010) for the panel data analysis, this study proposes a Bayesian hierarchical model. It incorporates the individual effect to estimate the residential water demand function under the separability condition using panel data of Japanese households. This is the first study that incorporates the individual effect in the discrete/continuous choice approach.

²Previous studies suggested that water demand is price inelastic. However, as is suggested in the meta-analysis (Dalhuisen, Florax, de Groot, and Nijkamp, 2003), the choice of the approach may affect the estimates, since the water demand is price inelastic in previous studies employing the reduced form approach, however, are price elastic in the discrete/continuous choice approach.

³Olmstead (2009) reported that, between 1963 and 2004, there were only three studies on water demand (Hewitt and Hanemann, 1995; Pint, 1999; Rietveld, Rouwendal, and Zwart, 2000) that adopted the discrete/continuous choice approach.

We organize this article as follows. Section 2 describes block rate pricing and reviews previous studies. Section 3 explains the derivation of the model, incorporating the individual effect to extend Miyawaki et al. (2010) for panel data analysis and proposes its MCMC estimation method. Section 4 shows the empirical analysis of Japanese residential water demand using panel data. Section 5 concludes.

2 Block rate pricing system and literature review

Figure 1 shows the example of a three-tier increasing block rate pricing where Y is the consumption of the good or service, P_k is the unit price of Y in block k ($k = 1, 2, 3$) and \bar{Y}_k is the boundary quantity between block k and $k + 1$, i.e., the upper limit of block k . Under this system, when consumption of Y exceeds \bar{Y}_k the unit price jumps from P_k to P_{k+1} .

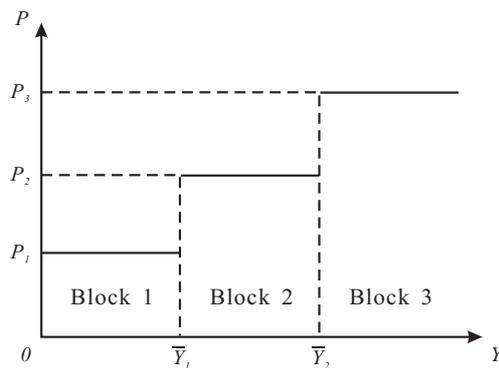


Figure 1: Three-Tier increasing block price structure.

A block rate pricing system generally has the following characteristics:

1. The total payment is comprised of the fixed charge, FC , and the variable charge, VC , depending on the volume consumed. A practical example of the fixed charge is a minimum access charge for water and electricity services.
2. In a block rate pricing system with K blocks, potential consumption is divided into K consumption regions—i.e., K blocks.

3. K unit prices correspond to K blocks. The k -th ($1 \leq k \leq K$) unit price in the k -th block is applied to consumption and is constant within the k -th block.

Now we consider two consumption goods. The first is a good, such as water or electricity, to which block rate pricing is applied. The second is a composite good, to which a single constant price is applied. We denote Y^i and Y_a^i as consumption of the first and the second good by the consumer i . We denote \bar{Y}_k^i as an upper limit (threshold) of the k -th block and P_k^i as the unit price of the k -th block. The superscript i is attached because we assume that different consumers face different block rate pricings. Then, the budget constraint of consumer i , who chooses the consumption, Y_k^i , within the k -th block, $\bar{Y}_{k-1}^i \leq Y_k^i < \bar{Y}_k^i$, is expressed as follows.

$$P_k^i Y_k^i + Y_a^i \leq Q_k^i \equiv I^i - FC^i - \sum_{j=1}^{k-1} (P_j^i - P_{j+1}^i) \bar{Y}_j^i, \quad (1)$$

where we denote I^i as the income of consumer i . The Q_k^i is called as the virtual income of consumer i , who chooses to consume at the k -th block. We also set $\bar{Y}_0^i = 0$ and $\bar{Y}_{K^i}^i = \infty$ without loss of generality. As shown in the above equation, the budget constraint becomes piecewise-linear (see also Figure 2).

The discrete/continuous choice approach is a common structural approach that solves the utility maximization problem under a piecewise-linear budget constraint first proposed by Burtless and Hausman (1978). Other studies include expenditures with food stamps (Moffitt, 1989), car ownership, and use (de Jong, 1990), electricity demand (Herriges and King, 1994; Reiss and White, 2005), water demand (Hewitt and Hanemann, 1995; Olmstead et al., 2007; Miyawaki et al., 2010), choice of wireless service calling plans (Iyengar, 2004), and consumer preference regarding multiple product categories and brands (Song and Chintagunta, 2007).

Although the discrete/continuous choice approach is based on economic theory and can be extended to allow various dependencies, it is difficult to estimate the model's parameters.

For example, Moffitt (1986) pointed out a problem with non-differentiability in the likelihood function and also a computational burden. This is why most previous studies estimate the demand function in a simplified manner in which, for example, all consumers face identical two-tier block rate pricing.

Several recent studies (Pint, 1999; Rietveld et al., 2000; Olmstead et al., 2007; Olmstead, 2009) consider multiple-block pricing (the number of blocks varies from two to four). However, they ignored the separability condition. Miyawaki et al. (2010) addressed this problem by implementing an MCMC simulation based on a hierarchical Bayesian approach, but they did not consider the application to the panel data.

The main contributions of our article are to extend Miyawaki et al. (2010) to incorporate the individual effect and to take a hierarchical Bayesian approach in estimating the residential water demand function under the separability condition using Japanese household panel data.

3 Derivation of the demand function under block rate pricing and application of MCMC

3.1 Derivation of the demand function under block rate pricing

In this subsection we explain the derivation of the demand function under a block rate pricing system within the setting of Section 2.

Suppose consumer i determines the consumption of a good that is subject to K^i -block rate pricing, Y^i , to maximize utility, $U(Y^i, Y_a^i)$, under the piecewise-linear budget constraint as is shown in (1). Since the increasing block rate pricing is used for the residential water supply in Japan, we suppose $P_k^i < P_{k+1}^i$ ($k = 1, \dots, K^i - 1$). Figure 2 illustrates an example of a utility maximization problem under a piecewise-linear budget constraint due to a block rate pricing system (the case of three-tier increasing block pricing), where we denote V^i as the level of an indifference curve and where the second block is optimal with its optimal demand

Y_{opt}^i .

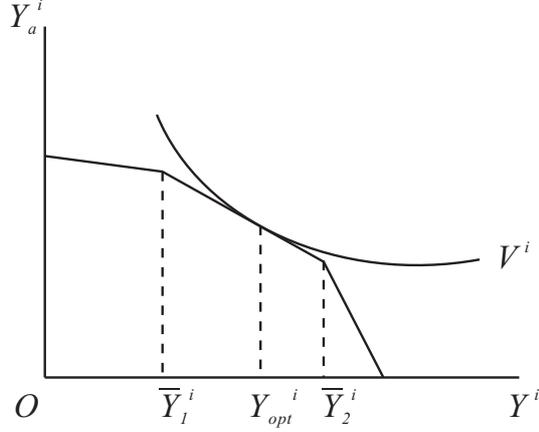


Figure 2: Utility maximization problem: three-block case.

Before derivating the demand function under block rate pricing, we consider K^i conditional utility maximization problems. For $k = 1, \dots, K^i$, the k -th conditional problem is given as follows:

$$\max_{Y_k^i, Y_a^i} U(Y_k^i, Y_a^i) \quad \text{s.t.} \quad P_k^i Y_k^i + Y_a^i \leq Q_k^i. \quad (2)$$

The optimal conditional consumption, Y_k^i , is determined as if consumer i faced a single price P_k^i and given income Q_k^i . With these K^i optimal conditional consumptions, the demand function under increasing block rate pricing for consumer i is given by

$$Y^i = \begin{cases} Y_k^i, & \text{if } \bar{Y}_{k-1}^i < Y_k^i < \bar{Y}_k^i \text{ and } k = 1, \dots, K^i, \\ \bar{Y}_k^i, & \text{if } Y_{k+1}^i \leq \bar{Y}_k^i \leq Y_k^i \text{ and } k = 1, \dots, K^i - 1. \end{cases} \quad (3)$$

Since we apply the log-linear conditional demand model used in previous studies, Equation (3) is rewritten as

$$y_i = \begin{cases} y_{ik}, & \text{if } \bar{y}_{i,k-1} < y_{ik} < \bar{y}_{ik} \text{ and } k = 1, \dots, K^i, \\ \bar{y}_{ik}, & \text{if } y_{i,k+1} \leq \bar{y}_{ik} \leq y_{ik} \text{ and } k = 1, \dots, K^i - 1, \end{cases} \quad (4)$$

$$y_{ik} = \beta_1 p_{ik} + \beta_2 q_{ik} \equiv \mathbf{x}'_{ik} \boldsymbol{\beta}, \quad (5)$$

where $(y_i, y_{ik}, \bar{y}_{ik}, p_{ik}, q_{ik}) = (\log Y^i, \log Y_k^i, \log \bar{Y}_k^i, \log P_k^i, \log Q_k^i)$, $\mathbf{x}_{ik} = (p_{ik}, q_{ik})'$, and $\boldsymbol{\beta} = (\beta_1, \beta_2)'$.

As per Miyawaki et al. (2010), we now introduce two unobserved random variables into the demand function of the i -th consumer: the heterogeneity, w_i^* , and the state variable, s_i^* . The heterogeneity is a stochastic term that models consumers' characteristics and is assumed to be an additive to the log conditional demand y_k^i . Thus, w_i^* is assumed to follow the linear model

$$w_i^* = \mathbf{z}_i' \boldsymbol{\delta} + v_i, \quad v_i \sim \text{i.i.d. } N(0, \sigma_v^2), \quad (6)$$

where \mathbf{z}_i and $\boldsymbol{\delta}$ are $d \times 1$ vectors of explanatory variables for the heterogeneity and corresponding parameters, respectively, and v_i is an independently and identically distributed disturbance term with a normal distribution of mean 0 and variance σ_v^2 . The state variable, s_i^* , is a discrete random variable that indicates which block is potentially optimal for consumers.

Then, the basic model for the demand function under increasing the block rate pricing is given by the following equations:

$$y_{ik} = \mathbf{x}'_{ik} \boldsymbol{\beta}, \quad \mathbf{x}_{ik} = (p_{ik}, q_{ik})', \quad k = 1, \dots, K^i, \quad (7)$$

$$w_i^* = \mathbf{z}_i' \boldsymbol{\delta} + v_i, \quad v_i \sim \text{i.i.d. } N(0, \sigma_v^2), \quad (8)$$

$$s_i^* = \begin{cases} 2k-1, & \text{if } w_i^* \in R_{i,2k-1} = (\bar{y}_{i,k-1} - y_{ik}, \bar{y}_{ik} - y_{ik}) \text{ and } k = 1, \dots, K^i, \\ 2k, & \text{if } w_i^* \in R_{i,2k} = (\bar{y}_{i,k-1} - y_{ik}, \bar{y}_{ik} - y_{ik}) \text{ and } k = 1, \dots, K^i - 1, \end{cases} \quad (9)$$

$$y_i^* = \begin{cases} y_{ik} + w_i^*, & \text{if } s_i^* = 2k-1 \text{ and } k = 1, \dots, K^i, \\ \bar{y}_{ik}, & \text{if } s_i^* = 2k \text{ and } k = 1, \dots, K^i - 1, \end{cases} \quad (10)$$

$$y_i = y_i^* + u_i, \quad u_i \sim \text{i.i.d. } N(0, \sigma_u^2). \quad (11)$$

The extension of the above model to incorporate the individual effect for parameter $\boldsymbol{\delta}$ in the panel data with n observations and T time periods, which we call a random coefficients

model (RC), is rewritten as follows:

$$y_{it,k} = \mathbf{x}'_{it,k} \boldsymbol{\beta}, \quad \mathbf{x}_{it,k} = (p_{it,k}, q_{it,k})', \quad k = 1, \dots, K^t, \quad (12)$$

$$\mathbf{w}_i^* = \mathbf{Z}_i \boldsymbol{\delta}_i + \mathbf{v}_i, \quad \mathbf{v}_i \sim \text{i.i.d. } N_T(\mathbf{0}, \sigma_v^2 \mathbf{I}), \quad (13)$$

$$s_{it}^* = \begin{cases} 2k-1, & \text{if } w_{it}^* \in R_{it,2k-1} \text{ and } k = 1, \dots, K^{it}, \\ 2k, & \text{if } w_{it}^* \in R_{it,2k} \text{ and } k = 1, \dots, K^{it} - 1, \end{cases} \quad (14)$$

$$y_{it}^* = \begin{cases} y_{it,k} + w_i^* = \mathbf{x}'_{it,k} \boldsymbol{\beta} + w_i^*, & \text{if } s_{it}^* = 2k-1 \text{ and } k = 1, \dots, K^{it}, \\ \bar{y}_{it,k}, & \text{if } s_{it}^* = 2k \text{ and } k = 1, \dots, K^{it} - 1, \end{cases} \quad (15)$$

$$y_{it} = y_{it}^* + u_{it}, \quad u_{it} \sim \text{i.i.d. } N(0, \sigma_u^2), \quad (16)$$

where subscripts i and t denote the observation i and time t , respectively, $N_T(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ denotes a T -variate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$, and \mathbf{I} is an identity matrix.

This RC model extends the basic model in three ways. First, it introduces a linear structure and a normal error \mathbf{v}_i to consumer heterogeneity, $\mathbf{w}_i^* = (w_{i1}^*, \dots, w_{iT}^*)'$. The $\mathbf{Z}_i = (\mathbf{z}_{i1}, \dots, \mathbf{z}_{iT})'$ and $\boldsymbol{\delta}_i$ are a $T \times d$ vector of explanatory variables and a $d \times 1$ vector of their coefficients, respectively. The heterogeneity intervals are given by

$$R_{it,2k-1} = (\bar{y}_{it,k-1} - \mathbf{x}'_{it,k-1} \boldsymbol{\beta}, \bar{y}_{it,k} - \mathbf{x}'_{it,k} \boldsymbol{\beta}), \quad R_{it,2k} = (\bar{y}_{it,k} - \mathbf{x}'_{it,k} \boldsymbol{\beta}, \bar{y}_{it,k} - \mathbf{x}'_{it,k+1} \boldsymbol{\beta}). \quad (17)$$

To capture the individual effect for the coefficient of \mathbf{Z}_i , the $\boldsymbol{\delta}_i$'s are assumed to be independently and identically distributed random samples from a normal distribution $N_d(\boldsymbol{\mu}_\delta, \sigma_v^2 \boldsymbol{\Sigma}_\delta)$ as in (18). Second, a discrete latent variable, s_{it}^* , is used to indicate potentially optimal demand chosen by the i -th consumer at time t . When s_{it}^* is odd ($s_{it}^* = 2k-1$ for $k = 1, \dots, K^{it}$), the i -th consumer would select the optimal conditional demand. When, on the other hand, s_{it}^* is even ($s_{it}^* = 2k$ for $k = 1, \dots, K^{it} - 1$), one of threshold values would be optimal for the

consumer. We augment the model parameter space by introduction of the s_{it}^* and exploit the data augmentation method to estimate parameters (see, e.g., Tanner and Wong (1987) for the description of the data augmentation). Third, another normal disturbance u_{it} is considered for the potential demand y_{it}^* . It represents the measurement error as well as the optimization error and the model misspecification error (see Hausman, 1985).

The RC model above includes two popular models in panel data analysis. When \mathbf{z}_{it} includes $y_{i,t-1}$ as an explanatory variable, the model becomes the dynamic panel data model. On the other hand, when the heterogeneity w_{it}^* has an AR(1) serial correlation, this model is interpreted as an AR(1) error component model (see Appendix A.2). The AR(1) process specification can be further extended with a heteroskedastic variance structure, $\mathbf{w}_i^* \sim N(\mathbf{Z}_i \boldsymbol{\delta}_i, \boldsymbol{\Sigma})$.

The next subsection describes a Bayesian estimation method for this RC model.

3.2 Bayesian analysis and MCMC implementation

To conduct a Bayesian analysis, we assume the prior distributions of $(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2)$, which are given by

$$\begin{aligned} \boldsymbol{\beta} | \sigma_u^2 &\sim N_2(\boldsymbol{\mu}_{\boldsymbol{\beta},0}, \sigma_u^2 \boldsymbol{\Sigma}_{\boldsymbol{\beta},0}), & \boldsymbol{\delta}_i | \sigma_v^2, \boldsymbol{\mu}_{\boldsymbol{\delta}}, \boldsymbol{\Sigma}_{\boldsymbol{\delta}} &\sim \text{i.i.d. } N_d(\boldsymbol{\mu}_{\boldsymbol{\delta}}, \sigma_v^2 \boldsymbol{\Sigma}_{\boldsymbol{\delta}}), \\ \sigma_u^2 &\sim IG\left(\frac{n_{u,0}}{2}, \frac{S_{u,0}}{2}\right), & \sigma_v^2 &\sim IG\left(\frac{n_{v,0}}{2}, \frac{S_{v,0}}{2}\right),^4 \end{aligned} \quad (18)$$

where $\boldsymbol{\mu}_{\boldsymbol{\beta},0} = (\mu_{\beta_1,0}, \mu_{\beta_2,0})'$ is a 2×1 known vector, $\boldsymbol{\Sigma}_{\boldsymbol{\beta},0} = \text{diag}(\sigma_{\beta_1,0}^2, \sigma_{\beta_2,0}^2)$ is a 2×2 known diagonal matrix with positive diagonal elements $(\sigma_{\beta_1,0}^2, \sigma_{\beta_2,0}^2)$, and $n_{u,0}, S_{u,0}, n_{v,0}, S_{v,0}$ are known positive constants. Let $\pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2 | \boldsymbol{\mu}_{\boldsymbol{\delta}}, \boldsymbol{\Sigma}_{\boldsymbol{\delta}})$ denote the prior probability density function of $(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2)$ conditional on $(\boldsymbol{\mu}_{\boldsymbol{\delta}}, \boldsymbol{\Sigma}_{\boldsymbol{\delta}})$. We further assume the proper

⁴The distribution $IG(a, b)$ denotes an inverse gamma distribution with a probability density function

$$\pi(x) \propto x^{-(a+1)} \exp(-b/x), \quad x > 0,$$

where a and b are positive constants.

hierarchical priors for $(\boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta)$ such that

$$\boldsymbol{\mu}_\delta \sim N_d(\boldsymbol{\mu}_{\bar{\delta},0}, \boldsymbol{\Sigma}_{\bar{\delta},0}), \quad \boldsymbol{\Sigma}_\delta \sim IW_d(n_{\bar{\delta},0}, \mathbf{S}_{\bar{\delta},0}),^5 \quad (19)$$

where $\boldsymbol{\mu}_{\bar{\delta},0}$ is a $d \times 1$ known vector, $\boldsymbol{\Sigma}_{\bar{\delta},0}$ and $\mathbf{S}_{\bar{\delta},0}$ are known $d \times d$ positive definite matrices, and $n_{\bar{\delta},0} > d - 1$ is known constant. Denoting the prior probability density function of $(\boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta)$ by $\pi(\boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta)$, the joint prior probability density function of model parameters is

$$\pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta) = \pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2 \mid \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta) \pi(\boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta). \quad (20)$$

Then the joint posterior probability density function of model parameters is given by

$$\begin{aligned} & \pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i, \mathbf{s}_i^*, \mathbf{w}_i^*\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta \mid \{\mathbf{y}_i\}_{i=1}^n) \propto \pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta) \\ & \times \sigma_u^{-nT} \sigma_v^{-nT} \exp \left[-\frac{1}{2} \sum_{i=1}^n \left\{ \sigma_u^{-2} (\mathbf{y}_i - \mathbf{y}_i^*)' (\mathbf{y}_i - \mathbf{y}_i^*) + \sigma_v^{-2} (\mathbf{w}_i^* - \mathbf{Z}_i \boldsymbol{\delta}_i)' (\mathbf{w}_i^* - \mathbf{Z}_i \boldsymbol{\delta}_i) \right\} \right] \\ & \times \prod_{i=1}^n \prod_{t=1}^T \left\{ I(w_{it}^* \in R_{it, s_{it}^*}) \prod_{k=1}^{K_{it}-1} I(\mathbf{x}'_{it, k+1} \boldsymbol{\beta} \leq \mathbf{x}'_{it, k} \boldsymbol{\beta}) \right\}, \quad (21) \end{aligned}$$

where $I(A)$ is the indicator function; $I(A) = 1$ if A is true and $I(A) = 0$ otherwise, $\mathbf{y}_i = (y_{i1}, y_{i2}, \dots, y_{iT})'$, $\mathbf{y}_i^* = (y_{i1, s_{i1}^*}^*, y_{i2, s_{i2}^*}^*, \dots, y_{iT, s_{iT}^*}^*)'$, and $\mathbf{s}_i^* = (s_{i1}^*, s_{i2}^*, \dots, s_{iT}^*)'$. The last term

⁵The $IW_d(n_{\bar{\delta},0}, \mathbf{S}_{\bar{\delta},0})$ denotes an inverse Wishart distribution with a probability density function given by

$$\pi(\boldsymbol{\Sigma}_\delta) \propto |\boldsymbol{\Sigma}_\delta|^{-\frac{n_{\bar{\delta},0}+d+1}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{S}_{\bar{\delta},0}^{-1} \boldsymbol{\Sigma}_\delta^{-1}) \right\}.$$

For $n_{\bar{\delta},0} > d + 3$, its mean and variance exist and are given by

$$\begin{aligned} E(\boldsymbol{\Sigma}_\delta) &= \frac{1}{n_{\bar{\delta},0} - d - 1} \mathbf{S}_{\bar{\delta},0}^{-1}, \\ \text{Var}(\sigma_{ii}) &= \frac{2(s^{ii})^2}{(n_{\bar{\delta},0} - d - 1)^2 (n_{\bar{\delta},0} - d - 3)}, \quad \text{Var}(\sigma_{ij}) = \frac{s^{ii} s^{jj} + \frac{n_{\bar{\delta},0} - d + 1}{n_{\bar{\delta},0} - d - 1} (s^{ij})^2}{(n_{\bar{\delta},0} - d)(n_{\bar{\delta},0} - d - 1)(n_{\bar{\delta},0} - d - 3)}, \end{aligned}$$

where σ_{ij} and s^{ij} are the ij -th element of $\boldsymbol{\Sigma}_\delta$ and $\mathbf{S}_{\bar{\delta},0}^{-1}$, respectively. When $(d, n_{\bar{\delta},0}, \mathbf{S}_{\bar{\delta},0}) = (4, 10, 10^{-1} \mathbf{I}_4)$ (which we use for our empirical analysis in Section 4), the mean and variance of $\boldsymbol{\Sigma}_\delta$ are $(2, 8/3)$ for σ_{ii} and $(0, 10/9)$ for σ_{ij} (see Chapter 3 of Gupta and Nagar (2000) for further characteristics of the inverse Wishart distribution).

$I(\mathbf{x}'_{it,k+1}\boldsymbol{\beta} \leq \mathbf{x}'_{it,k}\boldsymbol{\beta})$ is the separability condition that guarantees disjoint heterogeneity intervals (see (17)). Because $\boldsymbol{\beta}$ is a two-dimensional vector in our statistical modeling, this condition reduces to two inequality constraints:

$$\beta_2 \leq \bar{r}\beta_1 \text{ and } \beta_2 \leq \underline{r}\beta_1, \quad (22)$$

where $\bar{r} = \max_{i,t,k} -(p_{it,k+1} - p_{it,k})/(q_{it,k+1} - q_{it,k})$ and $\underline{r} = \min_{i,t,k} -(p_{it,k+1} - p_{it,k})/(q_{it,k+1} - q_{it,k})$. Further discussion is found in Miyawaki et al. (2010).

As all full conditional distributions are well-known (see Appendix A.1), we use a Gibbs sampler to draw samples from the posterior distribution, which is implemented in nine steps:

MCMC algorithm for the RC model

Step 1. Initialize $\boldsymbol{\beta}, \{\boldsymbol{\delta}_i, \mathbf{s}_i^*, \mathbf{w}_i^*\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta$, and $\boldsymbol{\Sigma}_\delta$.

Step 2. Generate β_1 given $\beta_2, \{\mathbf{s}_i^*, \mathbf{w}_i^*\}_{i=1}^n, \sigma_u^2$.

Step 3. Generate β_2 given $\beta_1, \{\mathbf{s}_i^*, \mathbf{w}_i^*\}_{i=1}^n, \sigma_u^2$.

Step 4. Generate $(\sigma_v^2, \{\boldsymbol{\delta}_i\}_{i=1}^n)$ given $\{\mathbf{w}_i^*\}_{i=1}^n, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta$.

(a) Generate σ_v^2 given $\{\mathbf{w}_i^*\}_{i=1}^n, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta$.

(b) Generate $\boldsymbol{\delta}_i$ given $\{\mathbf{w}_i^*\}_{i=1}^n, \sigma_v^2, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta$ for $i = 1, \dots, n$.

Step 5. Generate $\boldsymbol{\mu}_\delta$ given $\{\boldsymbol{\delta}_i\}_{i=1}^n, \boldsymbol{\Sigma}_\delta, \sigma_v^2$.

Step 6. Generate $\boldsymbol{\Sigma}_\delta$ given $\{\boldsymbol{\delta}_i\}_{i=1}^n, \boldsymbol{\mu}_\delta, \sigma_v^2$.

Step 7. Generate (s_{it}^*, w_{it}^*) given $\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2$ for $i = 1, \dots, n$ and $t = 1, \dots, T$.

(a) Generate s_{it}^* given $\boldsymbol{\beta}, \boldsymbol{\delta}_i, \sigma_u^2, \sigma_v^2$.

(b) Generate w_{it}^* given $\boldsymbol{\beta}, \boldsymbol{\delta}_i, s_{it}^*, \sigma_u^2, \sigma_v^2$.

Step 8. Generate σ_u^2 given $\boldsymbol{\beta}, \{\mathbf{s}_i^*, \mathbf{w}_i^*\}_{i=1}^n$.

Step 9. Go to Step 2.

4 Empirical analysis

4.1 Data

We use the household-level dataset collected by internet surveys concerning household water and energy consumption and garbage emissions, which we conducted twice (in June 2006 and June 2007) for individuals in the Tokyo and Chiba prefectures in collaboration with INTAGE, Inc., a marketing research company (www.intage.co.jp/english), which has more than 1.3 million monitors all over Japan. As respondents, 1,687 monitors were randomly selected from all INTAGE monitors, 47,239, in this area who are between age 20 and 79. The numbers of respondents in June 2006 and June 2007 were 1,276 and 760, respectively. The number of respondents in both June 2006 and 2007 was 515. The individuals' answers concerned attributes of the household to which they belong, including the number of household members, household annual income, number of rooms and floor space of their house or apartment, and the household's monthly water and sewerage bills. Because water and sewerage are billed every second month in Japan, reported usage is considered to be a two-month usage. In the survey, these attributes are collected only once yearly, and we used respondents collected in June 2006 and April 2007. Since sewerage and water bills are also calculated based on water consumption, the amount of water consumption was calculated from the water and sewerage bills using the corresponding information on water charge schedules and sewerage charge schedules in each city. Every household faces increasing block rate pricing; the number of blocks varies from two to eleven, depending on cities where respondents live.

The number of observations used for the empirical analysis in the next subsection was reduced to 135 because of respondents' missing or inappropriate answers or for technical reasons as follows:

1. Consumption within the zero unit price block is observed.
2. Living in cities that have discontinuous parts in their price system.
3. Living in cities that changed rate tables in June 2006.⁶

4. Using a well for water use because of its special charge system.

The histograms of the amount of water consumption, the dependent variable for the empirical analysis in the next subsection, are shown in Figure 3. Other variables used as explanatory

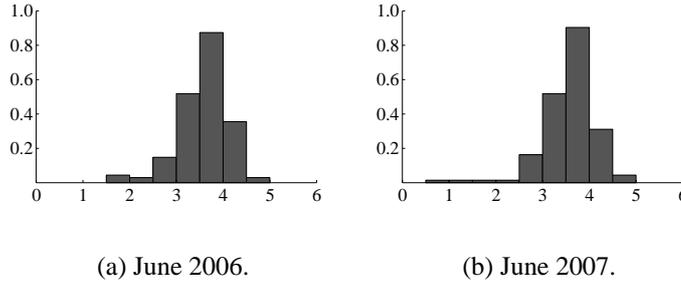


Figure 3: Histograms of the amount of water consumption ($\log \text{m}^3$).

variables for the empirical analysis are listed in Table 1. In Figure 4, we summarize the block

Table 1: Explanatory variables used in the water demand function

<i>Variable</i>	<i>Coefficient</i>	<i>Description</i>
price	β_1	water+sewer ($\log \text{¥}10^3/\text{m}^3$)
virtual income	β_2	income augmented by price ($\log \text{¥}10^3$)
variables for w_i^*	δ_0	the constant
	δ_1	the number of members in a household (person)
	δ_2	the number of rooms in a house/apartment (room)
	δ_3	the total floor space of a house/apartment (50m^2)

rate price structure. Each column of Figure 4 shows the histograms of the number of blocks, the unit price where the consumption is actually made, and the minimum access charge for June 2006 and June 2007.

Regarding the income variable, it is a sensitive issue to ask households their exact annual income level. Therefore, in our survey, instead of the actual values, the household is asked to choose one of eight categories for the annual income in million yen; 0-2, 2-4, 4-6, 6-8, 8-10,

⁶In June 2007, no cities changed the rate tables.

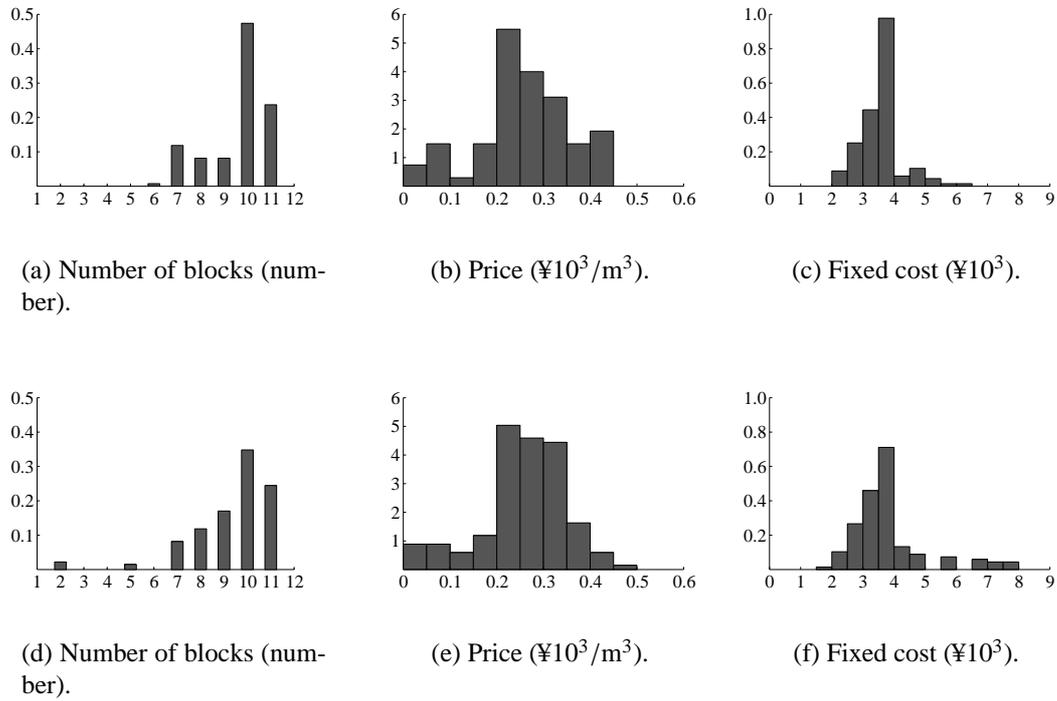


Figure 4: Histograms of the number of blocks, price, and fixed cost. Top row is for June 2006 and bottom row is for June 2007.

10-12, 12-15, over 15 million yen. The histograms for the income categories are shown in Figure 5. For the empirical analysis, we use the median of the interval of each categories

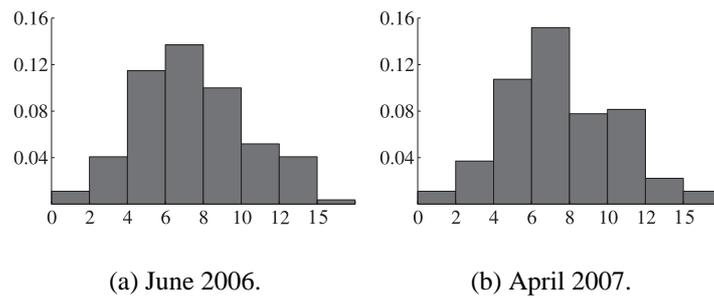


Figure 5: Histograms for the income (¥10⁶).

divided by six to estimate the two-month income for the households except of those who choose “over 15 million yen.” Households whose annual incomes are over 15 million yen are asked to answer the value of their annual income.

Basic statistics for heterogeneity are given in Table 2. We calculate the correlation coef-

Table 2: Basic statistics of explanatory variables for heterogeneity

<i>Variable</i>	<i>Unit</i>	<i>Year</i>	<i>Mean</i>	<i>SD</i>	<i>Min.</i>	<i>Max.</i>
the number of members in a household (δ_1)	person	2006	3.18	1.20	1	7
		2007	3.21	1.23	1	8
the number of rooms in a house/apartment (δ_2)	room	2006	4.41	1.08	2	8
		2007	4.39	1.07	2	8
the total floor space of a house/apartment (δ_3)	50m ²	2006	1.68	0.72	0.24	4.60
		2007	1.68	0.72	0.24	4.60

ficients among explanatory variables for heterogeneity. All correlation coefficients are less than .6, except for the correlation between the number of rooms and total floor space, which is .68 in 2006 and .67 in 2007.

4.2 Estimation results of panel data models

This subsection conducts the empirical analysis of Japanese residential water demand using the random coefficients (RC) model. It should be noted that use of two-period panel data conducted in June 2006 and June 2007 data is useful in removing the seasonality effect. The dependent variable is the amount of water consumption calculated from water and sewerage bills using the corresponding charge schedules. The explanatory variables are listed in Table

1. The separability condition on the parameter space of $\boldsymbol{\beta}$ implies

$$\beta_2 \leq -0.16\beta_1 \text{ and } \beta_2 \leq -3263.83\beta_1. \quad (23)$$

Prior distributions are parameterized by setting $\boldsymbol{\mu}_{\bar{\delta},0} = \mathbf{0}$, $\boldsymbol{\Sigma}_{\bar{\delta},0} = 10\mathbf{I}_4$, $n_{\bar{\delta},0} = 10$, $\mathbf{S}_{\bar{\delta},0} = 10^{-1}\mathbf{I}_4$, $\boldsymbol{\mu}_{\boldsymbol{\beta},0} = \mathbf{0}$, $\boldsymbol{\Sigma}_{\boldsymbol{\beta},0} = 10\mathbf{I}_2$, and $n_{u,0} = S_{u,0} = n_{v,0} = S_{v,0} = 0.1$. We adopt the Gibbs sampler described in Subsection 3.2. For Bayesian inferences, we generate 15 million samples after deleting the initial six million samples. The recorded values are reduced to 10,000 samples

by picking up every 1500-th value. Results are summarized in Figure 6 and Table 3.

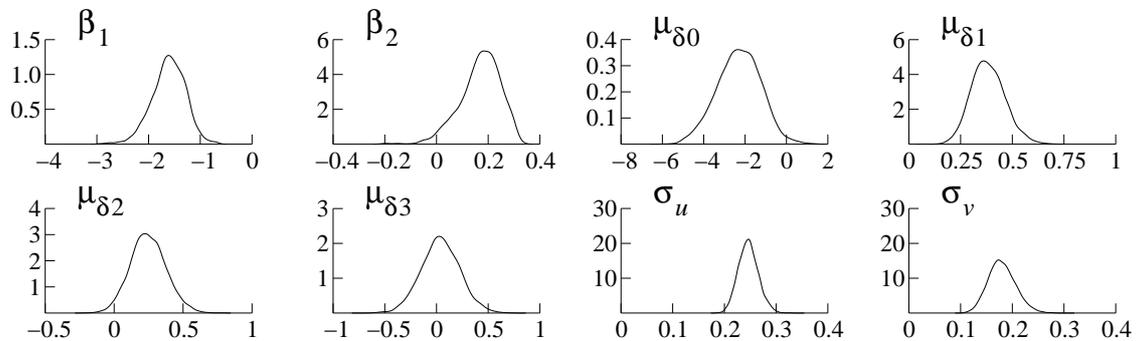


Figure 6: Estimated marginal posterior densities.

Table 3: Water demand function (RC model)

Parameter	Mean	SD	95% interval	INEF*	CD*
β_1 (price)	-1.61	.33	[-2.30 -1.02]	125	.593
β_2 (income)	.17	.079	[-.00 .30]	157	.787
$\mu_{\delta 0}$ (constant)	-2.30	1.06	[-4.42 -.34]	134	.705
$\mu_{\delta 1}$ (num. of members)	.38	.082	[.23 .56]	20	.814
$\mu_{\delta 2}$ (num. of rooms)	.25	.13	[.00 .52]	6	.870
$\mu_{\delta 3}$ (floor space)	.039	.19	[-.33 .42]	2	.387
σ_u (measurement error)	.25	.019	[.21 .29]	2	.636
σ_v (heterogeneity)	.18	.027	[.13 .24]	9	.853

* “INEF” and “CD” denote the inefficiency factor and the p -value of convergence diagnostic statistic, respectively.

Each column of Table 3 represents the parameter symbols (their corresponding variables), posterior means, posterior standard deviations, posterior 95% credible intervals, inefficiency factors, and p -value of convergence diagnostic statistics. The inefficiency factor is an indicator that measures the degree of autocorrelation of the Markov chain and is defined as $1 + 2 \sum_{j=1}^{\infty} \rho(j)$, where $\rho(j)$ is the lag j sample autocorrelation. As pointed out in Chib (2001), this value is interpreted as the ratio of the variance of the sample mean obtained by the Markov chain to that of the sample mean by an uncorrelated draw. When it is close to one, the Markov chain would be as efficient as an uncorrelated Monte Carlo draw. When, on

the other hand, it is much greater than one, we need to take a longer Markov chain. In contrast, the p -value is for the two-sided test of whether the convergence of the Markov chain is reached, proposed by Geweke (1992). The first 10% and last 50% MCMC samples are used to conduct this test as suggested by Geweke (1992).

Obtained MCMC samples for all parameters can be considered to be those from the posterior distribution judging from the p -values of their convergence diagnostics. The inefficiency factors also suggest that we took a sufficiently long Markov chain to conduct inferences.

Table 3 shows several aspects of the Japanese residential water demand function. First, price and income elasticities are highly credible to be negative and positive, respectively, in terms of their 95% credible intervals.⁷ The absolute value of price elasticity is much larger than that of income elasticity. Because the separability condition strongly restricts the parameter space, this result could be a consequence of this condition (see also Miyawaki et al. (2010)). These elasticities have theoretically correct signs. Second, the number of members in a household and the number of rooms in a household/apartment have a positive effect on water demand because the posterior probability $P(\mu_{\delta_j} > 0 | \text{Data}) > .95$ ($j = 1, 2$). In contrast, the total floor space in a household/apartment (δ_3) has no effect on water demand in terms of its 95% credible interval. This result is partly influenced by the correlation between the number of rooms and total floor space, as noted at the end of the preceding subsection.

We compare our results with those obtained in previous studies,⁸ all of which applied the maximum likelihood method to estimate the water demand function based on the discrete/continuous choice approach. Their statistical models to be estimated do not include the individual effect. Furthermore, the separability condition is also ignored in these studies.

Olmstead et al. (2007) used data from households in the United States and Canada. The

⁷Precisely, the 95% credible interval for β_2 includes zero, which means that β_2 does not differ from zero in terms of the credible interval. However, the posterior probability $P(\beta_2 > 0 | \text{Data}) = .97$ implies that we have credible evidence for the positive income elasticity with more than 95% posterior probability.

⁸Pint (1999) estimated the water demand function during the California drought. Because Pint (1999) used the level of unit price as an explanatory variable for the conditional demand, its estimation result cannot be simply compared with ours.

household faces one of three kinds of price schedules: two-block increasing block rate pricing, four-block increasing block rate pricing, and uniform pricing. The estimated price and income elasticities (the coefficients of price and virtual income) are $-.3407$ and $.1306$, respectively, and their standard errors are $.0298$ and $.0118$, respectively. While their income elasticity is similar to ours, their price elasticity is smaller. They used 21 explanatory variables for heterogeneity, including number of residents per household, number of bathrooms, approximate area of the home, approximate area of its lot, and the approximate age of the home as household attributes. Coefficients of these variables are all significant at the 5% level. In particular, the coefficients of the number of residents per household and the approximate area of the home are $.1960$ and $.1257$, respectively.

Hewitt and Hanemann (1995) also estimated the residential water demand function under two-block increasing block rate pricing in Denton, Texas. The price and income elasticities are estimated to be -1.8989 and $.1782$, respectively, and their asymptotic t statistics are -6.421 and 1.864 , respectively. These results are similar to ours. Among variables for heterogeneity, they found that the number of bathrooms has a positive effect on water demand at the 5% significance level. They consider that the number of bathrooms would represent the number of members in a household, which would better explain the variation in residential water use.

Rietveld et al. (2000) analyzed the water demand function under four-block increasing block rate pricing in Indonesia. The price and income elasticities are estimated to be -1.280 and $.501 \times 10^{-6}$, respectively, with standard errors $.235$ and $.348 \times 10^6$, respectively. The tendency for demand to be elastic with regard to price and inelastic with regard to income is coincident with the results of Hewitt and Hanemann (1995) and ours. Furthermore, the log of the number of members in a household has a positive effect on water demand at the 5% significance level.

Finally, we further considered another panel data model, the AR(1) error component model. The MCMC simulation following procedures described in Appendix A.2 is con-

ducted. Because the results are found to be very similar to those obtained for the RC model, their details are omitted. The parameter that represents the serial correlation is not credible to be positive or negative in the sense that its 95% credible interval includes zero. No serial correlation is also observed when we use the four-consecutive-months data—that is, the data from June 2006 to September 2006.

5 Conclusion

This paper conducted a structural analysis of the Japanese residential water demand using panel data. The random coefficients model result shows that the price and income elasticities are estimated to be negative and positive, respectively, and the coefficients of the number of members and the number of rooms are estimated to be positive. Furthermore, the AR(1) error component model suggests that the Japanese residential water demand does not have serial correlation.

We note two applications of our model. First, the proposed model is useful for making policies that continue several periods. For example, the price and income elasticities play an important role when the policy makers make decisions on efficient use and allocation of water. This is especially important in developing countries and transition economies (see, e.g., da Motta, Huber, and Ruitenbeek (1998)). Furthermore, our model is beneficial to formulate the policy on population. The water and sewerage services are one of the factors that determine the population growth (see, e.g., Robinson (1997)).

Second, our model can incorporate a spatial dependency through the consumer heterogeneity. When we analyze the interregional residential water demand, it is important to control such a spatial dependency. The analysis of spatial dependency in the demand for public utilities would be a subject for future research.

Appendices

A.1 Full conditional distributions for RC model

The full conditional distributions for the random coefficients (RC) model is described in detail, following the algorithm in Subsection 3.2. We assume $p_{it,1} > 0$, $q_{it,1} > 0$, and $\bar{y}_{it,1} > 0$ to avoid tedious expressions depending on the sign of these variables without loss of generality. Let $k_{it} = \lceil s_{it}^*/2 \rceil$ and $\mathcal{A} = \{(i, t) \mid s_{it}^* \text{ is odd and equal to } 2k_{it} - 1 \text{ for } t = 1, \dots, T\}$, where $\lceil x \rceil$ is the ceiling function returning the smallest integer that is larger than or equal to x .

Step 1. Initialize $\boldsymbol{\beta}$, $\{\boldsymbol{\delta}_i, \mathbf{s}_i^, \mathbf{w}_i^*\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta$, and $\boldsymbol{\Sigma}_\delta$.*

Step 2. Generate β_1 given $\beta_2, \{\mathbf{s}_i^, \mathbf{w}_i^*\}_{i=1}^n, \sigma_u^2$.* The full conditional distribution for β_1 is the truncated normal distribution with mean μ_1 , variance σ_1^2 , and truncation interval R_1 : $\beta_1 \sim TN_{R_1}(\mu_1, \sigma_1^2)$, where

$$\sigma_1^{-2} = \sigma_{\beta_1,0}^{-2} + \sum_{(i,t) \in \mathcal{A}} (p_{it,k_{it}})^2, \quad (\text{A.24})$$

$$\mu_1 = \sigma_1^2 \left\{ \sigma_{\beta_1,0}^{-2} \mu_{\beta_1,0} + \sum_{(i,t) \in \mathcal{A}} p_{it,k_{it}} (y_{it} - \beta_2 q_{it,k_{it}} - w_{it}^*) \right\}, \quad (\text{A.25})$$

$$R_1 = \left\{ \max_{i,t}(-\infty, BL_{it}^1), \min_{i,t,k} \left(BU_{it}^1, -\beta_2 \frac{q_{it,k+1} - q_{it,k}}{p_{it,k+1} - p_{it,k}} \right) \right\}, \quad (\text{A.26})$$

$$(BL_{it}^1, BU_{it}^1) = \begin{cases} \left(\frac{\bar{y}_{it,k-1} - \beta_2 q_{it,k} - w_{it}^*}{p_{it,k}}, \frac{\bar{y}_{it,k} - \beta_2 q_{it,k} - w_{it}^*}{p_{it,k}} \right), & \text{if } (i, t) \in \mathcal{A}, \\ \left(\frac{\bar{y}_{it,k} - \beta_2 q_{it,k} - w_{it}^*}{p_{it,k}}, \frac{\bar{y}_{it,k} - \beta_2 q_{it,k+1} - w_{it}^*}{p_{it,k+1}} \right), & \text{otherwise.} \end{cases} \quad (\text{A.27})$$

These (BL_{it}^1, BU_{it}^1) are constructed from the intervals R_{it, s_{it}^*} defined by (17) of Subsection 3.2.

Step 3. Generate β_2 given $\beta_1, \{\mathbf{s}_i^, \mathbf{w}_i^*\}_{i=1}^n, \sigma_u^2$.* The full conditional distribution for β_2 is the

truncated normal distribution, $\beta_2 \sim TN_{R_2}(\mu_2, \sigma_2^2)$, where

$$\sigma_2^{-2} = \sigma_{\beta_2,0}^{-2} + \sum_{(i,t) \in \mathcal{A}} (q_{it,k_{it}})^2, \quad (\text{A.28})$$

$$\mu_2 = \sigma_2^2 \left\{ \sigma_{\beta_2,0}^{-2} \mu_{\beta_2,0} + \sum_{(i,t) \in \mathcal{A}} q_{it,k_{it}} (y_{it} - \beta_1 p_{it,k_{it}} - w_{it}^*) \right\}, \quad (\text{A.29})$$

$$R_2 = \left\{ \max_{i,t} (-\infty, BL_{it}^2), \min_{i,t,k} \left(BU_{it}^2, -\beta_1 \frac{p_{it,k+1} - p_{it,k}}{q_{it,k+1} - q_{it,k}} \right) \right\}, \quad (\text{A.30})$$

$$(BL_{it}^2, BU_{it}^2) = \begin{cases} \left(\frac{\bar{y}_{it,k-1} - \beta_1 p_{it,k} - w_{it}^*}{q_{it,k}}, \frac{\bar{y}_{it,k} - \beta_1 p_{it,k} - w_{it}^*}{q_{it,k}} \right), & \text{if } (i,t) \in \mathcal{A}, \\ \left(\frac{\bar{y}_{it,k} - \beta_1 p_{it,k} - w_{it}^*}{q_{it,k}}, \frac{\bar{y}_{it,k} - \beta_1 p_{it,k+1} - w_{it}^*}{q_{it,k+1}} \right), & \text{otherwise.} \end{cases} \quad (\text{A.31})$$

Step 4. Generate $(\sigma_v^2, \{\delta_i\}_{i=1}^n)$ given $\{\mathbf{w}_i^*\}_{i=1}^n, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta$. Integrating the joint full conditional probability density of $(\sigma_v^2, \{\delta_i\}_{i=1}^n)$ with respect to $\{\delta_i\}_{i=1}^n$, we have the full conditional distribution of σ_v^2 as the inverse gamma distribution, $\sigma_v^2 \sim IG(n_{v,1}/2, S_{v,1}/2)$. Then, the full conditional distribution of δ_i is the multivariate normal distribution, $\delta_i | \sigma_v^2 \sim N_d(\boldsymbol{\mu}_{\delta_{i,1}}, \sigma_v^2 \boldsymbol{\Sigma}_{\delta_{i,1}})$. Parameters of these full conditionals are $n_{v,1} = n_{v,0} + nT$,

$$S_{v,1} = S_{v,0} + n \boldsymbol{\mu}'_\delta \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta + \sum_{i=1}^n (\mathbf{w}_i^{*'} \mathbf{w}_i^* - \boldsymbol{\mu}'_{\delta_{i,1}} \boldsymbol{\Sigma}_{\delta_{i,1}}^{-1} \boldsymbol{\mu}_{\delta_{i,1}}), \quad (\text{A.32})$$

$$\boldsymbol{\mu}_{\delta_{i,1}} = \boldsymbol{\Sigma}_{\delta_{i,1}} (\boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta + \mathbf{Z}'_i \mathbf{w}_i^*), \quad \boldsymbol{\Sigma}_{\delta_{i,1}}^{-1} = \boldsymbol{\Sigma}_\delta^{-1} + \mathbf{Z}'_i \mathbf{Z}_i. \quad (\text{A.33})$$

Step 5. Generate $\boldsymbol{\mu}_\delta$ given $\{\delta_i\}_{i=1}^n, \boldsymbol{\Sigma}_\delta, \sigma_v^2$. The full conditional distribution of $\boldsymbol{\mu}_\delta$ is the multivariate normal distribution, $\boldsymbol{\mu}_\delta \sim N_d(\boldsymbol{\mu}_{\bar{\delta},1}, \boldsymbol{\Sigma}_{\bar{\delta},1})$, where

$$\boldsymbol{\mu}_{\bar{\delta},1} = \boldsymbol{\Sigma}_{\bar{\delta},1} \left(\boldsymbol{\Sigma}_{\bar{\delta},0}^{-1} \boldsymbol{\mu}_{\bar{\delta},0} + \sigma_v^{-2} \boldsymbol{\Sigma}_\delta^{-1} \sum_{i=1}^n \delta_i \right), \quad \boldsymbol{\Sigma}_{\bar{\delta},1}^{-1} = \boldsymbol{\Sigma}_{\bar{\delta},0}^{-1} + n \sigma_v^{-2} \boldsymbol{\Sigma}_\delta^{-1}. \quad (\text{A.34})$$

Step 6. Generate $\boldsymbol{\Sigma}_\delta$ given $\{\delta_i\}_{i=1}^n, \boldsymbol{\mu}_\delta, \sigma_v^2$. The full conditional distribution of $\boldsymbol{\Sigma}_\delta$ is the inverse

Wishart distribution, $\boldsymbol{\Sigma}_\delta \sim IW_d(n_{\bar{\delta},1}, \mathbf{S}_{\bar{\delta},1})$, where $n_{\bar{\delta},1} = n_{\bar{\delta},0} + n$ and

$$\mathbf{S}_{\bar{\delta},1}^{-1} = \mathbf{S}_{\bar{\delta},0}^{-1} + \sigma_v^{-2} \sum_{i=1}^n (\boldsymbol{\delta}_i - \boldsymbol{\mu}_\delta)(\boldsymbol{\delta}_i - \boldsymbol{\mu}_\delta)'. \quad (\text{A.35})$$

Step 7. Generate (s_{it}^, w_{it}^*) given $\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2$ for $i = 1, \dots, n$ and $t = 1, \dots, T$. The full conditional distribution of s_i^* is the multinomial distribution. Its probability mass function is given by*

$$\pi(s_{it}^* = s \mid \boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2) \propto \tau_s \left[\Phi\left\{\tau_s^{-1}(RU_{it,s} - \theta_{it,s})\right\} - \Phi\left\{\tau_s^{-1}(RL_{it,s} - \theta_{it,s})\right\} \right] \exp\left(-\frac{m_{it,s}}{2}\right), \quad (\text{A.36})$$

for $s = 1, \dots, 2K^{it} - 1$, where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, $RU_{it,s}$ and $RL_{it,s}$ denote the respective upper and lower limits of $R_{it,s}$ (see (17)), and

$$(m_{it,s}, \theta_{it,s}, \tau_s) = \begin{cases} \left(\frac{\sigma_u^{-2}\sigma_v^{-2}(y_{it} - \mathbf{x}'_{it,k}\boldsymbol{\beta} - \mathbf{z}'_{it}\boldsymbol{\delta}_i)^2}{\sigma_u^{-2} + \sigma_v^{-2}}, \frac{\sigma_u^{-2}(y_{it} - \mathbf{x}'_{it,k}\boldsymbol{\beta}) + \sigma_v^{-2}\mathbf{z}'_{it}\boldsymbol{\delta}_i}{\sigma_u^{-2} + \sigma_v^{-2}}, (\sigma_v^{-2} + \sigma_u^{-2})^{-1} \right), \\ \text{if } s = 2k - 1 \text{ and } k = 1, \dots, K^{it}, \\ \left(\sigma_u^{-2}(y_{it} - \bar{y}_{it,k})^2, \mathbf{z}'_{it}\boldsymbol{\delta}_i, \sigma_v^{-2} \right), \text{ if } s = 2k \text{ and } k = 1, \dots, K^{it} - 1. \end{cases} \quad (\text{A.37})$$

Given $s_{it}^* = s$, we generate w_{it}^* from the truncated normal distribution, $w_{it}^* | s_{it}^* = s \sim TN_{R_{it,s}}(\theta_{it,s}, \tau_s^2)$.

Step 8. Generate σ_u^2 given $\boldsymbol{\beta}, \{\mathbf{s}_i^, \mathbf{w}_i^*\}_{i=1}^n$. The full conditional distribution of σ_u^2 is the inverse gamma distribution, $\sigma_u^2 \sim IG(n_{u,1}/2, S_{u,1}/2)$, where $n_{u,1} = n_{u,0} + 2 + nT$ and*

$$S_{u,1} = S_{u,0} + (\boldsymbol{\beta} - \boldsymbol{\mu}_{\boldsymbol{\beta},0})' \boldsymbol{\Sigma}_{\boldsymbol{\beta},0}^{-1} (\boldsymbol{\beta} - \boldsymbol{\mu}_{\boldsymbol{\beta},0}) + \sum_{i=1}^n (\mathbf{y}_i - \mathbf{y}_i^*)' (\mathbf{y}_i - \mathbf{y}_i^*). \quad (\text{A.38})$$

Step 9. Go to Step 2.

A.2 AR(1) EC model

To incorporate a serial correlation into the RC model, we consider AR(1) process for v_{it} ,

$$\begin{aligned} w_{it}^* &= \mathbf{z}'_{it} \boldsymbol{\delta}_i + v_{it}, \\ v_{it} &= \gamma v_{i,t-1} + \eta_{it}, \quad \eta_{it} \sim N(0, \sigma_v^2), \quad v_{i0} \sim N\left(0, (1 - \gamma^2)^{-1} \sigma_v^2\right), \end{aligned} \quad (\text{A.39})$$

where η_{it} is independent of $\eta_{it'}$ ($t \neq t'$) and $|\gamma| < 1$. We call (A.39) an AR(1) error component (AR(1) EC) model.

It is straightforward to implement an MCMC method for the AR(1) EC model. All prior distributions except for γ are assumed to be the same as those for the RC model ((18) and (19)). For γ , we assume a uniform prior on an interval $(-1, 1)$ given by

$$\gamma \sim U(-1, 1), \quad (\text{A.40})$$

and it is assumed to be independent of other parameters. Thus, the joint prior probability density function for the AR(1) EC model parameters is

$$\pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta, \gamma) = \pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta) U(\gamma | -1, 1). \quad (\text{A.41})$$

Taking account of conditional distributions for the heterogeneities,

$$w_{i0}^* \sim N\left(0, (1 - \gamma^2)^{-1} \sigma_v^2\right),^9 \quad w_{it}^* | w_{i,t-1}^* \sim N\left(\mathbf{z}'_{it} \boldsymbol{\delta}_i + \gamma(w_{i,t-1}^* - \mathbf{z}'_{i,t-1} \boldsymbol{\delta}_i), \sigma_v^2\right) \text{ for } t \geq 1, \quad (\text{A.42})$$

the posterior probability density function for the AR(1) EC model is given by

$$\begin{aligned}
\pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i, \mathbf{s}_i^*, \mathbf{w}_i^*\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta, \gamma \mid \{\mathbf{y}_i\}_{i=1}^n) &\propto \pi(\boldsymbol{\beta}, \{\boldsymbol{\delta}_i\}_{i=1}^n, \sigma_u^2, \sigma_v^2, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta, \gamma) \\
&\times \sigma_v^{-n} (\sigma_u^{-nT} \sigma_v^{-nT}) (1 - \gamma^2)^{n/2} \exp \left[-\frac{1}{2} \left\{ (1 - \gamma^2) \sigma_v^{-2} \mathbf{w}_0^* \mathbf{w}_0^* \right. \right. \\
&\quad \left. \left. + \sum_{i=1}^n (\sigma_u^{-2} (\mathbf{y}_i - \mathbf{y}_i^*)' (\mathbf{y}_i - \mathbf{y}_i^*) + \sigma_v^{-2} (\tilde{\mathbf{w}}_i^* - \tilde{\mathbf{Z}}_i \boldsymbol{\delta}_i)' (\tilde{\mathbf{w}}_i^* - \tilde{\mathbf{Z}}_i \boldsymbol{\delta}_i)) \right\} \right] \\
&\quad \times \prod_{i=1}^n \prod_{t=1}^T \left\{ I(w_{it}^* \in R_{it, s_{it}^*}) \prod_{k=1}^{K^{it}-1} I(\mathbf{x}'_{it, k+1} \boldsymbol{\beta} \leq \mathbf{x}'_{it, k} \boldsymbol{\beta}) \right\},
\end{aligned} \tag{A.43}$$

where $\mathbf{w}_0^* = (w_{10}^*, w_{20}^*, \dots, w_{n0}^*)'$, $\tilde{\mathbf{w}}_i^* = \mathbf{w}_i^* - \gamma \mathbf{w}_{i-1}^*$, $\tilde{\mathbf{Z}}_i = \mathbf{Z}_i - \gamma \mathbf{Z}_{i-1}$, $\mathbf{w}_{i-1}^* = (w_{i0}^*, w_{i1}^*, \dots, w_{iT-1}^*)'$, and $\mathbf{Z}_{i-1} = (\mathbf{z}_{i0}, \mathbf{z}_{i1}, \dots, \mathbf{z}_{iT-1})'$.

The full conditional posterior distributions of $\boldsymbol{\beta}, \boldsymbol{\mu}_\delta, \boldsymbol{\Sigma}_\delta, \sigma_u^2$ are identical to those of the RC model derived in Appendix A.1. Other full conditional distributions are described as follows.

Conditional posterior distributions of $(\sigma_v^2, \{\boldsymbol{\delta}_i\}_{i=1}^n)$. The blocking technique (Step 4 of Appendix A.1) is applied to draw samples of $(\sigma_v^2, \{\boldsymbol{\delta}_i\}_{i=1}^n)$. Then, the full conditional distributions of σ_v^2 and $\boldsymbol{\delta}_i$ are the inverse gamma, $\sigma_v^2 \sim IG(n_{v,1}/2, S_{v,1}/2)$ and the multivariate normal, $\boldsymbol{\delta}_i \mid \sigma_v^2 \sim N_d(\boldsymbol{\mu}_{\delta_i,1}, \sigma_v^2 \boldsymbol{\Sigma}_{\delta_i,1})$, respectively, where $n_{v,1} = n_{v,0} + n(T+1)$,

$$S_{v,1} = S_{v,0} + n \boldsymbol{\mu}'_\delta \boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta + (1 - \gamma^2) \mathbf{w}_0^* \mathbf{w}_0^* + \sum_{i=1}^n (\tilde{\mathbf{w}}_i^* \tilde{\mathbf{w}}_i^* - \boldsymbol{\mu}'_{\delta_i,1} \boldsymbol{\Sigma}_{\delta_i,1}^{-1} \boldsymbol{\mu}_{\delta_i,1}), \tag{A.44}$$

$$\boldsymbol{\mu}_{\delta_i,1} = \boldsymbol{\Sigma}_{\delta_i,1} (\boldsymbol{\Sigma}_\delta^{-1} \boldsymbol{\mu}_\delta + \tilde{\mathbf{Z}}_i' \tilde{\mathbf{w}}_i^*), \quad \boldsymbol{\Sigma}_{\delta_i,1}^{-1} = \boldsymbol{\Sigma}_\delta^{-1} + \tilde{\mathbf{Z}}_i' \tilde{\mathbf{Z}}_i. \tag{A.45}$$

Conditional posterior distributions of $w_{i0}^, (s_{it}^*, w_{it}^*)$.* The posterior distribution of w_{i0}^* is the

⁹We set $\mathbf{z}_{i0} \equiv \mathbf{0}$.

normal distribution, $w_{i0}^* \sim N(\gamma(w_{i1}^* - \mathbf{z}'_{i1}\boldsymbol{\delta}_i), \sigma_v^2)$. We draw (s_{it}^*, w_{it}^*) for $i, t \geq 1$ from the similar full conditional distributions to those of the RC model replacing $\mathbf{z}'_{it}\boldsymbol{\delta}_i$ by $\mathbf{z}'_{it}\boldsymbol{\delta}_i + \gamma(w_{i,t-1}^* - \mathbf{z}'_{i,t-1}\boldsymbol{\delta}_i)$.

Conditional posterior distribution of γ . The full conditional posterior probability density function is given by

$$\pi(\gamma \mid \{\boldsymbol{\delta}_i, \mathbf{w}_i^*\}_{i=1}^n) \propto (1 - \gamma^2)^{n/2} \exp\left\{-\frac{(\gamma - \mu_\gamma)^2}{2\sigma_v^2\sigma_\gamma^2}\right\} I\{\gamma \in (-1, 1)\}, \quad (\text{A.46})$$

where

$$\mu_\gamma = \sigma_\gamma^2 \sum_{i=1}^n (\mathbf{w}_i^* - \mathbf{Z}_i\boldsymbol{\delta}_i)' (\mathbf{w}_{i,-1}^* - \mathbf{Z}_{i,-1}\boldsymbol{\delta}_i), \quad \sigma_\gamma^{-2} = \sum_{i=1}^n (\mathbf{w}_i^* - \mathbf{Z}_i\boldsymbol{\delta}_i)' (\mathbf{w}_i^* - \mathbf{Z}_i\boldsymbol{\delta}_i). \quad (\text{A.47})$$

We adopt the MH algorithm to draw samples of γ . Using the normal approximation to this density, we generate a candidate γ' from the proposal distribution $TN_{(-1,1)}(\mu_\gamma, \sigma_v^2\sigma_\gamma^2)$, and accept γ' with probability

$$\alpha(\gamma, \gamma') = \min\left\{1, \left(\frac{1 - \gamma'^2}{1 - \gamma^2}\right)^{n/2}\right\}. \quad (\text{A.48})$$

Acknowledgment

This work is supported by the Grants-in-Aid for Scientific Research 21243018 from the Japanese Ministry of Education, Science, Sports, Culture and Technology and by the Academic Frontier Project of the Nihon University Population Research Institute. The authors thank H. Kozumi (Kobe University) and K. Kakamu (Chiba University) for their valuable comments. All computational results were obtained using Ox 5.10 for Linux (see Doornik, 2002).

References

- Burtless, G. and J. A. Hausman (1978). The effect of taxation on labor supply: Evaluating the Gary negative income tax experiment. *Journal of Political Economy* 86(6), 1103–1130.
- 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.
- da Motta, R. S., R. M. Huber, and H. J. Ruitenbeek (1998). Market based instruments for environmental policymaking in latin america and the caribbean: lessons from eleven countries. *Environment and Development Economics* 4, 177–201.
- Dalhuisen, J. M., R. J. G. M. Florax, H. L. F. de Groot, and P. Nijkamp (2003). Price and income elasticities of residential water demand: A meta-analysis. *Land Economics* 79(2), 292–308.
- de Jong, G. C. (1990). An indirect utility model of car ownership and private car use. *European Economic Review* (34), 971–985.
- Doornik, J. A. (2002). *Object-Oriented Matrix Programming Using Ox* (3rd ed.). London: Timberlake Consultants Press and Oxford.
- 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.
- Gupta, A. K. and D. K. Nagar (2000). *Matrix Variate Distributions*, Volume 104 of *Monographs and Surveys in Pure and Applied Mathematics*. Boca Raton: Chapman and Hall.
- Hausman, J. A. (1985). The econometrics of nonlinear budget sets. *Econometrica* 53(6), 1255–1282.

- Herriges, J. A. and K. K. King (1994). Residential demand for electricity under inverted block rates: Evidence from a controlled experiment. *Journal of Business and Economic Statistics* 12(4), 419–430.
- 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.
- Iyengar, R. (2004). A structural demand analysis for wireless services under nonlinear pricing schemes. unpublished working paper, Marketing Department, Wharton School, University of Pennsylvania.
- Miyawaki, K., Y. Omori, and A. Hibiki (2010). Bayesian estimation of demand functions under block rate pricing. University of Tokyo CIRJE Discussion Paper Series CIRJE-F-712, Revised version of CIRJE-F-631, CIRJE-F-424, and CIRJE-F-568.
- Moffitt, R. (1986). The econometrics of piecewise-linear budget constraint. *Journal of Business and Economic Statistics* 4(3), 317–328.
- Moffitt, R. (1989). Estimating the value of an in-kind transfer: the case of food stamps. *Econometrica* 57(2), 385–409.
- Olmstead, S. M. (2009). Reduced-form vs. structural models of water demand under nonlinear prices. *Journal of Business and Economic Statistics* 87(1), 84–94.
- Olmstead, S. M., W. M. Hanemann, and R. N. Stavins (2007). Water demand under alternative price structures. *Journal of Environmental Economics and Management* 54(2), 181–198.
- Pint, E. M. (1999). Household responses to increased water rates during the California drought. *Land Economics* 75, 246–266.
- Reiss, P. C. and M. W. White (2005). Household electricity demand, revisited. *Review of Economic Studies* 72, 853–883.

- Rietveld, P., J. Rouwendal, and B. Zwart (2000). Block rate pricing of water in Indonesia: An analysis of welfare effects. *Bulletin of Indonesian Economic Studies* 36(3), 73–92.
- Robinson, J. A. (1997). Long-term consequences of population growth: Technological change, natural resources, and the environment. In M. R. Rosenzweig and O. Stark (Eds.), *Handbook of Population and Family Economics*, Volume 1B, Chapter 21, pp. 1175–1298. Amsterdam: North-Holland.
- Song, I. and P. K. Chintagunta (2007). A discrete-continuous model for multicategory purchase behavior of households. *Journal of Marketing Research* 44(4).
- Tanner, M. A. and W. H. Wong (1987). The calculation of posterior distributions by data augmentation. *Journal of American Statistical Association* 82(398), 528–540.