A nonparametric Bayesian approach for counterfactual prediction with an application to the Japanese private nursing home market

Shinya Sugawara

Japan Society of Promotion of Science, University of Tokyo

23. October 2012

A nonparametric Bayesian approach for counterfactual prediction with an application to the Japanese private nursing home market

Shinya Sugawara

Research Fellow of Japan Society of Promotion of Science, University of Tokyo

Abstract

This paper proposes a new inferential framework for structural econometric models using a nonparametric Bayesian approach. Although estimation methods based on moment conditions can employ a flexible estimation without distributional assumptions, they have difficulty conducting a prediction analysis. I propose a nonparametric Bayesian methodology for an estimation and prediction analysis. My methodology is applied to an empirical analysis of the Japanese private nursing home market. This market has a sticky economic circumstance, and my prediction simulates an intervention that removes this circumstance. The prediction result implies that the outdated circumstance in this market is harmful for consumers today.

Keywords: Nonparametric Bayes; Nonlinear simultaneous equation model; Prediction; Industrial organization; Nursing home; Long-term care in Japan

1 Introduction

This paper proposes a new methodology for structural econometrics using a nonparametric Bayesian approach. In structural econometrics, especially when applied to the industrial organization, many studies adopt estimation techniques, such as the generalized method of moments(GMM), that rely on identification conditions in the form of moment conditions, not distributional assumptions. Alternatively, much empirical research focuses on further inferences in addition to estimation. Especially since the introduction of the important Lucas criterion(Lucas, 1976), the availability of prediction analysis for counterfactual economic situations is an important advantage of structural studies. For advanced inferences, such as a prediction, however, a numerical technique based on random number generation is often needed, and hence, distributional assumptions are necessary. Because of this requirement, we cannot always conduct prediction analysis with only moment conditions.

To resolve this problem, I adopt a nonparametric Bayesian approach. This inferential framework is based on a likelihood function that can represent an arbitrary distribution. Unlike the other nonparametric methodologies, such as kernel or spline functions, the nonparametric Bayes approach can yield a closed form of a predictive distribution function, which is required for numerical techniques. This property allows for simultaneous estimation without distributional assumptions and prediction analysis.

As an empirical application of my methodology, this paper analyzes the market of private nursing homes in Japan. This market has an economic circumstance in the form of a price mechanism where nursing homes assume all longevity risks of their residents. This uneven market structure was established when the private nursing homes were luxury goods. This situation has changed as a result of the launch of the radical national long-term care insurance program in Japan, and private homes today appear to be ordinary goods. Nevertheless, many homes still adopt this price mechanism. This price mechanism is apparently a relief for elders whose only income is their pension. However, a rational home must recover the financial burden of risks from a different channel of payment. Therefore, there is a possibility that this circumstance forces consumers to pay a larger amount of their lifetime payments than if this price mechanism was nonexistent. My research question is whether this mechanism is beneficial or not for consumers.

The structure of this market is understood with a nonlinear simultaneous equation model that extends that of Berry et al. (1995), hereafter BLP. For this model, the GMM estimation is a standard tool because it can easily manage the simultaneity of many structural equations. I intend to extend the original BLP model, which consists of demand and supply for a static good, to express the dynamic nature of the nursing home market. The demand side in my model is the consumer optimization of the present value of lifetime utility. For the supply side, following the conventional approach of prior nursing home studies as summarized in Norton (2000), it is assumed that homes maximize their profit at the equilibrium.

To evaluate the economic impact of the peculiar price mechanism, a prediction analysis is conducted to simulate an exogenous intervention that removes the mechanism. A previous study by Nevo (2000) provided a prediction technique for the BLP model. In an analysis of mergers in the US ready-to-eat cereal industry, he assumes that the effects of a merger would be so marginal that everything in the model,including unobserved terms, remains the same. Under this assumption, the prediction analysis is a deterministic problem that does not require any distributional assumption for the unobserved variables. Due to its tractability, this methodology has been adopted by many researchers, such as in the study of Petrin (2002) on the impact of a new product, the minivan, on the automobile market. Alternatively, in the Japanese nursing home market, a change in the price mechanism would have a fundamental, not marginal, influence on the economy. Then, a more flexible method, such as my nonparametric Bayesian approach, could simulate the counterfactual intervention in the nursing home market.

I employ an empirical analysis using real data taken from the list of nursing homes in the book Shuukan Asahi Mook (2011). My estimation results produce reasonable estimates for each model parameter. Furthermore, my prediction analysis shows that the intervention reduces the lifetime total payments for residents, except those who have unrealistically long lifetimes. A possible reason for the current overpayment is consumers' inaccuracy in predicting their remaining lifetime. This result implies that the outdated circumstance is harmful for consumers today.

This paper contributes to three fields: nonparametric Bayesian statistics, applied econometrics for the industrial organization field, and empirical studies of elder care. First, in the literature of statistics, this study provides a new application fields for the nonparametric Bayesian method. Nonparametric Bayes methods have a long history from when Ferguson (1973) presented the Dirichlet process, which is the most commonly used model in this literature today. Although the practical use of nonparametric Bayes has expanded, only recently, along with the development of computer-intensive Bayesian methodologies such as the Markov chain Monte Carlo(MCMC), the method has been applied to various statistical fields, as surveyed by Hjort et al. (2010). In econometrics, Hirano (2002) applied a nonparametric Bayesian model called the Dirichlet process mixture to dynamic panel models. Alternatively, this paper is the first to apply nonparametric Bayesian methodology to a prediction analysis in structural econometrics.

Second, as an applied econometric study, I present a new, flexible prediction technique for the BLP model. This model and its variants, such as those proposed by Berry et al. (2004) and Berry and Pakes (2007), have already been applied to various industrial markets, such as the automobile retail market by BLP themselves, a service sector by Davis (2006) and a durable good market by Gowrisankaran and Rysman (2012), among others. Considering such wide applications, our methodology can be useful for empirical industrial organization research. Regarding the statistical concerns of the BLP model, this paper follows prior Bayesian research such as that by Yang et al. (2003), Musalem et al. (2009) and Jiang et al. (2009). These studies adopt parametric assumptions to implement an efficient estimation algorithm via a Gibbs sampler. My methodology provides an advantage in its flexibility due to the nonparametric modeling and a disadvantage in terms of computational time because of the demanding estimation technique via the Metropolis-Hastings algorithm.

Third, this study contributes to empirical studies of long-term care for the elderly. In developed countries, as the Baby Boom generation reaches older ages, elder care has become an important policy issue. This situation has stimulated a wide variety of empirical studies of long-term care using recent economic tools, such as an analysis of long-term care insurance demands and information asymmetry by Finkelstein and McGarry (2006) and within-family bargaining games for parental care by Engers and Stern (2002) and Bryne et al. (2009). In the empirical industrial organization field, a study closely related to this one by Mehta (2006) analyzed the US nursing home market using the BLP model.

My study takes advantage of a unique dataset of the Japanese market. Japan is experiencing some of the most drastic population aging in the world. To address this situation, the country has adopted a national program of long-term care insurance with universal coverage. Since the establishment of this program, the service sector for long-term care has rapidly grown into a large industry. A purpose of this paper is to share the implications of the Japanese experience that the rest of the world might face in the future.

The rest of the paper is organized as follows. In Section 2, I provide a brief review of the Japanese private nursing home market. Section 3 introduces my econometric model, and Section 4 presents the corresponding econometric methodology. The proposed method is applied to real data in Section 5. Section 6 concludes the paper.

2 The Japanese private nursing home market

This section presents a brief review of the private nursing home market in Japan. I first describe the current long-term elder care situation, which is united under the national insurance program. Then, I provide to industrial details of the nursing home market, especially its peculiar price mechanism, which is a main target of my research.

2.1 The current status of elder care in Japan

Elder care once was a family task in Japan. The country has experienced overwhelmingly rapid population-aging; therefore, maintaining a voluntary care system has been difficult. To remove the burden from families, the government first situated long-term care in a welfare program till the 1970s and then in a medicine program. Because those programs were not specially designed to meet the complicated demands of long-term care, the programs' capacities were quickly exceeded. To manage the growing demand for elder care, the national long-term care insurance program(LTCI) was launched in the year 2000[1].

The LTCI is an insurance system with universal coverage. It covers long-term care costs for two categories of insured people: Category 1 consists of all elderly aged 65 years or more, and Category 2 consists of those aged 40 to 64 years with aging-related diseases. An insured person can ask the municipality authority to assign a care eligibility level for him or her. Care eligibility levels are based on several items, including activities of daily living(ADL) and instrumental activities of daily living(IADL). There are seven eligibility levels, Assistance Required 1 and 2 and Care Required 1 to 5; the latter levels cover costs for more intensive care. An upper bound of monetary coverage and a set of available care services are prescribed by assigned level. Out-of-pocket expenses are 10% of costs within the bound, and the rest are covered by the LTCI. Services beyond the upper bound are available, but the excess costs are completely out-of-pocket.

An important property of the LTCI is that it does not allow direct cash transfers to elders or care givers, unlike the programs in Germany. In other words, the LTCI covers only the care cost via the market, which has led to political debates. A main supporting force for this system was the feminist movement for the "socialization" of care. This movement demanded a release of women from providing voluntary care, which is traditionally assigned to female family members, especially to the wife of

[1]See Ikegami and Campbell (2000) for details about the beginning of the program, including the historical background. Tsutsui and Muramatsu (2007) summarized the current program after the 2005 reform.

the first son. The government consented to this coverage because without the cash allowance, many elders might not apply eligibility level assessment.

The LTCI has instantly created a large demand for long-term care services. This sector rapidly shifted to a big industry. Both volume and range increased for this sector. For volume, even during the long recession in the 2000s, the care service sector showed a continuous growth, not only in sales additionally in the size of the labor forces. The range of services has additionally increased in variety. In addition to institutional care, which is commonly supplied through a market in developed countries, other services include at-home care and short-term stay services.

2.2 The nursing home industry

This paper concentrates on nursing homes among the different forms of institutional care. There are both public and private homes in this sector[2]; their differences can be summarized in the following three features. First, public homes provide uniform care at uniform prices, while private homes provide divergent care at a variety of prices. An explicit distinction is in the eligibility of residents. Public homes accept only elders who cannot leave a bed by themselves, while private homes accept a variety of residents. Specifically, there are two general categories in private homes: homes in one category provide care services as an option, while homes in the other category provide care services as a default. In the former category, when a resident requires permanent long-term care, he or she needs to exit the home. In the latter category, a resident can stay in the home until his or her death. This paper focuses on the latter category.

[2]I use the terms 'public nursing home' and 'private nursing home' as translations of Japanese words 'Tokubetsu-Yougo Roujin Houmu' and 'Yuuryou Roujin Houmu,' respectively. Researchers have not reached an agreement on the English term for private homes. For example, Ikegami and Campbell (2000) called the private nursing home as 'residential care with private-pay,' and Nonaka et al. (2011) called the private nursing home as 'Quasi institutional care.' In contrast to those authors, I treat private homes as a form of nursing homes and juxtaposes them with public homes for two reasons. First, they have a function that matches the standard definition of a nursing home, to provide general long-term cares that does not specialize in medical care for those who permanently live in an institution. Second, the Japanese term 'Roujin Home,' which means nursing home, is commonly found in words 'Tokubetsu-Yougo Roujin Houmu' and 'Yuuryou Roujin Houmu,'. This must represent the fact that these institutions are perceived as similar service goods by Japanese people.

Second, public homes are operated by a municipal authority or a non-profit organization. For-profit organizations cannot enter the public home market, only the private home market[3]. Even though non-profit organizations also can operate private homes, our data indicate that more than 90% of private homes are for-profit.

Third, LTCI coverage is another difference between public and private homes. In public homes, the LTCI covers everything except "hotel costs," meaning rents and foods[4]. In private homes, the LTCI covers only costs strictly categorized as care costs. This type of cost is not included in an ordinary payment but is treated as a person-specific additional cost.

The co-existence of public and private homes has its roots in history[5]. The first Japanese nursing homes were established in the late nineteenth century by religious and philanthropic organizations as voluntary institutions for the poor and solitary elderly. The legal basis of these institutions was first provided in 1923 as a part of a governmental welfare program, but the amounts of the subsidies were limited. To manage their operating costs, homes accepted "free contract" dwellers whose financial status was beyond the eligibility level of the welfare program but who wanted to receive institutional care. However, the quality of homes occasioned considerable complaints from the free contractors because homes could provide only limited services in the range of the national welfare program.

Private nursing homes were formed during this time to meet the demand of the free contractors. Although the exact origin is blurry, a record from as early as 1948 has been found about an active private nursing home. In 1963, the Act on the Social Welfare Service for Elderly(Roujin Fukushi Hou) updated the legal system of the long-term care sector. This act prohibited public homes from accepting any free contract resident. The act additionally prescribed legal requirements for private nursing homes for the first time. In other words, this act explicitly separated public and private nursing homes.

[3]Mitchel et al. (2004) stated that for-profit firms are not allowed to enter the institutional care market. This statement is true because the authors define the institutional care only as public homes.

[4]The hotel costs for public homes were once in the range of coverage of the LTCI but were eliminated by the 2005 reform.

[5]This part is drawn primarily from Momose (1997) and Sudou (2006a,b).

Until the 1990's, the only residents in private homes were high income elders, and the public homes functioned as a safety net. Several luxury private homes attracted wide attention in the Japanese "bubble" economy in the late 1980s. Then the launch of the LTCI has had a considerable influence on the nursing home market.

Table 1 is here

Table 1 describes the number of facilities and residential capacities for public and private nursing homes from the years immediately before the launch of the LTCI until recently[6]. The ratio of private to public home capacity was approximately 1 to 2.5 in 2009, although the ratio was 1 to 9 in 1999. Clearly, the market for private homes has been expanding much more quickly than that of the public home market.

The slow growth of the public home market is a result of a regulatory policy, which was caused by a rapid increase in the financial burden. In the first several years of the LTCI, the number of eligible elders grew more quickly than prior governmental estimates[7]. To slow the budget expansion, municipalities suppressed the establishment of new homes. This policy induces a long waiting list of elders, estimated to be 421,000[8]. The elderly typically spend years in several rehabilitation facilities, which are temporary care institutions between hospitals and public homes, until a vacancy arises in a public home. Because of such exogenous restrictions on the supply of public homes, I do not consider the crowding-out effects of public homes on private home demand in this study.

Because the LTCI additionally covers costs for private homes, municipalities sought to control their number as well. To provide a legal basis for this motivation, the national government announced the "Regulation of Volume"(Souryou Kisei) in 2005[9].

[6]The numbers are taken from the Survey on Institutions and Establishments for Long-Term Care for public homes after 2000, and the Survey on Social Welfare Institutions for public homes in 1999 and private homes for the whole periods. The numbers can be traced through more recent years, but the figures after 2009 have a problem with consistency. Because the research agents have changed from the government to private firms, the response rate has drastically dropped.

[7]See Campbell et al. (2010).

[8]The figure is taken from the press release by the Japan Ministry of Health, Labor and Welfare, December 2009. http://www.mhlw.go.jp/stf/houdou/2r98520000003byd.html.

[9]As reviewed in the paper by Mehta (2006), there is a similar regulation for a number of the nursing homes in the United States that is called the control of need (CON), which restricts an expansion of the Medicaid budget.

This regulation stated that in 2014, the rate of elders who in care institutions must be less than 37% of those with eligibility levels of Care-Required or higher. This regulation forces many municipalities to prevent the entry of new private homes, because the average rate already reached 41% in 2004. After active debates on this regulation, the government abandoned it in 2012. As seen in Table 1, this regulation had an effect on the market, as implied by the reduction in the growth after 2005. This reduction implies that the amount of potential demand for the nursing homes must be larger than their actual supply.

2.3 A peculiar price mechanism of private nursing homes

Next, I explain the peculiar price mechanism of the private nursing home market. To live in a private home, typical contracts require a resident to make both of two forms of the payments. The first payment is a monthly fee that covers the costs of daily needs. The second payment is paid at the time of the resident's entrance to the home and is called an initial payment. The amount of the initial payment is determined as rent during an expiration period of which the length is predetermined by the home. If a resident exits the home before the expiration, the rents for the remaining periods are paid back. On the other hand, if a resident lives longer than the expiration, he or she does not need to pay additional rent[10].

Under this contract, homes assume all longevity risks of their residents. It is legally possible to offer a contract where the rent is collected not by the initial payment but by the monthly fee to avoid the longevity risk. However, the percent of homes that exercise a contract without an initial payment is less than 35% in our dataset. In addition, most of the homes offer only one expiration period, which implies that homes do not practice price discrimination to manage the longevity risk.

This mechanism is a residual of the past, when private homes were perceived as a luxury good and consumers paid an expensive initial payment with a long expiration.

[10] Some portion of the initial payment is called an initial depreciation, which is not returned when the resident exits before the expiration. In the empirical analysis, I assume that a home does not collect the initial payment if initial depreciation is 100% of the initial payment. Otherwise, I do not reflect the initial depreciation in the definition of the price variables for simplicity.

With such limited demand, the price mechanism could have been a Nash equilibrium in those days. However, as Table 1 indicates, the nursing home industry has been experiencing an overwhelming transition in which the private homes are more familiar goods for ordinary consumers as a substitute for the long waiting lists in the public homes. Thus, whether this sticky circumstance is preferable from the perspective of the social welfare today is uncertain.

A brief review of the nursing home market in the United States may be helpful. As summarized in Norton (2000), the United States has a very different system from Japan's, called the Medicaid spend-down. In this system, the care costs are completely paid by consumers. When consumers can no longer afford the costs, the government assume them in the form of Medicaid, which is a public insurance system for low-income individuals. This system is based on the perspective of the safety-net, which differs from the Japanese universal care policy.

3 A structural econometric model for the nursing home market

3.1 Defining a market environment

In this section, I construct a structural econometric model similar to the BLP model for the Japanese private nursing home market. To explore the properties of the nursing home market, I make several extensions to the conventional model. I begin with describing the basic setup and notation in this subsection and proceed to details of my extension in the following subsections.

There are M local markets that are geographically isolated. Each market m has both demand and supply. The demand side consists of I_m consumers, and the supply side consists of H_m private nursing homes. The consumers decide among H_m private homes and an outside option. The outside option represents anything other than a private home, such as public homes, formal at-home care or informal family care. The consumer population I_m contains all the consumers who might enter a private home. An important finding of Berry et al. (1995) found that all the parameters can be identified with only firms-side observations; consumer-side information is unnecessary. Due to this finding, the dataset for this study consists of the observed characteristics of $H = \sum_m H_m$ nursing homes.

The economy is assumed to be in equilibrium in two senses. First, the supply and demand sides equate as a general equilibrium. Second, the supply side is an oligopolistic game, which results in the Nash equilibrium strategy profiles. I assume that there is a unique Nash equilibrium. This assumption is stronger than the conventional assumption that the dataset is generated from a unique equilibrium among possibly multiple equilibria. As mentioned in footnote 12 of Berry et al. (1995, p.853), to conduct a policy prediction, the economy is required to play the same equilibrium under the counterfactual situation. This requirement is not guaranteed if we assume the uniqueness of only the realized data-generating process.

There are several observable variables for equilibrium prices: For each $h_m = 1, ..., H_m$, p_{h_m} is a monthly fee and F_{h_m} is an initial payment, which corresponds to rents for an expiration period, namely T_{h_m} months. I make a variable for a monthly rent f_{h_m} as $f_{h_m} = F_{h_m}/T_{h_m}$ if $T_{h_m} \neq 0$ or $f_{h_m} = 0$ if $T_{h_m} = 0$. f_{h_m} is assumed to be exogenously determined in the housing market. I introduce vector notations for home specific variables such as $\boldsymbol{p}_m = (p_{1_m}, ..., p_{H_m})'$ and $\boldsymbol{p} = (\boldsymbol{p}_1', ..., \boldsymbol{p}_M')'$ where each component is indexed as p_h for $h = 1, 2, ..., H$.

As a clear distinction between my model and the conventional, static framework of the BLP model, I consider the dynamic nature of nursing homes. An important factor for a dynamic model is consumers' lifetimes, which I denote τ_{i_m} for the i_mth consumer. Consumer lifetimes might be uncertain for both the homes and the consumers themselves. This particular uncertainty is carefully modeled in the following subsections.

3.2 Modeling the demand side

My model for the demand side represents consumer optimization. I assume that consumers maximize the present value of their lifetime utility. For simplicity, I assume that there is no voluntary exit from a home, and hence, any exit is due to the resident's death, to avoid the complication of endogenous exit decisions. Then, the control variable of consumers in the mth market is their one-shot entrance decision among H_m private homes and an outside option.

In formulating the dynamic optimization, there is uncertainty regarding each consumer's lifetime, which is a time horizon of the present value calculation. It is difficult to consider an expected utility with respect to the stochastic horizon. Instead, I assume that consumers evaluates their remaining lifetime with a subjective prediction, which might be misspecified. Although I do not construct a specific model, the misspecification can exist for various reasons, such as a simple fallacy or an arbitrary over-evaluation for a risk management. This predicted value is consumer-specific and denoted as $\tau_{i_m}^I$ for the i_mth consumer.

For the utility components, I add explicit functional assumptions in this paper. Specifically, the utility function for a period takes the form of a linear function of an observable $K_d \times 1$ vector $\tilde{\boldsymbol{x}}_{h_m}$, an unobservable home-specific effect ξ_{h_m} and an individual-home match specific effect $\tilde{\eta}_{i_m h_m}$. In the present values of the future utilities, each consumer i_m has private information regarding his or her remaining lifetime in month $\tau_{i_m} \geq 1$ and time-discount rate $\delta_{i_m} \in (0,1)$. Consequently, the present value of the lifetime utility of the consumer i_m from the choice of h_mth home is

$$U_{i_m h_m} = \sum_{t=1}^{\tau_{i_m}^I} \delta_{i_m}^{t-1} [\tilde{\boldsymbol{x}}_{h_m}' \tilde{\boldsymbol{\beta}}_d + \xi_{h_m} + \tilde{\eta}_{i_m h_m}] - P(p_{h_m}, f_{h_m}, T_{h_m}, \tau_{i_m}^I, \delta_{i_m})\alpha, \quad (3.1)$$

where $-\alpha$ measures a disutility from a unit expenditure in terms of the present value utility, and hence, $\alpha > 0$ is required. P denotes the present value of the payment stream defined as

$$P(p_{h_m}, f_{h_m}, T_{h_m}, \tau_{i_m}^I, \delta_{i_m}) = \sum_{t=1}^{\tau_{i_m}^I} \delta_{i_m}^{t-1} p_{h_m} + F_{h_m} - I[T_{h_m} \geq \tau_{i_m}^I] \delta_{i_m}^{\tau_{i_m}^I} (T_{h_m} - \tau_{i_m}^I) f_{h_m}. \quad (3.2)$$

The right-hand side of the above equation consists of three parts. The first term represents the present value of monthly fees during the lifetime. The second term is an initial payment in a lump sum, which does not depend on δ_{i_m} because it is paid in full upon the start of living in the home, i.e., $t = 1$. The third term corresponds to the returns to the initial payment if the consumer would dies before the expiration, i.e., $T_{h_m} \geq \tau^I_{i_m}$. A calculation yields

$$\begin{aligned} U_{i_m h_m} &= \frac{1-\delta_{i_m}^{\tau^I_{i_m}}}{1-\delta_{i_m}} \Bigg[\left(\tilde{\boldsymbol{x}}'_{h_m} \tilde{\boldsymbol{\beta}}_d - p_{h_m}\alpha + \xi_{h_m} + \tilde{\eta}_{i_m h_m} \right) \\ & \quad - F_{h_m} \frac{1-\delta_{i_m}}{1-\delta_{i_m}^{\tau^I_{i_m}}} \alpha + f_{h_m} \frac{1-\delta_{i_m}}{1-\delta_{i_m}^{\tau^I_{i_m}}} \delta_{i_m}^{\tau^I_{i_m}} I[T_{h_m} \geq \tau^I_{i_m}](T_{h_m} - \tau^I_{i_m})\alpha \Bigg] . \end{aligned} \tag{3.3}$$

In (3.3), I decompose the term of F_{h_m} into a mean and an individual variation as

$$-F_{h_m} \frac{1-\delta_{i_m}}{1-\delta_{i_m}^{\tau^I_{i_m}}} \alpha = F_{h_m}\alpha_F + \eta_{F,i_m}, \tag{3.4}$$

where $E[\eta_{F,i_m}] = 0$. I refine $\boldsymbol{x}_{h_m} = (\tilde{\boldsymbol{x}}'_{h_m}, F_{h_m})'$, $\boldsymbol{\beta}_d = (\tilde{\boldsymbol{\beta}}'_d, -\alpha_F)'$ and

$$\eta_{i_m h_m} = \eta_{F,i_m} + \tilde{\eta}_{i_m h_m} + f_{h_m} \frac{1-\delta_{i_m}}{1-\delta_{i_m}^{\tau^I_{i_m}}} \delta_{i_m}^{\tau^I_{i_m}} I[T_{h_m} \geq \tau^I_{i_m}](T_{h_m} - \tau^I_{i_m}). \tag{3.5}$$

Further, I define

$$\begin{aligned} V_{i_m h_m} &= [(1-\delta_{i_m})/(1-\delta_{i_m}^{\tau^I_{i_m}})] U_{i_m h_m} && (3.6) \\ &= \boldsymbol{x}'_{h_m}\boldsymbol{\beta}_d - p_{h_m}\alpha + \xi_{h_m} + \eta_{i_m h_m}. && (3.7) \end{aligned}$$

Now a consumer chooses to enter the h_mth home if $U_{i_m h_m} = \max_{k_m \in \{0,1,2,\ldots,H_m\}}\{U_{i_m k_m}\}$ or $V_{i_m h_m} = \max_{k_m \in \{0,1,2,\ldots,H_m\}}\{V_{i_m k_m}\}$, where the subscript 0 represents an outside option for which I assume $V_{i_m 0} = \eta_{i_m 0}$.

I do not observe individual consumers' decisions, only market shares of homes. To establish an econometric model without individual variations, I assume $\eta_{i_m h_m}$ follows the i.i.d. type I extreme value distribution and integrate it out from the demand side model as

$$
\begin{aligned}
s_{h_m} &= \int_{V_{imh_m}=\max_{j\in\{0,1,2,\dots,H_m\}}\{V_{imj}\}} V_{imh_m}\pi(d\boldsymbol{\eta}_m) \qquad (3.8)\\
&= \begin{cases} \frac{\exp[\boldsymbol{x}'_{h_m}\boldsymbol{\beta}_d - p_{h_m}\alpha+\xi_{h_m}]}{1+\sum_{k_m=1}^{H_m}\exp[\boldsymbol{x}'_{k_m}\boldsymbol{\beta}_d - p_{k_m}\alpha+\xi_{k_m}]} & \text{for} \quad h_m = 1,\dots,H_m \\ \frac{1}{1+\sum_{k_m=1}^{H_m}\exp[\boldsymbol{x}'_{k_m}\boldsymbol{\beta}_d - p_{k_m}\alpha+\xi_{k_m}]} & \text{for} \quad h_m = 0 \end{cases}. \qquad (3.9)
\end{aligned}
$$

To finish the demand side modeling, I let q_{h_m} be the logarithm of the share for $h_m = 1, 2, \dots, H_m$, which is expressed as

$$
q_{h_m} = \boldsymbol{x}'_{h_m}\boldsymbol{\beta}_d - p_{h_m}\alpha + \xi_{h_m} + \ln\Big(1 - \sum_{k_m=1}^{H_m}\exp(q_{k_m})\Big). \qquad (3.10)
$$

3.3 Modeling the supply side

On the supply side, I construct a model for the profit maximization of private nursing homes. Following previous nursing home studies, which are surveyed in Norton (2000), I assume that homes maximize their expected profit at a steady state for their occupancy status. This assumption is required to avoid a complicated situation in which there are residents whose durations of residence overlap. In the equilibrium, the marginal profit function from the ith consumer for the h_mth home takes the form

$$
\Pi_{h_m}(\tau_{im}) = p_{h_m} + f_{h_m} - mc_{h_m} - f_{h_m}I[T_{h_m} < \tau_{im}], \qquad (3.11)
$$

where mc_{h_m} is the marginal cost for the h_mth resident.

Homes maximize their expected profit in which the per capita profit is defined by equation (3.11). In the general equilibrium, the market clearing condition indicates that the expectation is taken with respect the consumer's subject value τ^I_{im} and not to the true value. For this purpose, I assume that homes know the distribution of the consumers' subjective values τ^I_{im}. Under this assumption, $I_m s_{h_m}$ is the equilibrium number of residents.

Homes play a Beltrand-type competition in which p_{h_m} and T_{h_m} are control variables. Because these two control variables are complementary, I assume that homes firstly decide T_{h_m} and choose p_{h_m} given T_{h_m}. I further assume that a home's expectations for their customers' lifetimes are determined only with T_{h_m} and f_{h_m} without p_{h_m}. In other words, homes choose T_{h_m} to control for whom they attract, and p_{h_m} determines how to manage expenditures given presumed customers. Let $Pr^s(T_{h_m} < \tau | f_{h_m}, T_{h_m})$ be a home's subjective probability that a resident in the home has a lifetime longer than T_{h_m} in the steady state. From the assumption on the decision process for T_{h_m} and p_m, this subjective probability does not depend on the monthly fee. Consequently, I obtain the expected profit function as

$$\Pi_{h_m} = I_m s_{h_m} \Big[p_{h_m} + f_{h_m} - mc_{h_m} - f_{h_m} Pr^s(T_{h_m} < \tau | f_{h_m}, T_{h_m}) \Big]. \tag{3.12}$$

I assume the existence of an interior solution for the profit maximization problem. The first order condition for p_{h_m} yields

$$\left\{ p_{h_m} - mc_{h_m} + f_{h_m} \Big[1 - Pr^s(T_{h_m} < \tau | f_{h_m}, T_{h_m}) \Big] \right\} \frac{\partial s_{h_m}}{\partial p_{h_m}} + s_{h_m} = 0. \tag{3.13}$$

The above equation can be more explicit using the general equilibrium value of s_{h_m} derived from the demand side. Specifically, I substitute the closed form of $\partial s_{h_m} / \partial p_{h_m}$ and obtain

$$p_{h_m} + \frac{1}{\alpha (s_{h_m} - 1)} + f_{h_m} [1 - Pr^s(T_{h_m} < \tau | f_{h_m}, T_{h_m})] = mc_{h_m}. \tag{3.14}$$

Following the conventional approach, I assume a log-linear form of the marginal cost function. Specifically, the logarithm of the marginal cost is set to be a linear function of an observable $K_s \times 1$ vector $\boldsymbol{w}_{h_m}$ and an unobservable home specific effect ω_{h_m}. Several manipulations yield

$$p_{h_m} = \exp(\boldsymbol{w}_{h_m}\boldsymbol{\beta}_s + \omega_{h_m}) - \frac{1}{\alpha[\exp(q_{h_m}) - 1]} - f_{h_m}\Big[1 - Pr^s(T_{h_m} < \tau | f_{h_m}, T_{h_m})\Big]\Bigg\}. \tag{3.15}$$

In the above first-order condition for p_{h_m}, I assume that the subjective probability function can be represented as

$$1 - Pr^s(T_{h_m} < \tau | f_{h_m}, T_{h_m}) = \Gamma(T_{h_m}, f_{h_m}; \boldsymbol{\gamma}) \tag{3.16}$$

$$\stackrel{def}{=} \frac{\exp(\gamma_0 + \gamma_{T1}T_{h_m} + \gamma_{T2}T_{h_m}^2 + \gamma_{f1}f_{h_m} + \gamma_{f2}f_{h_m}^2 + \gamma_{TF}T_{h_m}f_{h_m})}{1 + \exp(\gamma_0 + \gamma_{T1}T_{h_m} + \gamma_{T2}T_{h_m}^2 + \gamma_{f1}f_{h_m} + \gamma_{f2}f_{h_m}^2 + \gamma_{TF}T_{h_m}f_{h_m})}, \tag{3.17}$$

where $\boldsymbol{\gamma} = (\gamma_0, \gamma_{T1}, \gamma_{T2}, \gamma_{f1}, \gamma_{f2}, \gamma_{Tf})'$.

The above demand and supply sides modeling induces a simultaneous equation system that consists of $2H_m$ structural equations (3.10) and (3.15) for $h_m = 1, ..., H_m$. In this model, the dependent variables are $(\boldsymbol{q}_m, \boldsymbol{p}_m)$, unobserved variables are $(\boldsymbol{\xi}_m, \boldsymbol{\omega}_m)$, and the coefficient parameters are $\tilde{\boldsymbol{\theta}} = (\boldsymbol{\beta}_d', \boldsymbol{\beta}_s', \alpha, \boldsymbol{\gamma}')'$.

4 An econometric framework via nonparametric Bayes

This section details an econometric analysis for the model presented in the last section. I begin with a general form of the BLP model to explain the difficulty of using the GMM to conduct predictions. Then, I provide an alternative method via a nonparametric Bayesian approach for the general model. Next, I narrow my attention to my variant of the BLP model for the private nursing home market to present corresponding estimation procedures and a prediction. Last, I discuss identification conditions for this model and required assumptions.

4.1 Difficulty of the GMM for prediction analysis

I first explain a difficulty of estimation methodologies based only on moment conditions to conduct a prediction analysis. This statement holds for structural econometric models in general, but I concentrate on the BLP model in this paper.

Take any market m. For $h_m = 1, ..., H_m$, the BLP model can be written in a general form as

$$q_{h_m} = f_q(\boldsymbol{p}_m, \boldsymbol{q}_{(-h_m)}, \boldsymbol{w}_{h_m}, \omega_{h_m}; \boldsymbol{\theta}), \quad (4.1)$$

$$p_{h_m} = f_p(\boldsymbol{p}_{(-h_m)}, \boldsymbol{q}_m, \boldsymbol{x}_{h_m}, \xi_{h_m}; \boldsymbol{\theta}), \quad (4.2)$$

where $\boldsymbol{q}_{(-h_m)} = (q_1, ..., q_{h_m-1}, q_{h_m+1}, ..., q_{H_m})$ and $\boldsymbol{p}_{(-h_m)}$ is similarly defined. f_q and f_p denote nonlinear functions known up to the parameter $\boldsymbol{\theta}$. Particularly in my model, they represent the right-hand sides of (3.10) and (3.15). Equilibrium conditions yield that these $2H_m$ equations comprise a simultaneous equation system. Conventional estimation methods are employed based on moment conditions in the form of $E[z\xi] = E[z\omega] = 0$, where z is some instrument. For this nonlinear simultaneous model, the moment conditions induce more simpler estimation methods than a method via a likelihood function, which is a complicated joint distribution accompanied by $2H_m$ nonlinear structural equations. This is the reason why GMM estimation is a standard econometric tool for the BLP model.

As an illustrative example of a prediction analysis, let us consider a prediction problem for p given counterfactual $\boldsymbol{x}_m = \tilde{\boldsymbol{x}}$. For simplicity, I assume that the reduced form is analytically obtained for p_{h_m} as

$$p_{h_m} = g_p(\boldsymbol{x}_m, \boldsymbol{w}_m, \boldsymbol{\xi}_m, \boldsymbol{\omega}_m; \boldsymbol{\theta}). \quad (4.3)$$

There are three unknown factors, namely $\boldsymbol{\xi}_m$, $\boldsymbol{\omega}_m$ and $\boldsymbol{\theta}$, on the right-hand side of the above equations. For $\boldsymbol{\theta}$, GMM estimates can be used. However, there the problem of the unobservables $(\boldsymbol{\xi}_m, \boldsymbol{\omega}_m)$ still exists.

Nevo (2000) proposed a method to manage these unobserved variables. His study investigates the effects of a merger in the US cereal industry. In the prediction for a counterfactual merger, he assumed that everything other than the merger status is kept unchanged. Under this assumption, counterfactual values of the explanatory variables are predicted using the same values of $(\boldsymbol{\xi}_m, \boldsymbol{\omega}_m)$. Specifically, Nevo (2000) substituted the values of the estimated residuals into the reduced form equation (4.3). Due to its tractability, many studies adopt this methodology.

In my case, the counterfactual situation, a different payment mechanism imposes a more drastic change for the market structure than the counterfactual in those studies. To reflect this large change, I rather want to allow different values of $(\boldsymbol{\xi}_m, \boldsymbol{\omega}_m)$ under the counterfactual situation. Therefore, I set these unobservables as stochastic terms in the prediction analysis and integrate them from the reduced form (4.3).

Under the conventional moment condition, however, such integration is feasible only when the reduced form g_p takes specific forms. One example is the reduced form that is additive and separable with respect to $(\boldsymbol{\xi}_m, \boldsymbol{\omega}_m)$ such that

$$p_{h_m} = g_p(\boldsymbol{x}_m, \boldsymbol{w}_m, \boldsymbol{\xi}_m, \boldsymbol{\omega}_m; \boldsymbol{\theta}) = \tilde{g}_p(\boldsymbol{x}_m, \boldsymbol{w}_m; \boldsymbol{\theta}) + \boldsymbol{a}'\boldsymbol{\xi}_m + \boldsymbol{b}'\boldsymbol{\omega}_m, \tag{4.4}$$

where $\boldsymbol{a}$ and $\boldsymbol{b}$ are constant vectors. By multiplying both sides by z_m, we can integrate out the terms for $(\boldsymbol{\xi}_m, \boldsymbol{\omega}_m)$ using the moment conditions. However, it is difficult to guarantee such an assumption because g_p is a reduced form. In general, we need to conduct a numerical integration for $(\boldsymbol{\xi}_m, \boldsymbol{\omega}_m)$ using a Monte Carlo algorithm that requires distributional assumptions. In addition, because the model has as many as $2H_m$ simultaneous equations, another numerical step is often needed to obtain the reduced form.

4.2 Introducing a nonparametric Bayesian approach

In this subsection, I introduce a nonparametric Bayesian approach for both estimation and prediction without distributional assumptions. The nonparametric Bayesian analysis can be summarized as a statistical methodology via a likelihood function which can represent an arbitrary distribution. Unlike the other nonparametric models, such as kernel or spline methods, the nonparametric Bayes models are associated with well-defined closed forms for the likelihood function and predictive distributions.

This feature enables us to conduct a numerical integration in a prediction analysis.

I incorporate nonparametric modeling for ξ_{h_m} and ω_{h_m}. Specifically, I assume $\boldsymbol{x}_{h_m}$ and $\boldsymbol{w}_{h_m}$ include constant terms, and ξ_{h_m} and ω_{h_m} represent home-specific stochastic terms with constraints on their moments. Because it is difficult to adopt a mean restriction in nonparametric Bayes models, I incorporate medians conditions, such that $\text{Med}(\xi_{h_m}) = \text{Med}(\omega_{h_m}) = 0$ in this paper.

In the nonparametric Bayesian literature, there are two popular approaches for managing such median constraints. One is the Dirichlet process mixture with restrictions, such as implemented in the studies by Doss (1985) and Kottas and Gelfand (2001). Another approach is the Polya tree mixture, which I use in this paper. The Polya tree mixture was established by Hanson and Johnson (2002) as an extension of the Polya tree proposed by Ferguson (1974). This original Polya tree is a general, nonparametric Bayes model that includes the Dirichlet process as a special case. In the 1990s, several papers were published on the theoretical aspects of the Polya tree, such as those by Lavine (1992, 1994) and Mauldin et al. (1992). More recent papers have presented computational procedures with applications, such as those by Hanson (2006), Jara et al. (2009) and Hanson et al. (2011).

An intuitive way to define a nonparametric likelihood function is through predictive densities. Let us consider a problem of estimating a joint conditional distribution for $(\omega_1, \omega_2 ..., \omega_H)$ given $\tilde{\boldsymbol{\theta}}$, where a scalar random variable ω_h has a common support Ω for all i. Suppose we have nonparametric predictive densities $\pi(\omega_h | \omega_1, ..., \omega_{h-1}, \tilde{\boldsymbol{\theta}})$ for all $h = 1, ..., H$. Then, the joint conditional density function can be derived as

$$\pi(\omega_1, ..., \omega_H | \tilde{\boldsymbol{\theta}}) = \pi(\omega_1 | \boldsymbol{\theta}) \pi(\omega_2 | \omega_1, \tilde{\boldsymbol{\theta}}) \pi(\omega_3 | \omega_1, \omega_2, \tilde{\boldsymbol{\theta}}) ... \pi(\omega_H | \omega_1, ..., \omega_{H-1}, \tilde{\boldsymbol{\theta}}). \quad (4.5)$$

The main concern for this setting is a choice of a nonparametric prediction distribution. An intuitive candidate is a histogram, given previous ωs. The Polya tree mixture is a method that constructs nonparametric predictive distributions similar to the histogram. I describe the detailed definition in Appendix A.

The Polya tree mixture has three primitives that econometricians need to specify.

The first is the base measure, which is used to define the bins of the histogram. I employ $N(0, 1/\tau)$ as my base measure for the analysis of Japanese nursing homes. τ is a scale parameter that is to be estimated. The second is the hyperparameters of the Polya tree prior, α_{j,k_j} for $j = 1, ..., J$ and $k_j = 1, 2, ..., 2^j$. I adopt a conventional choice introduced by Hanson and Johnson (2002), which is $a_{j,k_j} = cj^2$ for all k_j with a constant c. The third is the truncation level J. Choices for c and J are case-specific subjects, and I discuss them further in the empirical analysis section.

4.3 A Bayesian estimation procedure

4.3.1 The likelihood function

From this subsection, I describe the inferential framework that is specific to the econometric model for the Japanese nursing home market, which consists of structural equations (3.10) and (3.15). Because of the mutual dependencies of the dependent variables, these structural equations cannot be directly used to define the likelihood function. Instead, I obtain the likelihood function using a change of variables from unobservables to dependent variables, as suggested by Chintagunta and Dubé (2005). For distributions of unobservable terms, I assume that they follow independent Polya tree mixtures whose scale parameters are τ_ξ and τ_ω. Then, the resulting likelihood function is

$$\begin{aligned}\pi(\boldsymbol{p}, \boldsymbol{q}|\boldsymbol{\theta}, \text{Data}) \quad = \quad & \Big[\prod_{m=1}^{M} |\det(\mathcal{J}_m)|\Big]\pi_{\boldsymbol{\omega},\boldsymbol{\xi}}[\omega_{1_1}(\boldsymbol{p}_1, \boldsymbol{q}_1; f_{1_1}, T_{1_1}, \tilde{\boldsymbol{\theta}}), \xi_{1_1}(\boldsymbol{p}_1, \boldsymbol{q}_1; f_{1_1}, T_{1_1}, \tilde{\boldsymbol{\theta}}), \\ & \omega_{2_1}(\boldsymbol{p}_1, \boldsymbol{q}_1; f_{2_1}, T_{2_1}, \tilde{\boldsymbol{\theta}}), \xi_{2_1}(\boldsymbol{p}_1, \boldsymbol{q}_1; f_{2_1}, T_{2_1}, \tilde{\boldsymbol{\theta}}), \\ & ..., \omega_{H_1}(\boldsymbol{p}_1, \boldsymbol{q}_1; f_{H_1}, T_{H_1}, \tilde{\boldsymbol{\theta}}), \xi_{H_1}(\boldsymbol{p}_1, \boldsymbol{q}_1; f_{H_1}, T_{H_1}, \tilde{\boldsymbol{\theta}}), \\ & ..., \omega_{H_M}(\boldsymbol{p}_M, \boldsymbol{q}_M; f_{H_M}, T_{H_M}, \tilde{\boldsymbol{\theta}}), \xi_{H_M}(\boldsymbol{p}_M, \boldsymbol{q}_M; f_{H_M}, T_{H_M}, \tilde{\boldsymbol{\theta}})], \quad (4.6)\end{aligned}$$

where $\mathcal{J}_m$ is the Jacobian matrix of the transformation, which is explicitly derived in Appendix B. $\boldsymbol{\theta} = (\tilde{\boldsymbol{\theta}}', \tau_\xi, \tau_\omega)'$ denotes a vector of all the parameters and

$$\xi_{h_m}(\boldsymbol{p}_m, \boldsymbol{q}_m; f_{h_m}, T_{h_m}, \tilde{\boldsymbol{\theta}}) = q_{h_m} - \ln\Big[1 - \sum_{k_m=1}^{H_m} \exp(q_{k_m})\Big] - \tilde{\boldsymbol{x}}'_{h_m}\tilde{\boldsymbol{\beta}}_d + T_{h_m} f_{h_m}\alpha_F + p_{h_m}\alpha, \tag{4.7}$$

$$\omega_{h_m}(\boldsymbol{p}_m, \boldsymbol{q}_m; f_{h_m}, T_{h_m}, \tilde{\boldsymbol{\theta}}) = \ln\left[p_{h_m} + \frac{1}{\alpha[\exp(q_{h_m}) - 1]} + f_{h_m}\Gamma(T_{h_m}, f_{h_m}; \boldsymbol{\gamma})\right] - \boldsymbol{w}'_{h_m}\boldsymbol{\beta}_s. \tag{4.8}$$

In the above likelihood function, I require additional restrictions for the supports of the dependent variables to have well-defined logarithmic terms. Specifically,

$$0 < 1 - \sum_{k_m=1}^{H_m} \exp(q_{k_m}), \tag{4.9}$$

$$0 < p_{h_m} + \frac{1}{\alpha[\exp(q_{h_m}) - 1]} + f_{h_m}\Gamma(T_{h_m}, f_{h_m}; \boldsymbol{\gamma}). \tag{4.10}$$

Condition (4.9) is automatically satisfied in the estimation step due to the construction of q_{h_m} in Section 3, but it must be verified in the prediction step described below. Another condition, (4.10), is required in the estimation and prediction steps. Furthermore, because (4.10) states that the support of the likelihood function depends on parameters α and $\boldsymbol{\gamma}$, (4.10) violates a regularity condition for maximum likelihood estimators to have preferable asymptotic properties. This fact is another motivation to adopt a Bayesian estimation procedure.

4.3.2 Prior and proposal distributions

I implement the Bayesian estimation using the Markov chain Monte Carlo(MCMC) algorithm. Due to the construction of the Polya tree mixture using histogram-like stochastic structures, the likelihood is not a smooth function of parameters. I then adopt the Metropolis-Hastings(MH) algorithm via random walk proposal distributions. The prior and proposal distributions are specified as follows.

For prior distributions, I assume that the coefficient parameters follow independent normal distributions and that the scale parameters of the Polya tree mixtures, each of which is an inverse of the variance, follow independent Gamma distributions. Specifically,

$$\boldsymbol{\beta}_d \sim N(\boldsymbol{\mu}_{\boldsymbol{\beta}_d 0}, \Sigma_{\boldsymbol{\beta}_d 0}), \quad \boldsymbol{\beta}_s \sim N(\boldsymbol{\mu}_{\boldsymbol{\beta}_s 0}, \Sigma_{\boldsymbol{\beta}_s 0}), \tag{4.11}$$

$$\boldsymbol{\gamma} \sim N(\boldsymbol{\mu}_{\boldsymbol{\gamma} 0}, \Sigma_{\boldsymbol{\gamma} 0}), \quad \alpha \sim N(\mu_{\alpha 0}, \sigma^2_{\alpha 0}), \tag{4.12}$$

$$\tau_\omega \sim Gamma(a_{\tau_\omega 10}, a_{\tau_\omega 20}), \quad \tau_\xi \sim Gamma(a_{\tau_\xi 10}, a_{\tau_\xi 20}), \tag{4.13}$$

where *Gamma* denotes the Gamma distribution.

For proposal distributions, I choose distributions that can impose support conditions described so far. First, for unconstrained parameters $\boldsymbol{\beta}_d$ and $\boldsymbol{\beta}_w$, I use the normal proposal distributions. Second, for τ_ξ and τ_ω, the proposal distributions are set as log-normal distributions to guarantee their positivity. Third, a truncated normal proposal is incorporated for $\boldsymbol{\gamma}$ to satisfy the support conditions (4.9) and (4.10). Finally, for α, which must be positive and satisfy the support condition, I use the truncated log-normal proposal distribution.

4.4 Numerical techniques for a counterfactual prediction

The main purpose of this study is to simulate an exogenous intervention that eliminates the initial payment mechanism. Although we want to compare consumer welfare before and after the intervention, it is difficult to derive utility functions because a consumer lifetime and the time-discount factor are not observed. Instead, I conduct a comparison of the total amount of lifetime payments. My prediction analysis consists of two parts: the prediction for monthly fees p_{h_m} after the intervention and the calculation of the total payments both before and after the intervention.

To predict the monthly fee, I assume that $\eta_{i_m h_m}$ has the same distribution after the intervention. Considering the fact that $\eta_{i_m h_m}$ depends on T_{h_m} and f_{h_m}, this is a strong assumption but is technically required to conduct a prediction analysis. Given this assumption, the intervention yields the same economic model as (3.10) and (3.15) in which $f_{h_m}^{new} = 0$ and $T_{h_m}^{new} = 0$.

4.4.1 Algorithms to predict monthly fees

To predict p_{h_m}, I consider the predictive mean:

$$
\begin{aligned}
& E[p_{h_m}^{new} | f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \text{Data}] \qquad (4.14) \\
= & \int \int p_{h_m} \pi(p_{h_m} | T_{h_m}^{new} = 0, f_{h_m}^{new} = 0, Data, \boldsymbol{\theta}) \pi(\boldsymbol{\theta} | Data) d\boldsymbol{p_{h_m}} d\boldsymbol{\theta}.
\end{aligned}
$$

I conduct a dual-loop Monte Carlo integral to calculate the double integral numerically. Further, p_{h_m} must be integrated on the marginal distribution, which corresponds to the reduced form. Because of the mutual dependency of p_{h_m} and the other dependent variables, it is difficult to obtain a closed form of the reduced form analytically. Thus, a numerical solution for the simultaneous equation is required. The numerical procedure is summarized as follows. Let L and R be appropriately large integers as the numbers of the iterations for outer and inner loops of the Monte Carlo integration, respectively. The inner loop is accompanied by the numerical solution.

The outer loop approximates the integral with respect to $\boldsymbol{\theta}$. I generate random numbers $\boldsymbol{\theta}^l$, $l = 1, 2, ..., L$ from the posterior distribution of $\boldsymbol{\theta}|\text{Data}$. I can adopt the posterior samples of the MCMC estimation as the random numbers in this step.

The inner loop implements the integral for p_{h_m}. Given θ^l, I generate $p_{h_m}^{rl}$, $r = 1, 2, ..., R$ from the distribution of $p_{h_m}^{lr} | f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \boldsymbol{\theta}^l, \text{Data}$. To conduct the numerical solution, I implement the MCMC sampling for the inner loop. As summarized in Appendix C, I have closed forms for conditional predictive distributions for $q_{h_m}^{new} | \boldsymbol{p}_m^{new}, \boldsymbol{q}_{(-h_m)}^{new}, f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \boldsymbol{\theta}, \text{Data}$ and $p_{h_m}^{new} | \boldsymbol{q}_m^{new}, \boldsymbol{p}_{(-h_m)}^{new}, f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \boldsymbol{\theta}, \text{Data}$ for $h_m = 1, 2, ..., H_m$. Then, I iteratively draw the random samples from these conditional distributions, given the previous draws. After an appropriate length of the burn-in periods R', I have random draws from marginal predictive distributions, which can serve as p_{new}^{rl}.

Finally, I approximate the double integral by $(1/L) \sum_{l=1}^{L} (1/R) \sum_{r=1}^{R} p_{h_m}^{rl}$. In the above steps, the inner loop is computationally burdensome, but it can be conducted separately for each market. Due to this separability, I focus on a specific market in the empirical analysis.

4.4.2 Policy evaluation through predicting lifetime payments

Hereafter, I assume that the predicted values for a monthly fee after the intervention $p_{h_m}^{new}$ is already derived using the dual-loop Monte Carlo integral. To calculate the lifetime payments before and after the intervention, I begin with considering a match of a resident i_m and h_m. The total payment after the intervention is the accumulated monthly fees throughout the resident's lifetime τ_{i_m}, because there is no other form of a payment than the monthly fee. Then, the predicted total payment is $p_{h_m}^{new}\tau_{i_m}$. Because I do not consider an interest rate but use a simple summation, this amount is a lower bound of the present value of the consumer's total payment.

There is a difficulty in that the amount of the lifetime payment before the intervention depends on values of τ_{i_m} and T_{h_m}. To illustrate the problem, I separately consider three cases where the first two cases are not troublesome but the last case is problematic. The first case is $T_{h_m} = 0$, where the home does not collect an initial payment, even in the current situation. I abbreviate this case for my prediction analysis because we do not have a particular interest in this case. For the remaining two cases, we assume $T_{h_m} > 0$.

The second case is $\tau_{i_m} \leq T_{h_m}$. This case is denoted as "short-lived" because the i_mth consumer has a shorter lifetime than the expiration date. For this case, the lifetime payment before the intervention is $(p_{h_m} + f_{h_m})\tau_{i_m}$. Because the payments before and after the intervention are both multiplied by τ_{i_m}, I can cancel this term out when comparing the lifetime payments. Thus, the intervention effect can be detected through a comparison between $p_{h_m} + f_{h_m}$ and $p_{h_m}^{new}$, regardless of τ_{i_m}.

The third case is $\tau_{i_m} > T_{h_m}$, which is denoted as "long-lived". In this case, the consumer before the intervention does not need to pay rents after the expiration date. As a result, the lifetime payment before the intervention is $p_{h_m}\tau_{i_m} + f_{h_m}T_{h_m}$. Unlike the previous case, I cannot ignore the unobservable lifetime τ_{i_m} when comparing the lifetime payments. In practice, I consider several representative consumers whose lifetimes are $\tau_{i_m} = 240$ and 360. Because the lifetime is measured in months, these values correspond to 20 and 30 years of remaining life.

4.5 Identification

This subsection illustrates the identification conditions for my econometric framework. Identification of my model depends on three factors: distributional assumptions, functional form assumptions and exclusion restrictions, discussed below.

For the distributional and functional forms, I have made several assumptions so far. The assumptions as a whole do not allow to specify individual heterogeneity, as a price to implement a nonparametric Bayesian method. Specifically, there are two elements which are commonly included in the previous studies but eliminated in my study.

The first is random coefficient modeling. For example, in equation (3.4), the traditional approach includes individual variations in coefficients, whereas I locate them in the error term and integrate out. This is because the existence of such a random coefficient term makes it difficult to obtain closed form expressions of error terms as in (4.7) and (4.8), which are required for the construction of the likelihood and the predictive density functions.

Second, there may be a consumer heterogeneity which affects the distribution of the individual-specific term η. In this paper, I used a logit model for the distribution of this term. Third, we eliminate any endogeneity among unobservables. In the previous studies, such an endogeneity is assumed to exist and is controlled using the instrument. This paper ignores this endogeneity because of a technical difficulty of multivariate histograms. For the above two shortcuts, the flexibility of nonparametric Bayesian model can reflect these effects.

To specify the requirements of the exclusion restrictions, given (3.10) and (3.15), we can obtain conditional distributions for $q_{h_m}|p_{h_m}, \boldsymbol{q}_{(-h_m)}, \xi_{h_m}$ and $p_{h_m}|q_{h_m}, \omega_{h_m}$. Therefore, for the demand side, a standard exclusion restriction that is not correlated to p_{h_m} and $\boldsymbol{q}_{(-h_m)}$ but is correlated to q_{h_m} is required. For the supply side, I have conditional independency, such that p_{h_m} is independent from $\boldsymbol{p}_{(-h_m)}$ and $\boldsymbol{q}_{(-h_m)}$ given

q_{h_m}. Thus, having exclusion restrictions that are not home-specific but market-specific variables is sufficient.

5 An empirical study of the Japanese nursing home market

5.1 Data

This section applies my methodology to real data on the Japanese private nursing home market. In principle, the required information for this study is found in the public domain in the sense that private nursing homes are legally obligated to disclose the information when asked. Because it is burdensome to obtain a disclosure for all the homes, I refer to a list in a consumers' guidebook, Shuukan Asahi Mook (2011), which is a special volume of a leading weekly news magazine in Japan.

The local markets $m = 1, 2, ..., M$ are defined as prefectures, which are the largest subnational jurisdictions in Japan. An important assumption for the BLP model is that markets are geographically isolated. To guarantee this assumption, I need to incorporate a relatively large area as a market. The prefecture is an ideal unit for this purpose.

My sample consists of 1,265 homes. The details of the sampling methodology is summarized as follows: the editors of Shuukan Asahi Mook (2011) sent a questionnaires to all private homes except those that had a past legal fault. The population consists of "approximately 5,000" homes, in their words. They edited the book using 2,343 responses. Of the listed homes in the book, I eliminated 745 homes where long-term care is optional. From the remaining 1,598 homes, 324 homes are removed from the sample because of missing information. I further excluded 9 homes in prefectures that have only one home because the monopoly market would yield a different market structure to my oligopoly model. Approximately half of homes do not have a response. The low response rate might be caused by the enforcement of an early deadline by editors, specifically three months.

The share of a home s_{h_m} is defined as the ratio of the number of the residents in the h_mth home over the number of the potential consumers in the market m. As seen in equation (3.9), there must be a positive share for the outside option. Then, I included those who did not choose to live in any private nursing home in the potential consumers. I adopted the Category 1 elders with an eligibility level of Care Required 1 or higher, which is the minimum level typically required to receive institutional care under coverage of the LTCI, in the potential consumers. Table 2 shows numbers of the homes and Category 1 (age 65 years or more) elders in prefectures.

Table 2 is here

I have several observed variables related to prices: a monthly fee p_{h_m}, an initial payment F_{h_m} and an expiration period T_{h_m}. The monthly rent f_{h_m} is created using F_{h_m} and T_{h_m}, as mentioned earlier. Several homes report two price variables for their minimum and maximum. Specifically, the expiration period is unique for 1,226 homes, but the monthly fee and the initial payment are unique only for approximately half of the homes. Because there are only a few homes that offer multiple options for expiration periods, the variation in initial payments must be caused by a variation in monthly rents. Variations in monthly fees and monthly rents may be caused by quality differences in services and rooms. However, the lack of variation in expiration periods implies that separating equilibria as a tool to manage the longevity risk do not seem to occur.

I have two categories of explanatory variables: components of $\boldsymbol{x}$ shift the consumer utility for decision making, and components of $\boldsymbol{w}$ are characteristics of the marginal cost per resident. In addition to the common elements for these two categories, as mentioned in Section 4.5, I need exclusion restrictions that are home-specific variables only in $\boldsymbol{x}$ and market-specific variables only in $\boldsymbol{w}$.

For common observable elements on the demand and supply sides, I adopt three variables from Shuukan Asahi Mook (2011). The first variable is the number of residents per worker(**Worker**). This variable clearly affects the utility because it determines the amount of time a worker can spends on each resident. This variable additionally determine a labor cost, which is an important element of the marginal cost. For private homes that provide long-term care as a default option, this number is legally required to be 3 or fewer.

The second common variable is a dummy variable that takes a value of unity when the home is operated by a chain(**Chain**). The demand side effect of this variable can be both positive or negative, because a chain operation might imply either efficient operation or a stereotypical care. Its cost effect is also ambiguous because chains might operate more efficiently in service provisions but spend more on advertisements. In the private nursing home market, there is no dominant chain that has more than 20% of the share. Then I create the dummy variable as a bundle of six chains, namely Benesse Style Care, Message, Watami no Kaigo, Nichii Group, Life Commune and Tsukui, which account for more than 20 homes in our dataset[11].

The third common variable is years since opening(**Years**). On the demand side, that a home has survived for a long time might imply both high quality from the accumulation of experiences and disutility from old facilities. In addition, the supply side effect is indeterminate due to the coexistence of the accumulation of knowledge and high repair costs.

Next, I propose my exclusion restrictions. On the demand side, I adopt a home-specific variable of the occupancy rate(**Occupancy**). This variable affects the consumer utility because an extremely small occupancy rate might be a signal that the home has some problems. On the other hand, this variable does not affect the per-resident cost.

For the supply side exclusion restrictions, I use two market-specific variables of cost shifters: local averages of rents(**Rent**) and wages(**Wage**). These variables affect the marginal cost but not the utility, given the other price variables. **Rent** is defined as an annual average of monthly rents per 3.3 m^2 in the capital city of a prefecture,

[11]From a similar motive to the chain dummy, I tried to adopt another dummy variable that takes unity when the home is operated by a non-profit organization. However, I eliminate this variable from my empirical study because the convergence of its coefficient is quite slow, and there is no serious difference for coefficient estimates of the other variables, with or without this variable. The slow convergence might be caused by an insufficient sample size because this dummy variable takes unity for less than 10% of my sample.

which is taken from the 2010 Annual Report on the Retail Price Survey. For the average wage, there are no reliable data specific to care workers. Therefore, I adapt information of the medical and welfare sectors, which include care workers. **Wage** is defined as the quotient of annual wages plus bonuses over 12. These components are from the 2010 Basic Survey on Wage Structure.

Table 3 is here

Table 3 presents descriptive statistics for the explanatory variables. I adjust several volatile variables to stabilize our estimation in the following exercises of my inference. First, I standardize **Years**, **Rent** and **Wage** to have means of zero and unit variances. Second, I divide several variables generated from T and f by constants. Specifically, F is divided by 1000 and T, T^2, f and f^2, which appear in $\Gamma(T, f; \boldsymbol{\gamma})$, are divided by $10, 1000, 10$ and 100, respectively.

5.2 Estimation results

Before proceeding to the results, I determine case-specific components in my econometric frameworks. First, the hyperparameters are set as follows:

$$\boldsymbol{\beta}_d \sim N(\mathbf{0}, 1000I), \quad \boldsymbol{\beta}_s \sim N(\mathbf{0}, 1000I), \tag{5.1}$$

$$\boldsymbol{\gamma} \sim N(\mathbf{0}, 10I), \quad \alpha \sim N(1, 10), \tag{5.2}$$

$$\tau_\omega \sim Gamma(3, 10), \quad \tau_\xi \sim Gamma(3, 10). \tag{5.3}$$

I use normal proposal distributions, as mentioned above. I adjust their variances to have modest rates of acceptance in the MH algorithm. Specifically, the acceptance rates for parameters are located within the range between 0.29 and 0.6.

For the Polya tree mixture, I choose primitives as $c = 10$ and $J = 5$. I additionally adopt several alternative values to check the robustness of this choice. For c, 1, 100 and $1,000$ are incorporated. Among them, $c = 100$ and $1,000$ yield posterior samples similar to my primary result. $c = 1$, which is used in the studies by Hanson and

Johnson (2002) and Hanson (2006), yields similar posterior means for parameters but shows a slow convergence. Additionally, I check $J = 8$. This value is recommended by Hanson and Johnson (2002) as their rule of thumbs, namely $J \simeq log_2 H$. However, this value induces a slow convergence, although it yields similar posterior means to $J = 5$. Both $c = 1$ and $J = 8$ impose finer definitions of bins than $c = 10$ and $J = 5$. The slow convergence with these primitives is caused by the complexity of my model relative to the previous statistical papers.

In our implementation of the MCMC samplers, I generated 1,000,000 posterior samples after discarding 100,000 initial samples as the burn-in period. The computation took approximately 20 days using three cores of the Intel Xeon X5470 processor (3.33GHz).

Table 4 is here

Table 4 reports the estimation results. The first and second columns show posterior means and standard deviations, the third column represents 95% credible intervals, and the last column reports the inefficiency factors(IF). The maximum of the inefficiency factors is $12,880$, which implies that we would obtain the same variance of the posterior sample means from more than 75 uncorrelated draws, even in the worst case. For the sake of the convergence diagnosis, I additionally present figures of posterior sample paths and the posterior densities of the MCMC samples in Appendix D.

Overall, the estimated posterior means for coefficient parameters take reasonable values. On the demand side, **Worker** affects consumer utility negatively, because consumers prefer homes with a sufficient capacity of care workers. The positive coefficient for **Chain** indicates efficient service provision of chains. The effect of **Years** is not clear, indicating a complicated role of the history. In addition, the exclusion restriction **Occupancy** has a strongly positive effect on consumer utility, which implies that the popularity of a home is a good proxy for its quality. Furthermore, the mean effect of an initial payment on the utility, α_F, takes a strongly positive value. It is difficult to interpret this result alone, but at least it indicates a complicated role of the initial payment. I provide the further consideration in the prediction analysis later.

On the supply side, because the logarithm of the marginal cost is defined as $\log(mc_{h_m}) = \boldsymbol{w}'_{h_m}\boldsymbol{\beta}_s + \omega_{h_m}$, a positive coefficient means that the corresponding explanatory variable increases the marginal cost and hence decreases the profit of homes. The negative effect of **Worker** means that a reduction in the labor forces decreases the marginal cost, as expected. The positive coefficient for **Chain** implies that advertisement costs are higher than the revenue from the efficient operation of chains. **Years** has a negative effect, which can be interpreted as the accumulation of knowledge decreases the running costs. For the exclusion restrictions, **Rent** has a strongly positive coefficient as an increasing factor of the marginal cost. However, the sign of **Wage** is ambiguous. This ambiguity might imply that the local average wage of the medical and welfare sectors does not precisely capture the wages of institutional care workers in private nursing homes.

It is interesting to find a bimodal posterior density of the scale parameter of the Polya tree mixture τ_ω. This density is not caused by a problem of incomplete convergence but by the true posterior shape, as seen in the sample path in which the chain repeatedly visits both peeks. The peculiar shape of the posterior density function indicates that the distribution of ω_m is different from common probability distributions, such as the normal distribution. This result supports our usage of nonparametric modeling.

5.3 Prediction results

Next, I conduct a prediction analysis based on the above estimation result. In the dual-loop Monte Carlo integral, the number of the iterations for the outer loop is set to be $L = 50$. $\boldsymbol{\theta}^l$s are taken from the posterior samples obtained in the estimation step at intervals of 13,000 periods. Because the maximum of the inefficiency factors is 12,880, I can treat these $\boldsymbol{\theta}^l$s as independent samples from posterior distributions. The inner loop is set to have $R = 5,000$ posterior sample generations of q_h^{new} and p_h^{new} for $h = 1, ..., 32$ after discarding $R' = 5,000$ initial samples as the burn-in period.

For each of $\boldsymbol{\theta}^l$, $l = 1, ..., 50$, posterior sample paths of the predicted values exhibit a sufficient convergence of the inner loops.

When H_m is extremely large, as in Tokyo, which has 281 homes, it is computationally burdensome to achieve the convergence for the prediction procedure in Section 4.4. Instead, I concentrate on Shizuoka prefecture, which has 32 homes. In the prediction analysis below, I do not compare the lifetime payments for homes that currently do not collect the initial payment, although their information is used for prediction. Then, my target is the remaining 19 homes.

Figure 1 is here

Figure 1 presents a prediction result for the short-Lived consumers, $\tau_i \leq T_h$. The X axis indexes homes, while the Y axis measures a payment in 10,000 yen. Each home has two bars of monthly payments for before and after the intervention. Figure 1 shows that lifetime payments after the intervention are smaller than before the intervention. In other words, short-lived consumers can reduce their lifetime payment without the initial payment mechanism. Under this circumstance, longevity risks are pooled and distributed uniformly to all the residents. Therefore, short-lived consumers cannot recollect the risk premium and forced overpayment.

Figures 2 and 3 are here

Next, I consider long-lived consumers, $T_h < \tau_i$. Figures 2 and 3 gives comparisons of $p_h\tau + f_hT_h$ and $p_h^{new}\tau$ for $\tau = 240$ and 360, which correspond to 20 and 30 years of remaining lifetime, respectively. The two bars show the lifetime payments before and after the intervention.

The lifetime payments after the intervention exceed the payments before the intervention only in the case where a consumer with 30 years of remaining lifetime chooses a specific home. In other words, to recollect the risk premium, consumers need to stay at a home at least for 30 years. However, in practice, 30 additional years of life are not realistic for entrants into the private nursing homes with long-term care.

In addition, the expenses without the initial payment can be further reduced if we consider an interest rate.

Combining the above results, I conclude that the initial payment mechanism generally forces consumers to pay more. This overpayment might be a result of a consumer's rational risk management. However, there must be a loss of the consumer welfare in the aggregate level, because the overpayment is common for most consumers. Consumers may be better off with a combined policy of abandonment of the initial payment mechanism and a government-driven management of the longevity risk, similar to the safety net mechanism in the United States.

6 Conclusion

This paper has proposed a nonparametric Bayesian approach for structural econometrics. This approach enables a flexible prediction analysis without a distributional assumption. Although I have adopted the model of Berry et al. (1995) in this paper, my framework can work for general structural models. The validity of my method is shown in an empirical study of the Japanese private nursing home market. My prediction result implies that an outdated circumstance forces higher payments for most consumers today.

For empirical researchers, the elder care industry in Japan is an attractive field because the radical long-term care insurance program has rich implications for other aging countries. Although this paper concentrates on a specific market of private nursing homes, the long-term care industry has various sectors due to the market-oriented insurance program. As a future task, more studies from different perspectives in this economy are required.

A Polya tree mixture

This appendix complements Section 4.2 by introducing the Polya tree mixture in a manner similar to the intuitive definition of Christensen et al. (2008). To begin with, I define a Polya tree, which is an original form of the Polya tree mixture. First, I define a prior distribution for the Polya tree. Econometricians need to specify a base measure G on the support Ω, which has a well-defined density g and the known median μ.

The prior distribution of the Polya tree is constructed using a J step iterative process. In the first level, the support Ω is separated into two parts: R_{11} and R_{12}, which are below and above μ, respectively. With respect to the base measure G, both of these regions originally have probabilities of 1/2 because μ is the median of the base measure. We change these probabilities to λ_{11} and λ_{12} such that $\lambda_{11} + \lambda_{12} = 1$. In this manipulation, the shape of G is kept unchanged, but the integration constants in these regions are changed. Using the analogy of the histogram, the regions on Rs are called bins.

The second level creates a binary separation for each of R_{11} and R_{12} at the 25 and 75 percentiles, respectively. Then, the probabilities are changed in the same manner as in the first level. For example, on R_{11}, the new bins R_{21} and R_{22} have probabilities λ_{21} and λ_{22} such that $\lambda_{21} + \lambda_{22} = \lambda_{11}$. Such binary separations are repeated until the terminal level J.

As a result, one has a histogram-like prior distribution. The parameters of this prior distribution are the probabilities of bins λ_{j,κ_j} for $j = 1, ..., J$ and $\kappa_j = 1, 2, ..., 2^j$. The probabilistic structure of these variables can be represented simply using an additional latent variable as follows. From the above construction via binary separations, each level creates new probabilities by splitting them from the previous level. Let the latent variable $\zeta_{j,\kappa_{j-1}} \in [0, 1]$ be a proportion of the probability of the previous level, λ_{j-1,κ_j}, which is distributed to a new bin $R_{j,2\kappa_{j-1}-1}$. Then, we have a representation for the new probability as

$$\lambda_{j,2\kappa_{j-1}-1} = \zeta_{j,\kappa_{j-1}} \lambda_{j-1,\kappa_{j-1}}, \tag{A.1}$$

$$\lambda_{j,2\kappa_{j-1}} = (1 - \zeta_{j,\kappa_{j-1}}) \lambda_{j-1,\kappa_{j-1}}. \tag{A.2}$$

Due to conjugacy, it is convenient to impose an independent and identical Beta prior distribution for $\zeta_{j,\kappa_{j-1}}$. The prior and posterior distributions are written as follows:

$$\zeta_{j,\kappa_{j-1}} \sim Beta(\alpha_{j,2\kappa_{j-1}-1}, \alpha_{j,2\kappa_{j-1}}), \tag{A.3}$$

$$\zeta_{j,\kappa_{j-1}}|\omega_1, ..., \omega_{i-1} \sim Beta(\alpha_{j,2\kappa_{j-1}-1} + n_{j,2\kappa_{j-1}-1}, \alpha_{j,2\kappa_{j-1}} + n_{j,2\kappa_{j-1}}), \tag{A.4}$$

where $\alpha_{\cdot,\cdot}$s are hyperparameters, and $n_{j,k} = \sum_{h=1}^{i-1} I[\omega_h \in R_{j,k}]$ denotes the sample frequency.

Our purpose is the construction of the nonparametric predictive density of $\omega_i|\omega_1, ..., \omega_{i-1}, \tilde{\boldsymbol{\theta}}$. For this purpose, I integrate out the nuisance parameters $\boldsymbol{\lambda} = \lambda_{11}, \lambda_{12}, ..., \lambda_{J,2^J}$ and obtain

$$f(\omega_i|\omega_1, ..., \omega_{i-1}, \tilde{\boldsymbol{\theta}}) = \prod_{j=1}^{J} \frac{\alpha_{j,k_j} + n_{j,k_j}}{\alpha_{j,2\kappa_{j-1}-1} + \alpha_{j,2\kappa_{j-1}} + n_{j-1,\kappa_{j-1}}} I[w_i \in R_{j,k_j}] g(\omega_i), \tag{A.5}$$

where k_j is $2\kappa_{j-1} - 1$ or $2\kappa_{j-1}$.

The original Polya tree that is defined above has a similar weakness as a histogram, namely discontinuity at the borders of bins. This problem is caused by the fact that borders are defined as percentiles of the unique and fixed base measure. To overcome this discontinuity problem, the Polya tree mixture employs smoothing of the borders by introducing a variable base measure, denoted by G_τ, where τ is a scale parameter. This scale parameter is also estimated and integrated out in the definition of the nonparametric predictive density such that

$$\pi(\omega_i|\omega_1, ..., \omega_{i-1}, \tilde{\boldsymbol{\theta}}) = \int \pi(\omega_i|\omega_1, ..., \omega_{i-1}, \tilde{\boldsymbol{\theta}}, \tau)\pi(\tau|\tilde{\boldsymbol{\theta}})d\tau. \tag{A.6}$$

To achieve the median constraint, I assume that μ does not depend on τ. Because τ determines the percentile of G_τ other than μ, this integration smooths the bins except for those that are defined in the first level. On the other hand, discontinuity at μ enforces a median restriction such that the marginal distribution for ω_i satisfies $Pr(\omega_i \leq \mu) = 1/2$.

B An explicit representation for the Jacobian matrix in the likelihood function

This appendix provides a supplement for the estimation methodology described in Section 4.3. Specifically, I present a closed form expression for the Jacobian matrix in the likelihood function (4.6). A straightforward calculation yields the following

$$\mathcal{J}_m = \begin{pmatrix} \partial\omega_{1_m}/\partial p_{1_m} & ... & \partial\omega_{1_m}/\partial p_{H_m} & \partial\omega_{1_m}/\partial q_{1_m} & ... & \partial\omega_{1_m}/\partial q_{H_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial\omega_{H_m}/\partial p_{1_m} & ... & \partial\omega_{H_m}/\partial p_{H_m} & \partial\omega_{H_m}/\partial q_{1_m} & ... & \partial\omega_{H_m}/\partial q_{H_m} \\ \partial\xi_{1_m}/\partial p_{1_m} & ... & \partial\xi_{1_m}/\partial p_{H_m} & \partial\xi_{1_m}/\partial q_{1_m} & ... & \partial\xi_{1_m}/\partial q_{H_m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \partial\xi_{H_m}/\partial p_{1_m} & ... & \partial\xi_{H_m}/\partial p_{H_m} & \partial\xi_{H_m}/\partial q_{1_m} & ... & \partial\xi_{H_m}/\partial q_{H_m} \end{pmatrix}, \tag{B.1}$$

in which

$$\frac{\partial \omega_{h_m}}{\partial p_{l_m}} = \begin{cases} \frac{1}{p_{h_m}+[\alpha(\exp(q_{h_m})-1)]^{-1}+f_{h_m}\Gamma(T_{h_m},f_{h_m};\boldsymbol{\gamma})}(\equiv Z_{h_m}) & \text{for} \quad l_m = h_m \\ 0 & \text{for} \quad l_m \neq h_m \end{cases}, \tag{B.2}$$

$$\frac{\partial \omega_{h_m}}{\partial q_{l_m}} = \begin{cases} \left(-\frac{\exp(q_{h_m})}{\alpha[\exp(q_{h_m})-1]^2}\right) Z_{h_m} & \text{for} \quad l_m = h_m \\ 0 & \text{for} \quad l_m \neq h_m \end{cases}, \tag{B.3}$$

$$\frac{\partial \xi_{h_m}}{\partial p_{l_m}} = \begin{cases} \alpha & \text{for} \quad l_m = h_m \\ 0 & \text{for} \quad l_m \neq h_m \end{cases}, \tag{B.4}$$

$$\frac{\partial \xi_{h_m}}{\partial q_{l_m}} = \begin{cases} 1 + \frac{\exp[q_{h_m}]}{1-\sum_{k_m=1}^{H_m}\exp[q_{k_m}]} & \text{for} \quad l_m = h_m \\ \frac{\exp[q_{l_m}]}{1-\sum_{k_m=1}^{H_m}\exp[q_{k_m}]} & \text{for} \quad l_m \neq h_m \end{cases}. \tag{B.5}$$

Using the formula for the determinant by parts and the fact $Z_{h_m} > 0$, which is guaranteed under the support condition (4.10), we have

$$|\det(\mathcal{J}_m)| = |\det(D_m)|\left(\prod_{k_m=1}^{H_m} Z_{k_m}\right), \tag{B.6}$$

where D_m is a matrix whose (i, j) element is defined as

$$d_{ij}^m = \begin{cases} 1 + \frac{\exp(q_{i_m})}{1-\sum_{k_m=1}^{H_m}\exp(q_{k_m})} + \frac{\exp(q_{i_m})}{[1-\exp(q_{i_m})]^2} & \text{for} \quad i = j \\ \frac{\exp(q_{i_m})}{1-\sum_{k_m=1}^{H_m}\exp(q_{k_m})} & \text{for} \quad i \neq j \end{cases}. \tag{B.7}$$

Consequently, D_m does not depend on parameters $\boldsymbol{\theta}$ but on a dependent variable $\boldsymbol{q}_m$. Thus, $|\det(D_m)|$ can be negligible for estimation, whereas it must be considered for a prediction analysis.

C Expression for conditional predictive densities

This appendix details the prediction technique described in Section 4.4. Specifically, I derive the conditional predictive distributions for $q_{h_m}^{new}|\boldsymbol{p}_m^{new}, \boldsymbol{q}_{(-h_m)}^{new}, f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \boldsymbol{\theta}, \text{Data}$ and $p_{h_m}^{new}|\boldsymbol{q}_m^{new}, \boldsymbol{p}_{(-h_m)}^{new}, f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \boldsymbol{\theta}, \text{Data}$ for $h_m = 1, 2, ..., H_m$.

First, in the similar manner to derivation of the likelihood function, I obtain the joint predictive density for the dependent variables using a change of variable as

$$
\begin{aligned}
&\pi(\boldsymbol{p}_m^{new}, \boldsymbol{q}_m^{new} | f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \boldsymbol{\theta}, \text{Data}) \\
&= \pi_{\omega,\xi}[\omega_{h_m}(\boldsymbol{p}_m^{new}, \boldsymbol{q}_m^{new}; f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \tilde{\boldsymbol{\theta}}), \xi_{h_m}(\boldsymbol{p}_m^{new}, \boldsymbol{q}_m^{new}; f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \tilde{\boldsymbol{\theta}}) | \text{Data}] \\
&\quad |\det(\mathcal{J}_m)| \qquad \text{(C.1)} \\
&\propto |D_m| \prod_{h_m=1}^{H_m} \Bigg[Z_{h_m}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta}) g_{\tau_\omega}[\omega_{h_m}^{new}(p_{h_m}^{new}, q_{h_m^{new}}; \boldsymbol{\theta})] g_{\tau_\xi}[\xi_{h_m}^{new}(p_{h_m}^{new}, \boldsymbol{q}_m^{new}; \boldsymbol{\theta})] \\
&\quad \prod_{j=1}^{J} \frac{cj^2 + n_{\boldsymbol{\epsilon}[j,\tau_\omega,\omega_{h_m}^{new}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\omega})}{2cj^2 + n_{\boldsymbol{\epsilon}[j-1,\tau_\omega,\omega_{h_m}^{new}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\omega})} \prod_{j=1}^{J} \frac{cj^2 + n_{\boldsymbol{\epsilon}[j,\tau_\xi,\xi_{h_m}^{new}(p_{h_m}^{new}, \boldsymbol{q}_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\xi})}{2cj^2 + n_{\boldsymbol{\epsilon}[j-1,\tau_\xi,\xi_{h_m}^{new}(p_{h_m}^{new}, \boldsymbol{q}_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\xi})} \Bigg], \qquad \text{(C.2)}
\end{aligned}
$$

where

$$
\begin{aligned}
\omega_{h_m}^{new}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta}) &= \ln\Big(p_{h_m}^{new} + \frac{1}{\alpha[\exp(q_{h_m}^{new}) - 1]}\Big) - \boldsymbol{w}_{h_m}\boldsymbol{\beta}_s, && \text{(C.3)} \\
\xi_{h_m}^{new}(p_{h_m}^{new}, \boldsymbol{q}_m^{new}; \boldsymbol{\theta}) &= q_{h_m}^{new} - \tilde{\boldsymbol{x}}'_{h_m}\tilde{\boldsymbol{\beta}}_d - \ln\Big[1 - \sum_{k_m}^{H_m} \exp(q_{k_m}^{new})\Big] + p_{h_m}^{new}\alpha, && \text{(C.4)} \\
Z_{h_m}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta}) &= \frac{1}{p_{h_m}^{new} + \{\alpha[\exp(q_{h_m}^{new}) - 1]\}^{-1}}, && \text{(C.5)}
\end{aligned}
$$

and the support conditions yield

$$
0 < 1 - \sum_{k_m=1}^{H_m} \exp(q_{k_m}^{new}), \quad 0 < p_{h_m}^{new} + \frac{1}{\alpha[\exp(q_{h_m}^{new}) - 1]}. \qquad \text{(C.6)}
$$

Given the above joint predictive densities, I can obtain the conditional predictive densities to implement an MCMC prediction sampler. The conditional distribution for $p_{h_m}^{new}$ is:

$$\pi(p_{h_m}^{new}|\boldsymbol{q}_m^{new}, \boldsymbol{p}_{(-h_m)}^{new}, f_{h_m}^{new} = 0, T_{h_m=0}^{new}, \boldsymbol{\theta}, \text{Data})$$
$$\propto \; Z_{h_m}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta}) g_{\tau_\omega}[\omega_{h_m}^{new}(p_{h_m}, q_{h_m}; \boldsymbol{\theta})] g_{\tau_\xi}[\xi_{h_m}^{new}(p_{h_m}, \boldsymbol{q}_m; \boldsymbol{\theta})]$$
$$\prod_{j=1}^{J} \frac{cj^2 + n_{\boldsymbol{\epsilon}[j,\tau_\omega,\omega_{h_m}^{new}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\omega})}{2cj^2 + n_{\boldsymbol{\epsilon}[j-1,\tau_\omega,\omega_{h_m}^{new}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\omega})} \prod_{j=1}^{J} \frac{cj^2 + n_{\boldsymbol{\epsilon}[j,\tau_\xi,\xi_{h_m}^{new}(p_{h_m}^{new}, \boldsymbol{q}_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\xi})}{2cj^2 + n_{\boldsymbol{\epsilon}[j-1,\tau_\xi,\xi_{h_m}^{new}(p_{h_m}^{new}, \boldsymbol{q}_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\xi})}, \tag{C.7}$$

where

$$p_{h_m}^{new} > \frac{1}{\alpha[1 - \exp(q_{h_m}^{new})]}. \tag{C.8}$$

On the other hand, the conditional predictive density for $q_{h_m}^{new}$ is

$$\pi(q_{h_m}^{new}|\boldsymbol{p}_m^{new}, \boldsymbol{q}_{(-h_m)}^{new}, f_{h_m}^{new} = 0, T_{h_m=0}^{new}, \boldsymbol{\theta}, \text{Data})$$
$$\propto \; \prod_{m=1}^{M} |D_m| Z_{h_m}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta}) g_{\tau_\omega}[\omega_{h_m}^{new}(p_{h_m}, q_{h_m}; \boldsymbol{\theta})]$$
$$\prod_{j=1}^{J} \frac{cj^2 + n_{\boldsymbol{\epsilon}[j,\tau_\omega,\omega_{h_m}^{new}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\omega})}{2cj^2 + n_{\boldsymbol{\epsilon}[j-1,\tau_\omega,\omega_{h_m}^{new}(p_{h_m}^{new}, q_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\omega})}$$
$$\prod_{h_m=1}^{H_m} \left[g_{\tau_\xi}[\xi_{h_m}^{new}(p_{h_m}, \boldsymbol{q}_m; \boldsymbol{\theta})] \prod_{j=1}^{J} \frac{cj^2 + n_{\boldsymbol{\epsilon}[j,\tau_\xi,\xi_{h_m}^{new}(p_{h_m}^{new}, \boldsymbol{q}_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\xi})}{2cj^2 + n_{\boldsymbol{\epsilon}[j-1,\tau_\xi,\xi_{h_m}^{new}(p_{h_m}^{new}, \boldsymbol{q}_{h_m}^{new}; \boldsymbol{\theta})]}(\boldsymbol{\xi})} \right],$$
$$, \tag{C.9}$$

where

$$q_{h_m}^{new} < \log[1 - \sum_{k_m \neq h_m} \exp(q_{k_m}^{new})], \quad q_{h_m}^{new} < \log\Big[1 - \frac{1}{\alpha p_{h_m}^{new}}\Big]. \tag{C.10}$$

D Estimated posterior densities and sample paths

Figures 4, 5, 6 and 7 are here

References

Berry, S., J. Levinsohn, and A. Pakes (1995) "Automobile price in market equilibrium," *Econometrica*, Vol. 63, No. 4, pp. 841-890.

——— (2004) "Differentiated products demand systems from a combination of micro and macro data: The new vehicle market," *Journal of Political Economy*, Vol. 112, pp. 68-105.

Berry, S. and A. Pakes (2007) "The pure characteristics demand model," *International Economic Review*, Vol. 48, No. 4, pp. 1193-1225.

Bryne, D., M. S. Goeree, B. Hiedemann, and S. Stern (2009) "Formal home health care, informal care, and family decision making," *International Economic Review*, Vol. 50, No. 4, pp. 1205-1242.

Campbell, J. C., N. Ikegami, and M. J. Gibson (2010) "Lessons from public long-term care insurance in Germany and Japan," *Health Affairs*, Vol. 29, No. 1, pp. 87-95.

Chintagunta, P. and G. K. Y. Dubé, J.-P. (2005) "Beyond the endogeneity bias: The effect of unmeasured brand characteristics on household-level brand choice models," *Management Science*, Vol. 51, No. 5, pp. 832-849.

Christensen, R., T. Hanson, and A. Jara (2008) "Parametric nonparametric statistics: An introduction to mixture of finite Polya trees," *The American Statistician*, Vol. 62, No. 4, pp. 296-306.

Davis, P. (2006) "Spatial competition in retail markets: Movie theaters," *RAND Journal of Economics*, Vol. 37, No. 4, pp. 964-982.

Doss, H. (1985) "Bayesian nonparametric estimation of the median," *Annals of Statistics*, Vol. 13, pp. 1445-1464.

Engers, M. and S. Stern (2002) "Long-term care and family bargaining," *International Economic Review*, Vol. 43, No. 1, pp. 73-114.

Ferguson, T. S. (1973) "A Bayesian Analysis of Some Nonparametric Problems," *Annals of Statistics*, Vol. 1, pp. 209-230.

——— (1974) "Prior distributions on spaces of probability measures," *Annals of Statistics*, Vol. 2, pp. 615-629.

Finkelstein, A. and K. McGarry (2006) "Multiple Dimensions of Private Information: Evidence from the long-term care insurance market," *American Economic Review*, Vol. 96, pp. 938-58.

Gowrisankaran, G. and M. Rysman (2012) "Dynamics of consumer demand of new durable goods." Working paper.

Hanson, T., J. V. O. Monteiro, and A. Jara (2011) "The Polya tree sampler: toward efficient and automatic independent Metropolis-Hastings proposals," *Journal of Computational and Graphical Statistics*, Vol. 20, No. 1, pp. 41-62.

Hanson, T. and W. O. Johnson (2002) "Modeling regression error with a mixture of Polya trees," *Journal of the American Statistical Association*, Vol. 97, pp. 1020-1033.

Hanson, T. (2006) "Inference for mixtures of finite Polya tree models," *Journal of the American Statistical Association*, Vol. 101, pp. 1548-1565.

Hirano, K. (2002) "Semiparametric Bayesian Inference in Autoregressive Panel Data Model," *Econometrica*, Vol. 70, No. 2, pp. 781-799.

Hjort, N. L., C. Holmes, P. Müller, and S. G. Walker eds. (2010) *Bayesian Nonparametrics*, New York: Cambridge University Press.

Ikegami, N. and J. Campbell (2000) "Long-term care insurance comes to Japan," *Health Affairs*, Vol. 19, No. 3, pp. 26-39.

Japan Ministry of Health, Labor and Welfare (1999) "Survey on Social Welfare Institutions (Shakai fukushi shisetsu-tou chousa, in Japanese)," Tokyo.

——— (2010) "Basic Survey on Wage Structure," Tokyo.

——— (Various years) "Survey on institutions and establishments for long-term care (Kaigo sabisu sisetsu jigyousho chousa, in Japanese)," Tokyo.

Japan Statistical Bureau (2010) "Annual report on the retail price survey," Tokyo.

Jara, A., T. Hanson, and E. Lesaffre (2009) "Robustifying generalized linear mixed models using a new class of mixtures of multivariate Polya trees," *Journal of Computational and Graphical Statistics*, Vol. 18, No. 4, pp. 838-860.

Jiang, R., P. Manchanda, and P. E. Rossi (2009) "Bayesian analysis of random coefficient logit models using aggregate data," *Journal of Econometrics*, Vol. 149, No. 2, pp. 136-148.

Kottas, A. and A. E. Gelfand (2001) "Bayesian semiparametric median regression modeling," *Journal of the American Statistical Association*, Vol. 96, No. 456, pp. 1458-1468.

Lavine, M. (1992) "Some aspect of Polya tree distributions for statistical modeling," *Annals of Statistics*, Vol. 20, No. 3, pp. 1222-1235.

——— (1994) "More aspect of Polya tree distributions for statistical modeling," *Annals of Statistics*, Vol. 22, No. 3, pp. 1161-1176.

Lucas, R. E. (1976) "Econometric policy evaluation: A critique," in Brunner, K. and A. Meltzer eds. *Theory, policy, institutions: Papers from the Carnegie-Rochester conference series on public policy*: Elsevier.

Mauldin, R. D., W. D. Sudderth, and S. C. Williams (1992) "Polya trees and random distributions," *Annals of Statistics*, Vol. 20, No. 3, pp. 1203-1221.

Mehta, A. (2006) "Spatial competition and market definition in the nursing home industry." Working paper.

Mitchel, O. S., J. Piggott, and S. Shimizutani (2004) "Aged-care support in Japan: Perspectives and challenges." NBER Working Paper 10882.

Momose, T. (1997) *History of the welfare for elders in Japan(in Japanese)*, Tokyo, Japan: Chuuou Houki.

Musalem, A., E. Bradlow, and J. S. Raju (2009) "Bayesian estimation of random-coefficient choice models using aggreagate data," *Journal of Applied Econometrics*, Vol. 24, No. 3, pp. 490-516.

Nevo, A. (2000) "Mergers with differentiated products: The case of the ready-to-eat cereal industry," *Rand Journal of Economics*, Vol. 31, pp. 395-421.

Nonaka, T., H. Noguchi, A. Nishi, M. R. Reich, N. Ikegami, H. Hashimoto, K. Shibuya, I. Kawauchi, and J. C. Campbell (2011) "Population ageing and well-being: lessons from Japan's long-term care insurance policy," *Lancet*, Vol. 378, pp. 1183-1192.

Norton, E. (2000) "Long-term care," in Newhouse, J. and A. Culyer eds. *Handbook of health economics*, Vol. 1B, Amsterdam: North-Holland, pp. 95-994.

Petrin, A. (2002) "Quantifying the benefits of new products: The case of the minivan," *Journal of Political Economy*, Vol. 110, pp. 705-729.

Shuukan Asahi Mook (2011) *Elderly home handbook 2011: With information of initial payment and care for Alzheimer's disease(in Japanese)*, Tokyo, Japan: Asahi Shinbunsha.

Sudou, Y. (2006a) "The history and future perspective of the business of private nursing homes 1(in Japanese)," *Aioi Kiso-ken Review*, Vol. 1, pp. 82-94.

——— (2006b) "The history and future perspective of the business of private nursing homes 2(in Japanese)," *Aioi Kiso-ken Review*, Vol. 2, pp. 52-71.

Tsutsui, T. and N. Muramatsu (2007) "Japan's universal long-term care system reform of 2005: Containing costs and realizing a vision," *Journal of the American Geriatrics Society*, Vol. 55, No. 9, pp. 1458-1463.

Yang, S., Y. Chen, and G. M. Allenby (2003) “Bayesian analysis of simultaneous demand and supply,” *Quantitative Marketing and Economics*, Vol. 1, No. 3, pp. 251-275.

E Tables and Figures

		1999	2000(LTCI)	2001	2002	2003
Private	# Homes	298	350	400	508	694
	Capacities	32,302	37,467	41,445	46,561	56,837
Public	# Homes	4,214	4,463	4,651	4,870	5,084
	Capacities	283,822	298,912	314,192	330,916	346,069
		2004	2005	2006	2007	2008
Private	# Homes	1,045	1,406	1,968	2,671	3,400
	Capacities	76,128	964,12	123,155	147,981	176,935
Public	# Homes	5,291	5,535	5,716	5,892	6,015
	Capacities	363,747	383,326	399,352	412,807	422,703

Table 1: Numbers of the institutions and the residential capacities of public and private nursing homes

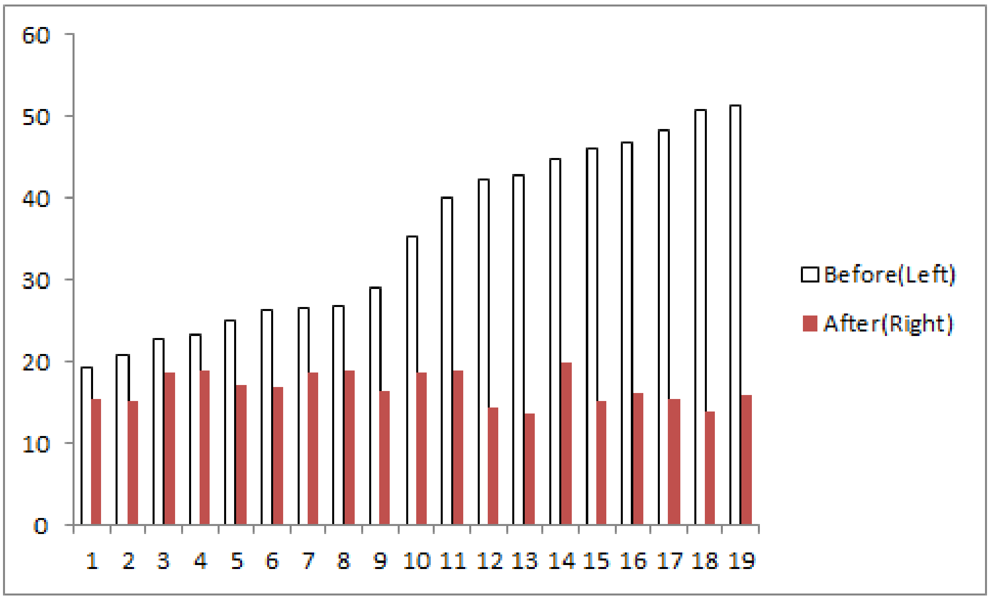

Figure 1: Prediction result for short-lived consumers ($\tau_i \leq T_h$):
X axis indexes homes and Y axis measures monthly fees in 10,000 yen

Prefecture	# Homes	# Category 1 elders	Prefecture	# Homes	# Category 1 elders
Hokkaido	38	234,434	Aichi	70	214,087
Iwate	2	62,053	Mie	3	78,731
Miyagi	15	84,786	Shiga	4	45,764
Akita	2	61,281	Kyoto	7	108,892
Yamagata	3	55,587	Osaka	120	358,001
Fukushima	7	84,428	Hyogo	61	223,140
Ibaraki	16	90,099	Nara	9	53,548
Tochigi	6	64,671	Shimane	3	40,650
Gumma	12	75,409	Okayama	28	93,412
Saitama	76	189,482	Hiroshima	17	128,505
Chiba	74	174,744	Yamaguchi	5	71,385
Tokyo	281	423,639	Kagawa	6	46,256
Kanagawa	232	264,673	Ehime	12	74,667
Niigata	15	109,182	Fukuoka	52	203,339
Ishikawa	4	48,238	Saga	7	37,445
Yamanashi	3	31,571	Nagasaki	4	78,863
Nagano	13	92,933	Kumamoto	3	86,886
Gifu	7	75,766	Oita	9	60,433
Shizuoka	32	128,088	Kagoshima	7	87,718

Table 2: Numbers of homes and elders in prefectures, excluding prefuctures with zero or one home

Variable	Notation in paper	Mean	S.D.
Monthly fee (10,000yen)	p	19.90	6.28
Initial payment per month (10,000yen)	f	10.90	13.96
Expiration period(month)	T	46.34	42.17
Share	s	0.00032	0.00045
Log(Share)	q	-8.418	0.798
# Residents per worker	**Worker**	2.523	0.476
Years from opening	**Years**	7.378	5.986
Occupancy rate	**Occupancy**	0.914	0.146
Chain dummy	**Chain**	0.315	0.465
Local average rent	**Rent**	6051	1813
Local average wage (1,000 yen)	**Wage**	1025	120
Sample size	H	1265	

Table 3: Descriptive statistics

Variable		Mean	S.D.	95% Interval	IF
$\boldsymbol{\beta}_d$	Constant	-4.298***	0.227	[-4.728,-3.837]	607
	# Residents per worker	0.678***	0.054	[-0.787,-0.579]	250
	Years from opening	0.017	0.029	[-0.042,0.071]	35
	Occupancy rate	0.440**	0.161	[0.114,0.752]	411
	Chain dummy	0.378***	0.052	[0.278,0.484]	40
α_F		0.051***	0.020	[0.012,0.089]	92
$\boldsymbol{\beta}_s$	Constant	2.928***	0.058	[2.823,3.052]	12924
	# Residents per worker	-0.415***	0.022	[-0.462,-0.369]	12288
	Years from opening	-0.095***	0.016	[-0.124,-0.068]	3159
	Chain dummy	0.217***	0.022	[0.163,0.255]	3603
	Local average rent	0.189***	0.014	[0.165,0.223]	811
	Local average wage	0.009	0.014	[-0.044,0.012]	706
α		0.139***	0.001	[0.137,0.140]	477
$\boldsymbol{\gamma}$	γ_0	0.283	2.982	[-5.884,5.741]	16
	γ_{T1}	-2.472	2.235	[-7.392,1.309]	22
	γ_{T2}	-2.191	2.328	[-7.286,1.568]	13
	γ_{f1}	-0.757	2.964	[-6.682,4.853]	24
	γ_{f2}	-2.488	1.993	[-7.159,0.171]	10
	γ_{fT}	-0.509	3.130	[-6.672,5.597]	6
τ	τ_ω	2.899***	0.360	[2.258,3.404]	3043
	τ_ξ	0.919***	0.080	[0.769,1.081]	45
	Sample size				1265

Table 4: Estimation result for real data

$***$, $**$ and $*$ indicate that 99%, 95% and 90% credible intervals do not include zero, respectively.

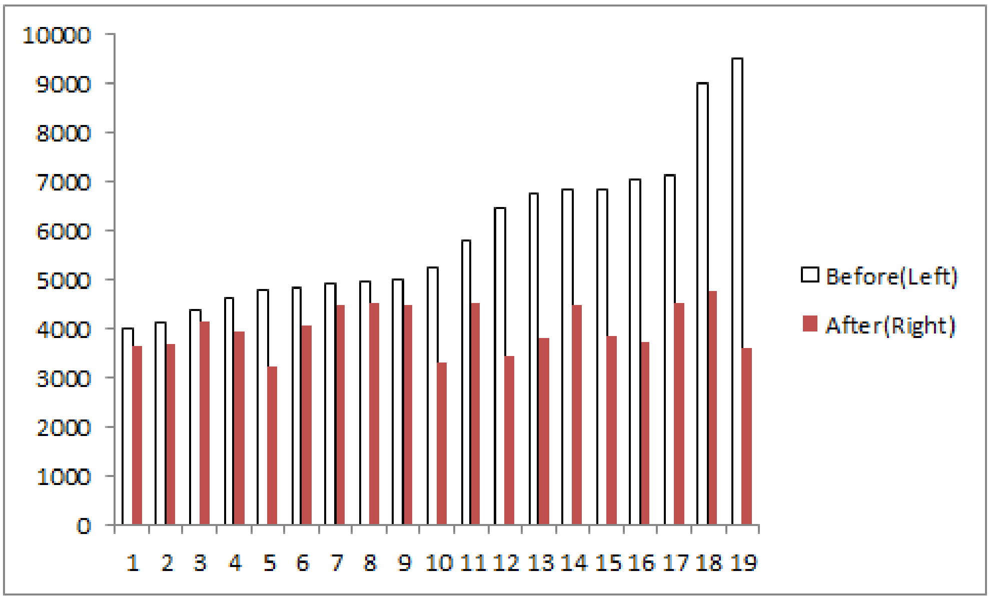

Figure 2: Prediction results for long-lived consumers: 20 year lifetime ($\tau_i > T_h$): $\tau_i = 240$

X axis indexes homes and Y axis measures lifetime payments in 10,000 yen

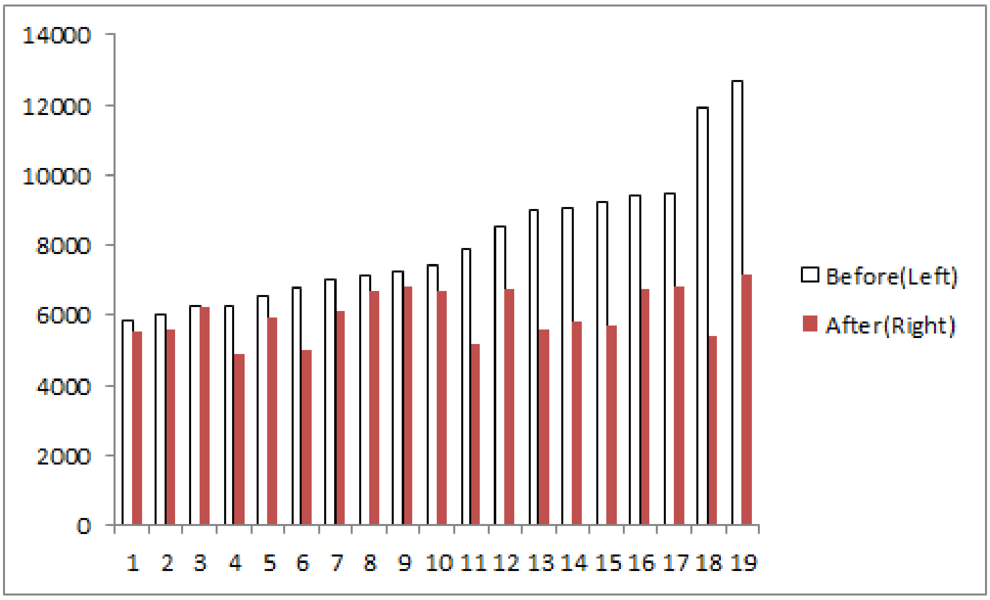

Figure 3: Prediction results for long-lived consumers: 30 year lifetime ($\tau_i > T_h$): $\tau_i = 360$

X axis indexes homes and Y axis measures lifetime payments in 10,000 yen

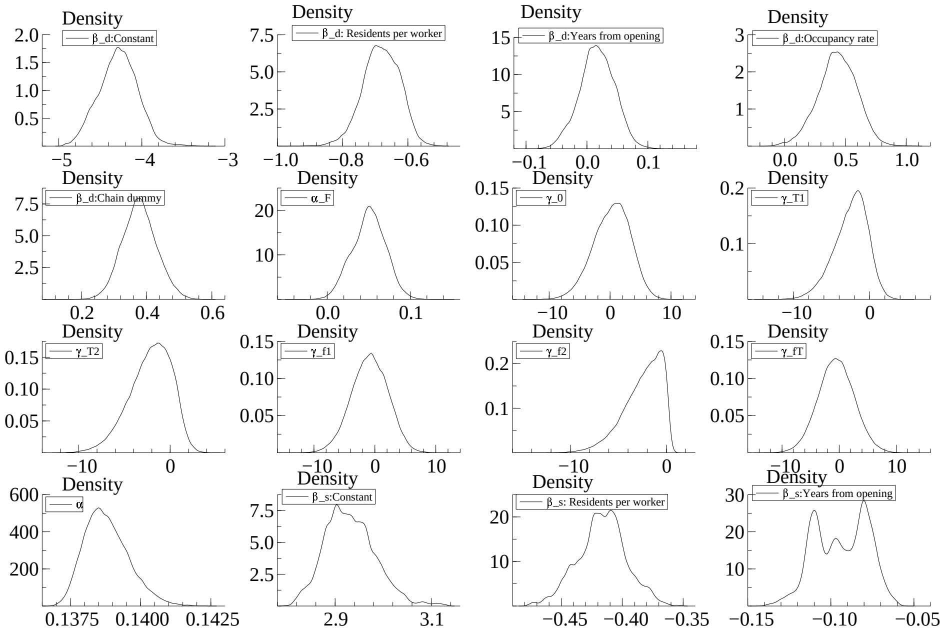

Figure 4: Posterior densities for the MCMC sampler, 1

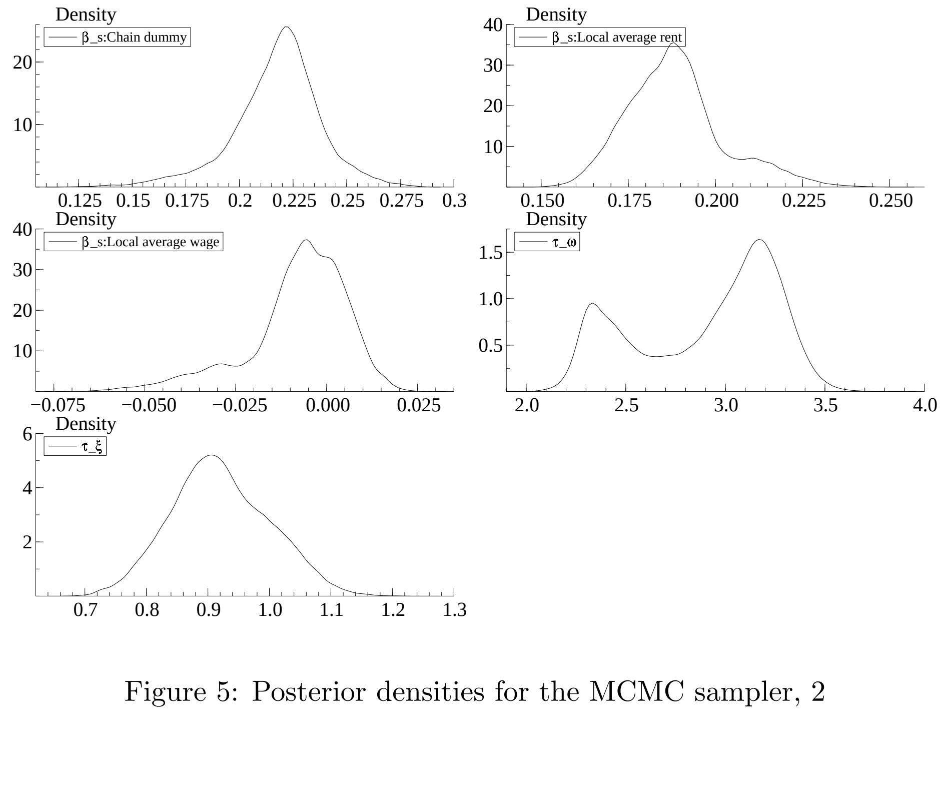

Figure 5: Posterior densities for the MCMC sampler, 2

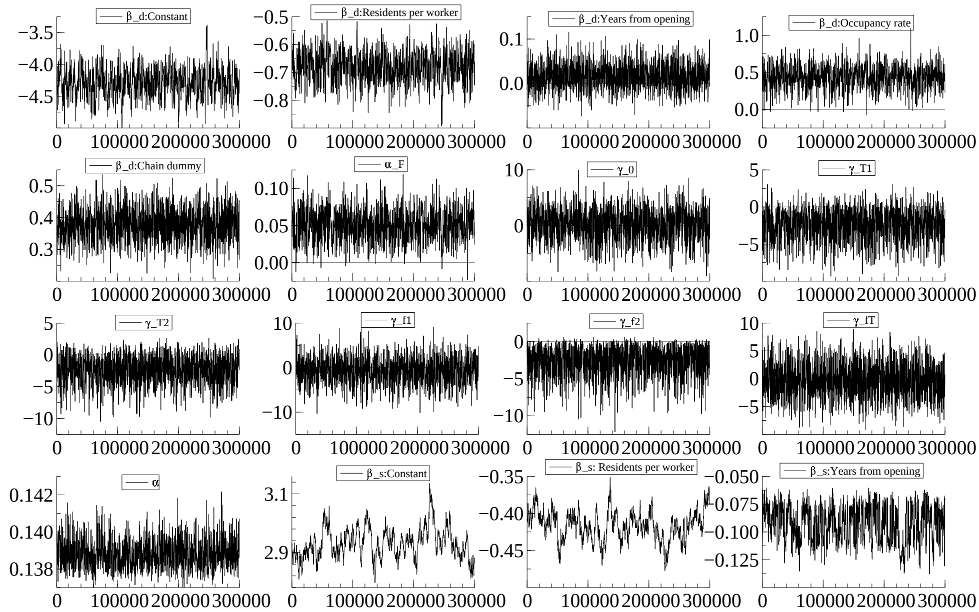

Figure 6: Sample paths for the MCMC sampler, 1

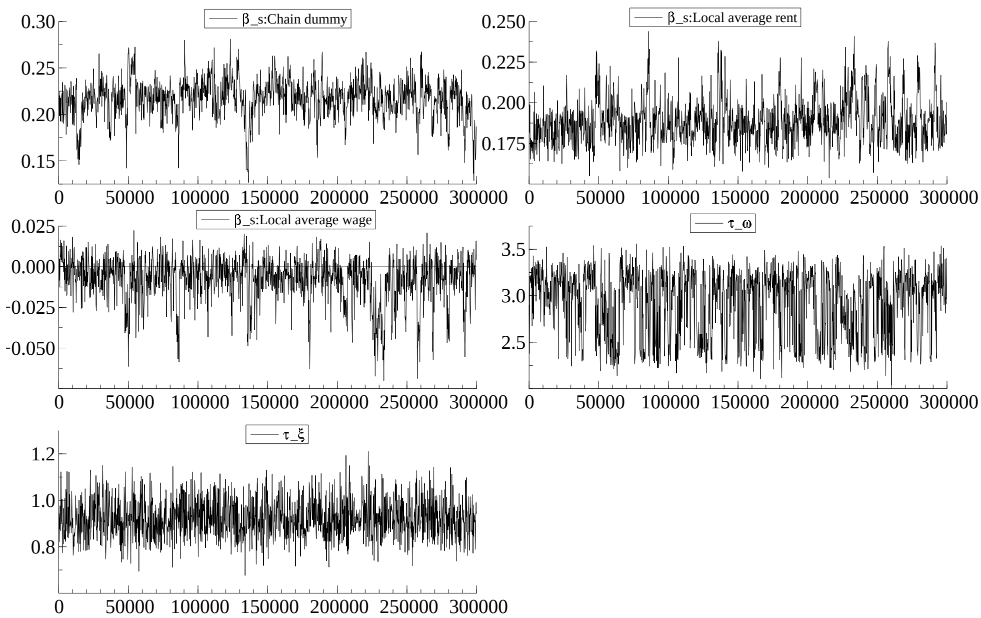

Figure 7: Sample paths for the MCMC sampler, 2