CIRJE-F-908

Firm-driven Management for Longevity Risk: Analysis of Lump-sum Forward Payments in Japanese Nursing Homes

Shinya Sugawara
The University of Tokyo

November 2013

CIRJE Discussion Papers can be downloaded without charge from:
http://www.cirje.e.u-tokyo.ac.jp/research/03research02dp.html

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.

Firm-driven management for longevity risk: Analysis of lump-sum forward payments in Japanese nursing homes

Shinya Sugawara*

This paper analyzes a system of insuring longevity risk in Japan using a lump-sum forward payment of rent. The system is unique to the Japanese private nursing home market and forces homes to cover most of the longevity risk of residents. To consider the impact of this system, this study conducts a prediction analysis of a counterfactual situation in which the system is eliminated, based on a structural model for the industrial organization of the private nursing home market. For the prediction analysis, a flexible technique is proposed using a nonparametric Bayesian approach. The results indicate there is an excess consumer payment under the current system, for which the consumer can only be compensated if he or she lives for an unrealistically long time in a nursing home. This result provides rich implications for the amount of longevity risk nursing homes assume, as well as possible consumer welfare to be gained from a government intervention to change the situation.
Keywords: Nursing home; Longevity risk; Risk premium; Long-term care in Japan; Nonparametric Bayes; Industrial organization

1. Introduction

In spite of a growing demand for longevity risk insurance caused by aging populations worldwide, this remains one of the largest uninsured risks. Many researchers have tried to explain the small market size of insurance-like products, including Finkelstein and Poterba (2004) for annuities, Brown and Finkelstein (2007) for private long-term care insurance, and Davidoff and Welke (2006) for reverse mortgages. According to Cutler (1996), a possible reason for the small size of the longevity risk insurance market might be the difficulty of evaluating the risk, as a result of continuous improvements in medical technologies. In fact, Mitchell et al. (1999) showed that there is a large difference between the premiums and the expected return from annuities.

*University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033, Japan and Nihon University 12-5 Gobancho, Chiyoda, Tokyo 102-8251, Japan, email: sugawara@e.u-tokyo.ac.jp, Tel.: +81-3-5841-5539

This study is concerned with a system of insuring longevity risk currently employed in the Japanese private nursing home market, in which residents pay a lump-sum forward payment of rent. The system is unique to this market, and forces nursing homes to assume most of longevity risk associated with their residents. The mechanism is apparently a relief for risk-averse elderly who have a limited income. However, a rational home must increase the other forms of prices to manage the burden of the longevity risk. This study considers the impact of this firm-driven management of longevity risk.

The system was established when access to the private nursing home market was limited to wealthy consumers. This situation has since changed with the launch of the social insurance program for elderly care in Japan, and private homes today appear to be more accessible to ordinary consumers. Nevertheless, many homes still adopt this price mechanism.

By way of comparison, as summarized in Norton (2000), private nursing homes in the United States have a different system to those in Japan. In the American system, rents are primarily paid by consumers. Once consumers have spent their assets and can no longer afford the cost of a nursing home, the government covers the rent as a benefit of Medicaid, the public insurance for low-income earners. Through this mechanism, which is called the Medicaid spend-down, the government provides a safety net for consumers to manage their longevity risk. To compare the systems in Japan and the United States, we first need to analyze the implications of the Japanese lump-sum forward payment system.

To consider the impact of the Japanese system, this study conducts a prediction analysis for a system in which consumers do not pay the lump-sum forward payment mechanism. A comparison of the systems provides two implications. First, the difference in the payments can measure the impact of a possible intervention to remove the sticky circumstance in terms of consumer welfare. Second, the difference can indicate the longevity risk premium that homes assume. These two findings can then be used to suggest policies to establish a sustainable mechanism for elderly care.

The prediction analysis is based on a structural econometric model for the industrial organization of the private nursing home market. Owing to considerable sunk costs and barriers to entry resulting from regulation policies, homes are likely to operate in incomplete competition. To model this situation, the approach of the Berry et al. (1995) was adopted, hereafter referred to as BLP. The BLP model simultaneously formulates the demand and supply sides of an economy in which firms play a Bertrand competition on the supply side. Owing to its tractability, the BLP model has been widely applied to various industrial markets, including an automobile retail market by BLP themselves, a service sector by Davis (2006), and a durable goods market by Gowrisankaran and Rysman (2012), among others. In this study, the original model is extended from a static BLP model to a dynamic model to capture the nature of the nursing home market.

The BLP model is conventionally analyzed using the generalized method of moments (GMM). However, the moment conditions for the GMM are not always sufficient to conduct a prediction analysis. Therefore, this study proposes a flexible prediction technique using a nonparametric Bayesian approach. This methodology allows for an estimation without distributional assumptions and a prediction analysis. The approach used here

follows that of previous Bayesian studies, such as Yang et al. (2003), Musalem et al. (2009), and Jiang et al. (2009). These previous studies adopt parametric assumptions to implement an efficient estimation algorithm via a Gibbs sampler. The methodology in this study has the advantage of flexibility, owing to the nonparametric modeling, but has a disadvantage in terms of its computational burden.

An empirical analysis is employed using real data taken from the list of nursing homes in Shuukan Asahi Mook (2011). The estimation results produce reasonable estimates for the model parameters. Furthermore, the prediction analysis shows that the outdated system induces a large increase in the lifetime total payments for residents, except for those who have unrealistically long lifetimes. This result may be due to risk-averseness on the part of both consumers and the nursing homes. Therefore, government help to hedge the longevity risk can increase the consumer surplus in the nursing home market.

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 the econometric model used in this study, and Section 4 presents the corresponding econometric methodology. The proposed method is applied to real data in Section 5. Finally, Section 6 concludes the paper.

2. The Japanese private nursing home market

2.1. The nursing home industry

To manage the world's most rapidly aging population, the Japanese government launched a national program, Long-Term Care Insurance (LTCI), in 2000. One of the main policy objectives of the LTCI is the "Socialization of Care," which requests the release of females from informal family care. To achieve this purpose, the LTCI does not provide per-capita cash transfers to the elderly, something that is allowed in a program in Germany, but covers only purchased service costs. To satisfy the considerable demand induced by this policy, the formal elderly care market has quickly grown into a large industry.

The LTCI is a social insurance system with universal coverage. The system compensates 90% of purchased service costs, the prices of which are prescribed by a detailed remuneration points system. Since there are few regional variations in the exchange rate between points and money, this has become an almost fixed rewards system for care services.

Among the wide variety of formal elderly care services, this study is concerned specifically with nursing homes. A nursing home is defined as an institution for permanent care. The LTCI covers both the public and private nursing home sectors.[1]

[1] The terms "public nursing home" and "private nursing home" are translations of the Japanese words "Tokubetsu-Yougo Roujin Houmu" and "Yuuryou Roujin Houmu," respectively. Tamiya et al. (2011) used the term "private nursing home" for "Yuuryou Roujin Houmu," but there is no consensus among researchers on this English term. For example, Ikegami and Campbell (2000) refers to a private nursing home as "residential care with private-pay." In this study, private homes are treated as a form of nursing home and juxtaposed with public homes for two reasons. First, their function matches that of the standard definition of a nursing home, namely to provide general long-term care

The two sectors are differentiated according to two features. First, LTCI coverage is different for public and private homes. For public homes, the LTCI covers everything except rent and food.[2] Public homes provide uniform care at uniform prices under the remuneration point system only for those who cannot leave a bed by themselves.

For private homes, the LTCI covers only those fees strictly categorized as care costs. Care costs are not included in an ordinary payment, but instead are treated as a person-specific additional payment. Then, private homes can provide divergent care at various prices, and accept a variety of residents. There are two general categories of private homes, according to resident eligibility. The first refers to homes for independent elderly people. These homes do not provide care services. In this category, when a resident requires permanent long-term care, he or she needs to exit the home. Private homes in the other category, called homes with care, provide care services as a default option, and a resident can stay in the home until his or her death. This study focuses on the latter category, private nursing homes with care.

The second distinction is that public homes are operated by a local authority or a non-profit organization. For-profit firms are not allowed to enter this market, but they can operate private homes[3]. Owing to the history of private homes, described below, most are owned by for-profit firms. Our data indicate that more than 90% of private homes are owned by for-profit firms.

The background to the co-existence of public and private homes is as follows.[4] The first Japanese nursing homes were established in the late 19th century by religious and philanthropic organizations as voluntary institutions for the poor and solitary elderly. A legal basis for 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 then formed to meet the demand of the free contractors. Although the exact origin of private nursing homes is blurry, there is a record of an active home from 1948. In 1963, the Act on the Social Welfare Service for the Elderly (Roujin Fukushi Hou) updated the legal system of the long-term care sector. This Act prohibited public homes from accepting free contract residents and prescribed legal requirements for private nursing homes for the first time. In other words, this Act explicitly separated public and private nursing homes. Until the 1990s, most residents in private homes

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 such as "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.

[2]These costs were once included in what was covered by the LTCI, but were eliminated by the amendment to the LTCI, as described in Tsutsui and Muramatsu (2007).

[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 institutional care only as public homes.

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

were wealthy elders, while residents in public homes were less well-off and solitary elders targeted by the welfare program. Several luxury private homes attracted wide attention in the Japanese "bubble" economy in the late 1980s.

Figure 1 is here

The impact of the LTCI has been different for the private and public home sectors. Private homes increased their total capacity from 32,302 in 1999 to 183,245 in 2009, while public homes showed slower growth from 283,822 in 1999 to 414,668 in 2009. Figure 1 shows the annual growth rates of the capacity of public and private homes.[5] 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 the result of a regulatory policy that, in response to the rapid increase of the financial burden of the LTCI, replaced costly institutional care with home care. To relax the considerable expansion of the fiscal budget, local authorities suppressed the establishment of new homes. This policy induced a long waiting list for those seeking elderly care, estimated to be 421,000 in 2009.[6] As a result of these exogenous restrictions on the supply of public homes, this study does not consider the crowding-out effects of public homes on private home demand.

The government sought to control the number of private homes as well. To provide a legal basis for this motivation, they announced the "Regulation of Volume" (Souryou Kisei)[7] in 2005. This regulation stated that, in 2014, the number of elderly in institutions, including public homes and private homes with care, must be less than 37% of the eligible number of elderly. This regulation became a new barrier to entry, because the rate had already reached 41% in 2004. After active debates on this regulation, it was abandoned in 2012. Figure 1 shows how the regulation caused a reduction in the growth rate of private home capacity after 2005. This reduction implies that the potential demand for nursing homes must be larger than the supply.

2.2. A price mechanism of private nursing homes

This section explains the price mechanism in the private nursing home market. In this market, there is a traditional contract that requires a resident to make two forms of payment. The first payment is a monthly fee that covers the costs of daily needs. The second payment is a once-off initial payment (*Nyuukyo kin*), which is a lump-sum forward

[5] These figures 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 for private homes for the entire period. There are more recent figures, but those after 2009 have a problem with consistency. Since the research agents changed from government officials to private firms, the response rate has dropped drastically.

[6] The figure is taken from a press release by the Japan Ministry of Health, Labor and Welfare. The web page is http://www.mhlw.go.jp/stf/houdou/2r98520000003byd.html and was accessed on October 26, 2013.

[7] As reviewed in Mehta (2006), there is a similar regulation in the US that specifies the number of beds in nursing homes, called the control of need (CON), which restricts the expansion of the Medicaid budget caused by the Medicaid spend-down.

payment of rent until an expiration period. The length of time until the expiration period is determined by the home. In addition to these two forms of payment, consumers need to pay the costs of long-term care, 90% of which are covered by the LTCI.

The lump-sum forward payment has the following unique property. If a resident exits the home before the expiration date, the rent for the remaining periods is paid back. However, if a resident remains in the home beyond the expiration date, he or she does not need to pay any additional rent.[8] Under this contract, homes assume most of the longevity risk associated with their residents.

This mechanism comes from the past, when access to the private home market was limited to wealthy consumers who paid an expensive lump-sum forward payment with a long expiration date. In those days, with such limited demand, the price mechanism could have been a Nash equilibrium strategy for homes. However, the nursing home industry has since experienced a significant transition, so that private homes are now more accessible to ordinary consumers as a substitute for public homes with their long waiting lists.

The lump-sum forward payment is not a formal rule, but rather an economic circumstance. Thus, it is legally possible to offer a contract in which rent is collected using a monthly fee rather than a lump-sum forward payment, thus avoiding the longevity risk. However, fewer than 35% of the homes in our dataset exercise this option. In addition, most homes offer only one expiration period, which implies that homes do not practice price discrimination in managing their longevity risk.

This payment mechanism is apparently a relief to risk-averse elderly without a source of income other than their public pension. However, rational homes must compensate for the longevity risk by using a different channel of the payment. In this case, they charge a higher monthly fee. Thus, the monthly fee under the lump-sum forward payment mechanism might reflect the longevity risk premium that homes cover. The amount of the total payment under this price mechanism might be much larger by risk-averse homes. Nevertheless, the mechanism exists, probably because of risk-averseness of consumers. This study considers the impacts of this outdated, but sticky system of firm-driven management of longevity risk.

3. A structural econometric model for the nursing home market

3.1. Defining a market environment

This section presents a structural econometric model for the Japanese private nursing home market as an extension of the BLP model. The basic setup of the economy is first

[8] A portion of the lump-sum forward payment is called an initial depreciation, which is not returned if a resident exits before the expiration date. The empirical analysis assumes that a home does not collect the lump-sum forward payment if the initial depreciation is 100% of the lump-sum forward payment. Otherwise, for the sake of simplicity, the initial depreciation is not reflected in the definition of the price variables.

described in this subsection, with the details of the model following in the subsequent subsections.

There are M local markets that are geographically isolated. For the mth market, the demand side consists of I_m consumers, and the supply side consists of H_m private nursing homes. The homes play incomplete competition of the Bertrand type owing to high sunk costs for the institution and the Regulation of Volume as barriers to entry. An important finding of Berry et al. (1995) is that all the parameters can be identified using only firms-side observations, making information on each consumer unnecessary. Therefore, the sample used here consists only of information for $H = \sum_m H_m$ nursing homes.

For the h_mth home, we can observe three variables that describe payment information: the monthly fee, p_{h_m}; the lump-sum forward payment, F_{h_m}; and the expiration period, T_{h_m} (month). Based on F_{h_m} and T_{h_m}, a new variable is created for monthly rent, f_{h_m}, which is defined as $f_{h_m} = F_{h_m}/T_{h_m}$ if $T_{h_m} \neq 0$, or $f_{h_m} = 0$ if $T_{h_m} = 0$. Vector notation is used here for notational simplicity, such as $\mathbf{p}_m = (p_{1_m}, ..., p_{H_m})'$ and $\mathbf{p} = (\mathbf{p}_1', ..., \mathbf{p}_M')'$, where each component is indexed as p_h for $h = 1, 2, ..., H$.

To consider the dynamic nature of the nursing home market, the original, static BLP model needs to be extended. For dynamic modeling, an important factor is the lifetime of consumers, denoted here as τ_{i_m} for the i_mth consumer. The lifetime is uncertain for both homes and consumers, and this uncertainty is carefully modeled in the following subsections.

3.2. Modeling the demand side

The model for the demand side represents the consumers' optimization problem. The consumers in the mth market decide to choose one home among H_m private homes, or an outside option. The outside option represents anything other than a private home, such as public homes, formal home care, or informal family care. From the dynamic nature of the nursing home as a good, consumers optimize the present value of their expected lifetime utility.

Two assumptions are used to simplify the demand-side modeling. The first assumption is that there is no voluntary exit from a home, and hence, any exit is the result of the resident's death. This avoids complications caused by endogenous exit decisions. To justify this assumption, the samples are restricted to private homes with care, and we do not consider homes for independent elderly.

The second assumption is that consumers evaluate their remaining lifetime using a subjective prediction. A consumer's lifetime is a time horizon in the present value calculation, and its uncertainty induces a difficulty in obtaining the expected lifetime utility. For mathematical simplicity, a consumer-specific, subjective prediction for the lifetime is introduced, but is integrated out in the formation of the econometric model. For the i_mth consumer, the subjective lifetime is denoted as $\tau^I_{i_m}$. Note that risk-averse consumers are likely to overestimate this value.

In addition, there is an assumption of an explicit functional form for the components of the consumers' lifetime utility. There are two components in the lifetime utility, namely

the utility gains for each period and the disutility from the lifetime payment. The utility gain for each period is assumed to take the form of a linear function of an observable $K_d \times 1$ vector, $\tilde{\mathbf{x}}_{h_m}$, an unobservable home-specific effect, ξ_{h_m}, and an unobservable individual home match-specific effect, $\tilde{\eta}_{i_m h_m}$. The disutility from the lifetime payment is defined as $-P(\cdot)\alpha$. The function $P(\cdot)$ measures the present value of the lifetime payment. The argument to the function is discussed in more detail later. The coefficient $-\alpha$ measures the disutility from a unit expenditure, and hence $\alpha > 0$ is required. An individual is allowed a specific time-discount factor, denoted as $\delta_{i_m} \in (0,1)$. Consequently, the present value of the lifetime utility of consumer i_m from a choice of $the h_m$th home is

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

The present value of the lifetime payment is defined as

$$P(p_{h_m}, f_{h_m}, T_{h_m}, \tau_{i_m}^I, \delta_{i_m}) = \sum_{t=1}^{\tau_{i_m}^I} \delta_{i_m}^{t-1} p_{h_m} + F_{h_m} - I[T_{h_m} \geq \tau_{i_m}^I] \delta_{i_m}^{\tau_{i_m}^I} (T_{h_m} - \tau_{i_m}^I) f_{h_m}, \quad (3.1)$$

where the right-hand side of (3.1) consists of three parts. The first term represents the present value of the monthly fees during the lifetime. The second term is the initial payment in a lump sum, which does not depend on δ_{i_m} because it is paid in full upon arriving at the home (i.e., $t = 1$). The third term corresponds to the return on the initial payment should a consumer die before the expiration date (i.e., $T_{h_m} \geq \tau_{i_m}^I$). This amount would be returned to residents after their death, but it is assumed to be contained in their utility. A dynasty model can justifies this assumption.

In Equation (3.1), I assume that one can decompose the terms F_{h_m} and f_{h_m} into a between-home mean and a home-and-individual-match variation, as follows:

$$\begin{aligned} & -F_{h_m} \frac{1-\delta_{i_m}}{1-\delta_{i_m}^{\tau_{i_m}^I}} \alpha + f_{h_m} \frac{1-\delta_{i_m}}{1-\delta_{i_m}^{\tau_{i_m}^I}} \delta_{i_m}^{\tau_{i_m}^I} I[T_{h_m} \geq \tau_{i_m}^I](T_{h_m} - \tau_{i_m}^I)\alpha \Bigg] \\ = \; & F_{h_m}\alpha_F + f_{h_m}\alpha_{f,i_m h_m} + \eta_{Ff,i_m h_m}. \end{aligned} \quad (3.2)$$

Here, $\mathbf{x}_{h_m} = (\tilde{\mathbf{x}}_{h_m}', F_{h_m}, f_{h_m})'$, $\beta_d = (\tilde{\beta}_d', \alpha_F, \alpha_{f_m})'$, and $\eta_{i_m h_m} = \tilde{\eta}_{i_m h_m} + \eta_{Ff,i_m h_m}$. For notational simplicity,

$$V_{i_m h_m} = [(1-\delta_{i_m})/(1-\delta_{i_m}^{\tau_{i_m}^I})] U_{i_m h_m} = \mathbf{x}_{h_m}' \beta_d - p_{h_m}\alpha + \xi_{h_m} + \eta_{i_m h_m}.$$

A consumer chooses to enter the h_mth home if $V_{i_m h_m} = \max_{k_m \in \{0,1,2,\ldots,H_m\}} \{V_{i_m k_m}\}$, where the subscript 0 represents an outside option for which $V_{i_m 0} = \eta_{i_m 0}$. To establish an econometric model without consumer information, the conditional distribution of $\eta_{i_m h_m}$, given $\mathbf{x}_{h_m}$, and p_{h_m} are assumed to be i.i.d. type I extreme value distributions. Then, this term is integrated out from the demand-side model as

$$
\begin{aligned}
s_{h_m} &= \int_{V_{i_m h_m} = \max_{j \in \{0,1,2,\ldots,H_m\}} \{V_{i_m j}\}} V_{i_m h_m} \pi(d\eta_m) \\
&= \begin{cases} \frac{\exp[\mathbf{x}'_{h_m}\beta_d - p_{h_m}\alpha + \xi_{h_m}]}{1+\sum_{k_m=1}^{H_m} \exp[\mathbf{x}'_{k_m}\beta_d - p_{k_m}\alpha + \xi_{k_m}]} & \text{for} \quad h_m = 1, ..., H_m \\ \frac{1}{1+\sum_{k_m=1}^{H_m} \exp[\mathbf{x}'_{k_m}\beta_d - p_{k_m}\alpha + \xi_{k_m}]} & \text{for} \qquad h_m = 0 \end{cases} .
\end{aligned}
$$

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

$$
q_{h_m} = \mathbf{x}'_{h_m}\beta_d - p_{h_m}\alpha + \xi_{h_m} + \ln\Big(1 - \sum_{k_m=1}^{H_m} \exp(q_{k_m})\Big). \tag{3.3}
$$

3.3. Modeling the supply side

The supply-side model represents the profit maximization problem of private nursing homes. Following previous nursing home studies, as surveyed in Norton (2000), two assumptions are placed on the homes' optimization problem. First, to simplify the problem, the capacity of homes is not considered. Second, what homes maximize is their expected profit for a period at a steady state, given their occupancy status. This assumption is required to avoid a complicated situation in which residence durations overlap. In equilibrium, the marginal profit function from the ith consumer for the h_mth home takes the following form:

$$
\Pi_{h_m}(\tau_{i_m}) = p_{h_m} + f_{h_m} - mc_{h_m} - f_{h_m} I[T_{h_m} < \tau_{i_m}], \tag{3.4}
$$

where mc_{h_m} is the marginal cost per resident, which is assumed to be homogeneous with respect to residents.

The home maximizes its expected profit, using the per capita profit defined by equation (3.4). Four assumptions are used to ensure a well-defined supply side competition. First, homes are assumed to know the distribution of $\eta_{i_m h_m}$. Under this assumption, $I_m s_{h_m}$ is the equilibrium number of residents.

The second assumption is to do with the control variables for the homes. The assumption here is that homes play the Bertrand-type price competition, but that the monthly rent, f_{h_m}, is determined exogenously in the housing market. Under this assumption, the home has two control variables, p_{h_m} and T_{h_m}, which are complementary. Then, I further assume that homes first decide T_{h_m}, and then choose p_{h_m} based on this decision.

Third, a home's subjective expectations for their customers' lifetimes are assumed to be determined only by T_{h_m} and f_{h_m}. Therefore, these expectations are not affected by p_{h_m} or the actions of other homes. In other words, homes choose T_{h_m} to control which customers they wish to attract, and p_{h_m} determines how to manage their expenditure, given these 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. Furthermore, I assume that this subjective probability can be modeled using the logit polynomial, which is given as

$$\begin{aligned} 1 - Pr^s(T_{h_m} < \tau | f_{h_m}, T_{h_m}) &= \Gamma(T_{h_m}, f_{h_m}; \gamma) \\ &= \frac{\exp(\gamma_0 + \gamma_{T1} T_{h_m} + \gamma_{T2} T_{h_m}^2 + \gamma_{f1} f_{h_m} + \gamma_{f2} f_{h_m}^2 + \gamma_{TF} T_{h_m} f_{h_m})}{1 + \exp(\gamma_0 + \gamma_{T1} T_{h_m} + \gamma_{T2} T_{h_m}^2 + \gamma_{f1} f_{h_m} + \gamma_{f2} f_{h_m}^2 + \gamma_{TF} T_{h_m} f_{h_m})}, \end{aligned}$$

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

Fourth is a conventional assumption of a log-linear form for 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, $\mathbf{w}_{h_m}$, and an unobservable home specific effect, ω_{h_m}.

Under the above four assumptions, the expected profit function is specified as follows:

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

A further assumption is that there exists an interior solution for the profit maximization problem. The first-order condition for p_{h_m}, given T_{h_m}, yields

$$p_{h_m} = \exp(\mathbf{w}_{h_m} \beta_s + \omega_{h_m}) - \frac{1}{\alpha[\exp(q_{h_m}) - 1]} - f_{h_m} \Big[1 - Pr^s(T_{h_m} < \tau | f_{h_m}, T_{h_m}) \Big] \Big\}. \quad (3.5)$$

Consequently, the structural equations (3.3) and (3.5) provide a simultaneous model for the demand and supply for $h_m = 1, ..., H_m$. This model consists of $2H_m$ (not 2) simultaneous equations, because in (3.3), q_{h_m} depends not only on p_{h_m}, but also on q_{k_m}, for $k_m \neq h_m$. The dependent variables are $(\mathbf{q}_m, \mathbf{p}_m)$, the unobserved variables are (ξ_m, ω_m), and the coefficient parameters are $\tilde{\theta} = (\beta_d', \beta_s', \alpha, \gamma')'$.

4. An econometric framework via nonparametric Bayes

This section provides details of the econometric methodology for the model given in (3.3) and (3.5). The section begins by explaining the difficulty with prediction analysis based only on moment conditions. This holds for structural econometric models in general, although the explanation concentrates on the model used in this study.

Conventional estimation methods are employed based on moment conditions in the form of $E[z\xi|\mathbf{x}, \mathbf{w}] = E[z\omega|\mathbf{x}, \mathbf{w}] = 0$, where z is some instrument. For this nonlinear simultaneous model, there is a simpler estimation method, based on moment conditions, than the method based on a likelihood function. The latter method requires evaluating the joint distribution function accompanied by $2H_m$ nonlinear structural equations. This is a reason why the GMM estimation is the standard econometric tool for the BLP model.

Prediction analysis provides a different situation. As an illustrative example, consider a prediction problem for p, given counterfactual $\mathbf{x}_m = \tilde{\mathbf{x}}$ and $\mathbf{w}_m = \tilde{\mathbf{w}}$. For simplicity, assume that the reduced form is analytically obtained for p_{h_m} as

$$p_{h_m} = g_p(\mathbf{x}_m, \mathbf{w}_m, \xi_m, \omega_m; \theta).$$

There are three unknown factors on the right-hand side of the equation, namely ξ_m, ω_m, and θ. For θ, GMM estimates can be used. However, there is the problem of the unobservables, (ξ_m, ω_m).

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, the counterfactual values of the explanatory variables are predicted using the same values of (ξ_m, ω_m). Specifically, he substituted the values of the estimated residuals into the reduced form equation, (4).

On the other hand, in the counterfactual situation in this study for the private nursing home market, eliminating the lump-sum forward payment would impose a drastic change in the market structure. To reflect this large change, I try to allow different values of (ξ_m, ω_m) under the counterfactual situation. To do so, these unobservables are set as stochastic terms in the prediction analysis and integrated from the reduced form, (4).

However, using the conventional econometric method via moment conditions, 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 (ξ_m, ω_m), such that

$$p_{h_m} = g_p(\mathbf{x}_m, \mathbf{w}_m, \xi_m, \omega_m; \theta) = \tilde{g}_p(\mathbf{x}_m, \mathbf{w}_m; \theta) + \mathbf{a}'\xi_m + \mathbf{b}'\omega_m,$$

where $\mathbf{a}$ and $\mathbf{b}$ are constant vectors. Multiplying both sides by z_m, we can integrate out the terms for (ξ_m, ω_m) using the moment conditions. However, it is difficult to guarantee such an assumption of the functional form, because g_p is a reduced form. Thus, in general, we need to conduct a numerical integration for (ξ_m, ω_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.

4.1. Estimation via nonparametric Bayesian analysis

This subsection provides an intuitive introduction to the nonparametric Bayesian approach. A practical and detailed explanation of the estimation procedure is provided in in Appendix A. The nonparametric Bayesian analysis can be summarized as a statistical methodology that uses a likelihood function, which can represent an arbitrary distribution. Unlike other nonparametric models, such as models that use kernel or spline statistics, the nonparametric Bayes models are associated with well-defined closed forms of the likelihood function and predictive distributions. This feature enables us to conduct a numerical integration in a prediction analysis.

This study incorporates nonparametric modeling for ξ_{h_m} and ω_{h_m}. Here, $\mathbf{x}_{h_m}$ and $\mathbf{w}_{h_m}$ are assumed to include constant terms to measure the factors common to all homes, and ξ_{h_m} and ω_{h_m} represent home-specific residual terms. Since it is difficult to adopt a mean restriction in nonparametric Bayes models, this study incorporates median conditions for the residual terms, such that $\text{Med}(\xi_{h_m}|\mathbf{x}_{h_m}, \mathbf{w}_{h_m}) = \text{Med}(\omega_{h_m}|\mathbf{x}_{h_m}, \mathbf{w}_{h_m}) = 0$. The Polya tree mixture is adopted here as a nonparametric Bayesian method that works under the median constraint.

An intuitive way to define a nonparametric likelihood function is through predictive densities. Consider a problem of estimating a joint conditional distribution for $(\omega_1, \omega_2 ..., \omega_H)$ given $\tilde{\theta}$. Suppose we have nonparametric predictive densities $\pi(\omega_h|\omega_1, ..., \omega_{h-1}, \tilde{\theta})$, for all $h = 1, ..., H$. Then, the joint conditional density function can be derived as

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

The remaining concern in this setting is the 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 that of a histogram. See Appendix A.1 for a detailed discussion of this method.

4.2. Counterfactual prediction analysis

The main purpose of this study is to simulate a counterfactual situation in which the lump-sum forward payment would be eliminated. Although we want to compare consumer welfare with and without the mechanism, it is difficult to derive the lifetime utility because consumers' lifetimes and the time-discount factor are not observed. Instead, the total amounts of lifetime payments are compared. Because the interest rate is not considered, this comparison is conservative in the sense that the present value of the lifetime payment without the lump-sum forward payment is over-evaluated in terms of utility.

For this prediction analysis, three additional assumptions are added to yield the same economic model described by (3.3) and (3.5) under the counterfactual $f_{h_m}^{new} = 0$ and $T_{h_m}^{new} = 0$. First, new nursing homes do not enter the market in the counterfactual situation. This assumption is justified, at least in the short term, because the high sunk costs and the Regulation of Volume are significant barriers to entry. In other words, the prediction analysis is limited to the short-run effects of a change to the payment mechanism.

Second, the current and the counterfactual situations are assumed to be generated from the same, single equilibrium, even if there are potentially multiple equilibria for the Bertrand competition on the supply side. When conducting estimations in game theoretic models, it is commonly assumed that the dataset is generated by a single equilibrium to guarantee identification[9]. In addition, the analysis in this study requires that the single equilibrium should be played in the counterfactual situation, as mentioned in footnote 12 of Berry et al. (1995, p.853).

Third, the home-specific error terms, ξ_{h_m} and ω_{h_m}, have the same distributions under the counterfactual situation. This is still a restrictive assumption, but it relaxes the assumption of Nevo (2000), which used the same values for the error terms in the counterfactual situation.

Under these assumptions, the prediction analysis consists of two steps. The first step is the prediction of the monthly fees, p_{h_m}, for the counterfactual situation. The second

[9]Several studies, such as Ciliberto and Tamer (2009), have proposed estimation methods that allow multiple equilibria as a data generating process. However, it is difficult to adopt these methods into this already complicate model.

step calculates the lifetime payments with and without the lump-sum forward payment mechanism. Details of the prediction for p_{h_m} can be found in Appendix B. Hereafter, the predicted monthly fee values, $p_{h_m}^{new}$, are assumed to be derived using this prediction technique.

The calculation of the lifetime payments with and without the lump-sum forward payment begins by considering a match between a resident, i_m, and a home, h_m. If there is no lump-sum forward payment, the lifetime payment consists of the accumulated monthly fees throughout the resident's lifetime, τ_{i_m}, because this is the only form of payment. Then, the predicted total payment is $p_{h_m}^{new}\tau_{i_m}$. Because interest rates are not considered and a simple summation is used, this amount is a lower bound for the present value of the consumer's total payment.

For the lifetime payment with the lump-sum forward payment mechanism, three cases are considered separately based on the value of τ_{i_m} and T_{h_m}. The first case is $T_{h_m} = 0$, where the home does not collect a lump-sum forward payment, even in the current situation. This case is abbreviated in the prediction analysis as it is not of particular interest in this study. The remaining two cases assume that $T_{h_m} > 0$.

The second case is $\tau_{i_m} \leq T_{h_m}$, where the consumer exits the home before the expiration date. In this case, the lifetime payment with the lump-sum forward payment mechanism is $(p_{h_m} + f_{h_m})\tau_{i_m}$. Because the payments with and without the mechanism are both multiplied by τ_{i_m}, this term cancels out when comparing the lifetime payments for the current and counterfactual situations. Thus, the difference between the lifetime payments can be detected by comparing $p_{h_m} + f_{h_m}$ and $p_{h_m}^{new}$, regardless of τ_{i_m}.

The third case is $\tau_{i_m} > T_{h_m}$, where the consumer lives longer than the lump-sum forward payment and there is no additional rent payment. The resulting lifetime payment is $p_{h_m}\tau_{i_m} + f_{h_m}T_{h_m}$. Unlike the previous case, the unobservable true lifetime τ_{i_m} cannot be ignored when comparing the lifetime payments with and without the lump-sum forward payment mechanism. In practice, several representative consumers are considered with lifetimes of $\tau_{i_m} = 240$ and 360, or 20 and 30 years of remaining life.

4.3. Identification

Identification in this model depends on three factors. In addition to the functional form assumption and the distributional assumptions on (ξ_m, ω_m) described earlier, we also need exclusion restrictions.

To specify the requirements for the exclusion restrictions, note that, (3.3) and (3.5) define the models for q_{h_m} given $p_{h_m}, \mathbf{q}_{(-h_m)}$ and ξ_{h_m} and for p_{h_m} given q_{h_m} and ω_{h_m}. Then, the demand side needs a standard exclusion restriction that it is not correlated to p_{h_m} or $\mathbf{q}_{(-h_m)}$, but is correlated to q_{h_m}. For the supply side, p_{h_m} does not depend on $\mathbf{p}_{(-h_m)}$ and $\mathbf{q}_{(-h_m)}$, given q_{h_m}. Thus, it is sufficient to have exclusion restrictions that is not correlated to q_{h_m}, but is correlated to p_{h_m}. In other words, we do not need to consider the correlation to $\mathbf{p}_{(-h_m)}$. Thus, the required exclusion restriction for the supply side should not always be home-specific variables, but rather market-specific variables.

As a price of the flexible prediction analysis via nonparametric Bayesian modeling, this study has a disadvantage in handling heterogeneity among consumers. Many previous

studies have adopted random coefficient modeling for individual heterogeneity. On the other hand, this study locates the heterogeneity in the error term, then integrates it out of the econometric model, in a similar manner described in Berry (1994). This approach allows for closed-form expressions for the error terms, which are required when constructing the likelihood function via a change of variables.

5. An empirical study of the Japanese nursing home market

5.1. Data

This section applies the 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, because private nursing homes are legally obligated to disclose the information when asked. However, as it is burdensome to obtain a disclosure for all the homes, a list in a consumers' guidebook is used, Shuukan Asahi Mook (2011), which is a special volume of a leading weekly news magazine in Japan.

Prefectures are used as the unit for local markets. Prefectures are the largest sub-national jurisdictions in Japan. An important assumption for the BLP model is that markets are geographically isolated. To guarantee this assumption, the markets need to incorporate a relatively large area. The prefecture is an ideal unit for this purpose. Furthermore, since the prefecture authority has the power to grant an operating license, private homes are likely to make a location decision at the prefecture level.

The sample consists of 1,265 homes. The sampling methodology is as follows: the editors of Shuukan Asahi Mook (2011) sent 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 homes listed in the book, 745 were eliminated because they provide optional long-term care. From the remaining 1,598 homes, a further 324 homes were removed from the sample because of missing information. Then, nine homes were excluded as they were the only home in their prefecture, because a monopoly market would yield a different market structure to the oligopoly model being studied here. Approximately half the homes did not respond to the questionnaire. The low response rate might have been caused by the enforcement of an early deadline by the editors (three months).

The share of a home, s_{h_m}, is defined as the ratio of the number of residents in the h_mth home to the number of the potential consumers in the market, m. As seen in equation (3.3), there must be a positive share for the outside option. Then, those who did not choose to live in a private nursing home should be included as potential consumers. As the potential consumers, I adopt elderly(age 65 and older) with the LTCI eligibility level of care required one or more, because they typically receive institutional care under coverage of the LTCI. Table 1 shows the number of homes and the elderly in the different prefectures.

Table 1 is here

For variables related to the payment mechanism, there is a monthly fee, p_{h_m}, a lump-sum forward payment, F_{h_m}, and an expiration period, T_{h_m}. Several homes report two price variables, namely a minimum and maximum. Specifically, the expiration period is unique for 1,226 homes, but the monthly fee and the lump-sum forward payment are only unique for approximately half the homes. Since there are only a few homes that offer multiple expiration period options, the variation in lump-sum forward 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 does not seem to occur.

There are two categories of explanatory variables: components of $\tilde{\mathbf{x}}$ are characteristics of the consumer utility, and components of $\mathbf{w}$ are characteristics of the marginal cost per resident. In addition to the common elements for these two categories, exclusion restrictions are required that are home-specific variables in $\tilde{\mathbf{x}}$ and market-specific variables in $\mathbf{w}$, as mentioned in Section 4.3.

There are three variables for the common observable elements on the demand and supply sides, as per Shuukan Asahi Mook (2011): the number of residents per worker, a dummy variable that takes a value of unity when the home is operated by a chain, and the number of years since opening[10].

With respect to chains, most of chains consist of franchise homes in the private nursing home market. Therefore, these chains are not likely to make a network-level decision, but are likely to share operational methods, such as advertising. Thus, there is no distinct behavioral model for chains in this study, and their behavioral difference is considered as an explanatory variable. There is no dominant chain that has more than 20% of the share in the Japanese private nursing home market. Therefore, a dummy variable is created as a bundle of six chains, namely Benesse Style Care, Message, Watami no Kaigo, Nichii Group, Life Commune, and Tsukui. These six chains account for more than 20 homes in the dataset.

The exclusion restrictions are as follows. On the demand side, there is a home-specific variable for the occupancy rate. This variable affects the consumer utility as a proxy for the quality of the home, because an extremely small occupancy rate might be a signal that the home has some problems. In contrast, this variable does not affect the per-resident cost, because this is an aggregate value with regard to residents.

The supply-side exclusion restrictions are two market-specific variables of cost shifters, namely the local averages of rents and wages. These variables naturally affect the marginal cost, but not the utility, given the other price variables. The local average for rent is defined as an annual average of monthly rents per 3.3 m^2 in the capital city of a prefecture. These figures are taken from the 2010 Annual Report on the Retail Price Survey. For the average wage, there are no reliable data specific to care workers.

[10] Another dummy variable was tried that takes unity when the home is operated by a non-profit organization. However, this was eliminated from the empirical study because the convergence of its coefficient is quite slow, and there is no serious difference to the 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 the sample.

Therefore, information is adapted from the medical and welfare sectors, which include care workers. The local average wage is calculated by dividing the annual wages plus bonuses by 12. These components are from the 2010 Basic Survey on Wage Structure.

Table 2 is here

Table 2 presents the descriptive statistics for the explanatory variables. Several volatile variables have been adjusted to stabilize the estimation in the following inference exercises. First, the number of years since opening, the local average rent, and the local average wage are standardized to have zero means and unit variances. Second, several variables generated from T and f are divided by constants. Specifically, F is divided by 1,000 and T, T^2, f and f^2, which appear in $\Gamma(T, f; \gamma)$, are divided by $10, 1,000, 10$ and 100, respectively.

5.2. Estimation results

In the implementation of the MCMC samplers, 1,000,000 posterior samples were generated after discarding 100,000 initial samples as the burn-in period. The computation took approximately 20 days using three cores of Intel Xeon X5470 processors (3.33GHz). The values of the hyperparameters and their sensitivity are discussed in Appendix A.2.3.

Table 3 is here

Table 3 reports the estimation results. The first and second columns show the posterior means and standard deviations, the third column represents the 95% credible intervals, and the last column reports the inefficiency factors (IF). The maximum of the inefficiency factors is $1,414$, which implies that we would obtain the same variance in the posterior sample means from more than $1,000,000/1,414 \simeq 707$ uncorrelated draws, even in the worst case. For the sake of the convergence diagnosis, the figures for the posterior sample paths and the posterior densities of the MCMC samples are presented in Appendix C.

Overall, the estimated posterior means for the coefficient parameters take reasonable values. On the demand side, the number of residents per worker has a negative effect on consumer utility, because consumers prefer homes with sufficient care workers. The positive coefficient for the chain dummy indicates that chains provide efficient service through scale effects. The effect of years since opening is not clear, indicating that history plays a complicated role. A home that has existed for a long time might imply that it is of a high quality, based on an accumulation of experience. However, it might also indicate disutility from old facilities. In addition, the occupancy rate, which is an exclusion restriction, has a positive effect on consumer utility, which implies that the popularity of a home is a good proxy for its quality. The mean effects of the lump-sum forward payment and the monthly rent on utility, α_F and α_f, contain zero in their 95% credible intervals. It is difficult to interpret these results on their own. The payment mechanism will be discussed in further detail in the prediction analysis.

On the supply side, because the logarithm of the marginal cost is defined as $\log(mc_{h_m}) = \mathbf{w}'_{h_m}\beta_s + \omega_{h_m}$, a positive coefficient means that the corresponding explanatory variable

increases the marginal cost, and hence decreases the profit of homes. The negative effect of the number of residents per worker means that a reduction in the labor force decreases the marginal cost, as expected. The positive coefficient for the chain dummy might imply that the advertisement costs incurred by chains are greater than the revenue generated from the efficient operation of the chains. Years since opening has a negative effect. This can be interpreted to mean that an accumulation of operational experience decreases running costs. For the exclusion restrictions, the local average rent has a positive coefficient as a factor that increases marginal cost. However, the sign of the local average 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.

The results show a bimodal posterior density for 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. This is seen in the stable sample path in which the chain repeatedly visits both peaks. The peculiar shape of this posterior density function indicates that the distribution of ω_m should be different to the common probability distributions, such as the normal distribution. This result shows the advantage of using a nonparametric estimation.

5.3. Prediction results

This section presents the prediction analysis based on the estimation results. 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 4.2. Instead, the prediction analysis concentrates on the Shizuoka prefecture, which has 32 homes. The Shizuoka prefecture has divergence in industries such as agriculture, fishing, tourism, and manufacturing. Owing to this economic status, there is a wide variety of nursing homes, from reasonable quality homes to luxury private nursing homes, which were formerly hotels. Thus, the Shizuoka prefecture is a suitable target for the prediction analysis

In the dual-loop Monte Carlo integral, the number of iterations for the outer loop is set to $L = 50$. The values for θ^l are taken from the posterior samples obtained in the estimation step at intervals of 1,500 periods. Because the maximum of the inefficiency factors is 1,414, these values of θ^l can be treated 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 θ^l, $l = 1, ..., 50$, the posterior sample paths of the predicted values exhibit a sufficient convergence of the inner loops.

In the following prediction analysis, the lifetime payments for homes that currently do not collect the lump-sum forward payment are not compared, although their information is used for the prediction. Therefore, the target is the remaining 19 homes.

Figure 2 is here

Figure 2 presents a prediction result for the match of short-lived consumers and longer expiration periods, $\tau_i \leq T_h$. The X-axis indexes homes, while the Y-axis measures

payments in 10,000 yen. Each home has two bars of monthly payments to indicate the situation with and without the lump-sum forward payment mechanism. Figure 2 shows that lifetime payments under the counterfactual are smaller. In other words, short-lived consumers can reduce their lifetime payment by abandoning the lump-sum forward payment mechanism. In this situation, longevity risks are pooled and distributed uniformly to all residents. Therefore, short-lived consumers cannot recoup their risk premium and forced overpayment.

Figures 3 and 4 are here

Next, consider a match between long-lived consumers and shorter expiration periods, $T_h < \tau_i$. Figures 3 and 4 gives comparisons of $p_h\tau + f_hT_h$ and $p_h^{new}\tau$ for $\tau = 240$ and 360, or 20 and 30 years of remaining life, respectively. The two bars show the lifetime payments in the situation with and without the lump-sum forward payment mechanism.

The lifetime payments in the counterfactual situation exceed the actual payments only in the case in which a consumer with 30 years of remaining life chooses a specific home. In other words, to recoup the risk premium, consumers need to live in the home for at least 30 years. However, in practice, 30 additional years of life is not realistic for people entering a private nursing home with care. In addition, the expenses without the lump-sum forward payment can be further reduced if we consider an interest rate.

The above results show that the lump-sum forward payment mechanism generally forces consumers to pay more. This overpayment might be a result of risk averseness, in the case of both homes and consumers. There must be a loss of consumer welfare at the aggregate level, because the overpayment is common for most consumers.

The empirical results present several ideas for the efficient management for longevity risk in Japan. The consumer surplus can be improved by abandoning the lump-sum forward payment mechanism. On the other hand, the results also show the sharp risk-averseness of both consumers and homes. If consumers accept the very high-risk premium, there might be a demand for annuities and private long-term care insurance, in addition to the public LTCI. On the other hand, the risk-averseness of homes can be managed if there is a method to hedge the longevity risk, such as the survivor bond described in Blake and Burrows (2001).

6. Conclusion

This study analyzed the unique lump-sum forward payment that is used to manage longevity risk in the Japanese private nursing home market. To consider the impact of this system, a prediction analysis was conducted using a structural model based on the work of Berry et al. (1995). This study also proposed a flexible nonparametric Bayesian approach for the econometric analysis. The prediction results imply that the outdated system forces consumers today to pay higher lifetime payments than necessary.

Acknowledgment

I would like to thank Yasuhiro Omori, Jiro Nakamura, Yuji Genda, Katsumi Shimotsu, Naoya Sueishi, Hirokazu Ishise, Mototsugu Fukushige, Yuichiro Kanazawa, Naoto Kunitomo, Akimichi Takemura, Munenori Ando, Yasusada Murata, Ryo Okui, Yoichi Nishiyama, and seminar and workshop participants at the University of Tokyo, Nihon University, Kyoto University, Okinawa International University, ISBA, JuBES, JEA, JJSS, JYSG, and NASM for their helpful comments. The remaining errors are my responsibility. This work is supported by the Grant-in-Aid for Research Activity start-up No. 25885021 from the Japanese Society for the Promotion of Science. Also, the previous version of this paper, which was titled "A nonparametric Bayesian approach for counterfactual prediction with an application to the Japanese private nursing home market," was supported by Grant-in-Aid for Scientific Research (A) No. 21243018 from the Japanese Society for the Promotion of Science. The computational results are obtained using Ox version 6 (Doornik (2011)).

References

S. Berry. Estimating discrete-choice models of product differenciation. *RAND Journal of Economics*, 25:242–262, 1994.

S. Berry, J. Levinsohn, and A. Pakes. Automobile price in market equilibrium. *Econometrica*, 63(4):841–890, 1995.

D. Blake and W. Burrows. Survivor bonds: Helping to hedge mortality risk. *Journal of Risk and Insurance*, 68(2):339–348, 2001.

Jerrrey R. Brown and Amy Finkelstein. Why is the market for long-term care insurance so small? *Journal of Public Economics*, 91:1967–1991, 2007.

P. Chintagunta, J.-P. Dubé, and Goh K. Y. Beyond the endogeneity bias: The effect of unmeasured brand characteristics on household-level brand choice models. *Management Science*, 51(5):832–849, 2005.

R. Christensen, T. Hanson, and A. Jara. Parametric nonparametric statistics: An introduction to mixture of finite Polya trees. *The American Statistician*, 62(4):296–306, 2008.

Federico Ciliberto and Ellie Tamer. Market strucuture and multiple equilibria in airline markets. *Econometrica*, 77(6):1791–1828, 2009.

D. Cutler. Why don't markets insure long-term risk? Working paper, 1996.

T. Davidoff and G. Welke. Selection and moral hazard in the reverse mortgage market. Working paper, 2006.

P. Davis. Spatial competition in retail markets: Movie theaters. *RAND Journal of Economics*, 37(4):964–982, 2006.

J. A. Doornik. *Object-Oriented Matrix Programming Using Ox*. Timberlake Consultants Press and Oxford, London, 6th edition, 2011.

T. S. Ferguson. Prior distributions on spaces of probability measures. *Annals of Statistics*, 2:615–629, 1974.

Amy Finkelstein and James Poterba. Adverse selection in insurance market: Policyholder evidence from the U. K. annuity market. *Journal of Political Economy*, 112(1):183–208, 2004.

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

T. Hanson. Inference for mixtures of finite Polya tree models. *Journal of the American Statistical Association*, 101:1548–1565, 2006.

T. Hanson and W. O. Johnson. Modeling regression error with a mixture of Polya trees. *Journal of the American Statistical Association*, 97:1020–1033, 2002.

N. Ikegami and J. Campbell. Long-term care insurance comes to Japan. *Health Affairs*, 19(3):26–39, 2000.

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

Japan Ministry of Health, Labor and Welfare. Basic Survey on Wage Structure. Tokyo, 2010.

Japan Ministry of Health, Labor and Welfare. Survey on Institutions and Establishments for Long-Term Care (Kaigo Sabisu Sisetsu Jigyousho Chousa, in Japanese). Tokyo, Various years.

Japan Statistical Bureau. Annual Report on the Retail Price Survey. Tokyo, 2010.

A. Jara, T. Hanson, and E. Lesaffre. Robustifying generalized linear mixed models using a new class of mixtures of multivariate Polya trees. *Journal of Computational and Graphical Statistics*, 18(4):838–860, 2009.

R. Jiang, P. Manchanda, and P. E. Rossi. Bayesian analysis of random coefficient logit models using aggregate data. *Journal of Econometrics*, 149(2):136–148, 2009.

M. Lavine. Some aspect of Polya tree distributions for statistical modeling. *Annals of Statistics*, 20(3):1222–1235, 1992.

M. Lavine. More aspect of Polya tree distributions for statistical modeling. *Annals of Statistics*, 22(3):1161–1176, 1994.

R. D. Mauldin, W. D. Sudderth, and S. C. Williams. Polya trees and random distributions. *Annals of Statistics*, 20(3):1203–1221, 1992.

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

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

O. Mitchell, J. M. Poterba, M. Warshawsky, and J. R. Brown. New evidence of the money's worth of individual annuities. *American Economic Review*, 89(5):1299–1318, 1999.

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

A. Musalem, E. Bradlow, and J. S. Raju. Bayesian estimation of random-coefficient choice models using aggreagate data. *Journal of Applied Econometrics*, 24(3):490–516, 2009.

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

E.C. Norton. Long-term care. In J. Newhouse and A. Culyer, editors, *Handbook of health economics*, volume 1B, pages 95–994. North-Holland, Amsterdam, 2000.

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

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

Y. Sudou. The history and future perspective of the business of private nursing homes 2(in Japanese). *Aioi Kiso-ken Review*, 2:52–71, 2006b.

N. Tamiya, H. Noguchi, A. Nishi, M. R. Reich, N. Ikegami, H. Hashimoto, K. Shibuya, I. Kawauchi, and J. C. Campbell. Population ageing and wellbeing: lessons from japan's long-term care insurance policy. *Lancet*, 378:1183–1192, 2011.

T. Tsutsui and N. Muramatsu. Japan's universal long-term care system reform of 2005: Containing costs and realizing a vision. *Journal of the American Geriatrics Society*, 55(9):1458–1463, 2007.

S. Yang, Y. Chen, and G. M. Allenby. Bayesian analysis of simultaneous demand and supply. *Quantitative Marketing and Economics*, 1(3):251–275, 2003.

A. Details of the nonparametric Bayesian estimation

This appendix complements the description of the nonparametric Bayesian analysis in Section 4.1. First, an intuitive definition of the Polya tree mixture is provided, followed by a description of the estimation procedure corresponding to the nonparametric Bayesian modeling. Lastly, the values of the hyperparameters and primitives actually used in the empirical study in Section 5 are presented.

A.1. Polya tree mixture

This section first describes the Polya tree mixture in a similar way to the intuitive description provided by Christensen et al. (2008). The Polya tree, as invented by Ferguson (1974) and developed theoretically by Lavine (1992, 1994) and Mauldin et al. (1992), is first defined, and then the Polya tree mixture is introduced as an extension to the original Polya tree.

As in Section 4.1, the estimation problem is first considered for the joint conditional distribution for $(\omega_1, \omega_2 ..., \omega_H)$, given $\tilde{\theta}$. The scalar random variable, ω_h, is assumed to have a common support, Ω, for all h.

To formulate a Polya tree, econometricians need to specify a base measure, G, on the support, Ω, which has a well-defined density, g, and a known median, μ. Using the base measure, the prior distribution of the Polya tree is constructed using a J-step iterative process. In the first level, 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}, respectively, such that $\lambda_{21} + \lambda_{22} = \lambda_{11}$. Such binary separations are repeated until the terminal level, J.

In the above construction, the probabilities of bins λ_{j,κ_j}, for $j = 1, ..., J$ and $\kappa_j = 1, 2, ..., 2^j$, are unknown parameters. Therefore, we use a hierarchical Bayesian model to introduce a simple probabilistic structure for these parameters. From the above construction using binary separations, we can see that 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

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

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

$$
\begin{aligned}
\zeta_{j,\kappa_{j-1}} &\sim Beta(\alpha_{j,2\kappa_{j-1}-1}, \alpha_{j,2\kappa_{j-1}}), \\
\zeta_{j,\kappa_{j-1}}|\omega_1, ..., \omega_{i-1} &\sim Beta(\alpha_{j,2\kappa_{j-1}-1} + n_{j,2\kappa_{j-1}-1}, \alpha_{j,2\kappa_{j-1}} + n_{j,2\kappa_{j-1}}),
\end{aligned}
$$

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

The result is histogram-like probability models for the J levels. The Polya tree utilizes information from all the J levels by multiplying them together. Furthermore, the purpose is to construct a nonparametric predictive density of $\omega_i|\omega_1, ..., \omega_{i-1}, \tilde{\theta}$. For this purpose, the nuisance parameters, $\lambda = \lambda_{11}, \lambda_{12}, ..., \lambda_{J,2^J}$, are integrated out. Consequently, we have

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

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

The original Polya tree as defined above has a similar weakness to histograms, namely discontinuity at the borders of the bins. 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

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

To achieve the median constraint, it is assumed that μ does not depend on τ but τ determines the percentiles of G_τ other than μ. Then, this integration smooths the bins defined in the second and later levels. On the other hand, the border in the first level at μ is kept unchanged to guarantee the median restriction that the marginal distribution for ω_i satisfies $Pr(\omega_i \leq \mu) = 1/2$.

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. Here, $N(0, 1/\tau)$ is employed as the base measure for the analysis of Japanese nursing homes. Then, τ is a scale parameter that needs to be to be estimated. The second primitive is the set of hyperparameters of the Polya tree prior, α_{j,k_j}, for $j = 1, ..., J$ and $k_j = 1, 2, ..., 2^j$. This study adopts 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 primitive is the truncation level, J. Choices for c and J are case-specific subjects, and are discussed further in Appendix A.2.3.

A.2. The Bayesian Estimation procedure

A.2.1. The likelihood function

This section describes the estimation framework specific to the econometric model for the Japanese nursing home market, as described by the structural equations (3.3) and (3.5). As a result of the mutual dependencies of the dependent variables, these structural equations cannot be used directly to define the likelihood function. Instead, the likelihood function is obtained using a change of variables from unobservables to dependent variables, as suggested by Chintagunta et al. (2005). It is assumed that the distributions of the unobservable terms follow independent Polya tree mixtures, with scale parameters τ_ξ and τ_ω.[11]. Thus, the parameter is $\theta = (\tilde{\theta}', \tau_\xi, \tau_\omega)'$. The resulting likelihood function is

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

where

$$\xi_{h_m}(\mathbf{p}_m, \mathbf{q}_m; f_{h_m}, T_{h_m}, \tilde{\theta}) = q_{h_m} - \ln\Big[1 - \sum_{k_m=1}^{H_m} \exp(q_{k_m})\Big] - \tilde{\mathbf{x}}'_{h_m}\tilde{\beta}_d + T_{h_m} f_{h_m} \alpha_F + p_{h_m}\alpha, \quad \text{(A.2)}$$

$$\omega_{h_m}(\mathbf{p}_m, \mathbf{q}_m; f_{h_m}, T_{h_m}, \tilde{\theta}) = \ln\left[p_{h_m} + \frac{1}{\alpha[\exp(q_{h_m}) - 1]} + f_{h_m}\Gamma(T_{h_m}, f_{h_m}; \gamma)\right] - \mathbf{w}'_{h_m}\beta_s. \quad \text{(A.3)}$$

,

and $\mathcal{J}_m$ is the following Jacobian matrix of the transformation:

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

[11] Jara et al. (2009) considered the multivariate version of the Polya tree mixture, which can capture the correlation between random variables. However, this methodology was not used in this study as it is difficult to include in the already complicated model.

in which

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

Using the formula for the determinant by parts, and the fact that $Z_{h_m} > 0$, which is guaranteed under the support condition described later in (A.5), we have

$$
|\det(\mathcal{J}_m)| = |\det(D_m)| \Big(\prod_{k_m=1}^{H_m} Z_{k_m} \Big),
$$

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

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

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

The likelihood function requires additional restrictions for the supports of the dependent variables to have well-defined logarithmic terms in (A.2) and (A.3). Specifically,

$$
0 < 1 - \sum_{k_m=1}^{H_m} \exp(q_{k_m}), \tag{A.4}
$$

$$
0 < p_{h_m} + \frac{1}{\alpha[\exp(q_{h_m}) - 1]} + f_{h_m}\Gamma(T_{h_m}, f_{h_m}; \gamma). \tag{A.5}
$$

Condition (A.4) is automatically satisfied in the estimation step owing to the construction of q_{h_m} in Section 3, but it must be verified in the prediction step described below. Another condition, (A.5), is required in the estimation and prediction steps. Furthermore, because (A.5) states that the support of the likelihood function depends on parameters α and γ, (A.5) violates a regularity condition for the maximum likelihood estimators to have preferable asymptotic properties. This fact is another motivation to adopt a Bayesian estimation procedure.

A.2.2. Prior and proposal distributions

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

The prior distributions 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,

$$\begin{aligned}
\beta_d &\sim N(\mu_{\beta_d 0}, \Sigma_{\beta_d 0}), \quad \beta_s \sim N(\mu_{\beta_s 0}, \Sigma_{\beta_s 0}), \\
\gamma &\sim N(\mu_{\gamma 0}, \Sigma_{\gamma 0}), \quad \alpha \sim N(\mu_{\alpha 0}, \sigma^2_{\alpha 0}), \\
\tau_\omega &\sim Gamma(a_{\tau_\omega 10}, a_{\tau_\omega 20}), \quad \tau_\xi \sim Gamma(a_{\tau_\xi 10}, a_{\tau_\xi 20}),
\end{aligned}$$

where $Gamma$ denotes the Gamma distribution.

For the proposal distributions, distributions are selected that can impose the support conditions described so far. First, for the unconstrained parameters, β_d and β_w, the normal proposal distributions are used. Second, for τ_ξ and τ_ω, the proposal distributions are set as log-normal distributions to guarantee positive values. Third, a truncated normal proposal is incorporated for γ to satisfy the support conditions (A.4) and (A.5). Finally, the truncated log-normal proposal distribution is used for α, which must be positive and satisfy the support condition.

A.2.3. Hyperparameters and primitives for the empirical study

The empirical analysis in Section 5 adopts the following hyperparameters and primitives. The hyperparameters are set as follows:

$$\begin{aligned}
\beta_d &\sim N(\mathbf{0}, 1000I), \quad \beta_s \sim N(\mathbf{0}, 1000I), \\
\gamma &\sim N(\mathbf{0}, 10I), \quad \alpha \sim N(1, 10), \\
\tau_\omega &\sim Gamma(3, 10), \quad \tau_\xi \sim Gamma(3, 10).
\end{aligned}$$

For the normal proposal distributions, their variances are adjusted to have modest rates of acceptance in the MH algorithm. Specifically, the acceptance rates for parameters are located within the range 0.29 to 0.6.

For the Polya tree mixture, primitives are chosen as $c = 10$ and $J = 5$. In addition, several alternative values are adopted to check the robustness of this choice. Here, 1, 100, and $1,000$ are incorporated for c. Of these values, $c = 100$ and $1,000$ yield posterior samples similar to the primary result. The value of $c = 1$, which is used in the studies by Hanson and Johnson (2002) and Hanson (2006), yields similar posterior means for the parameters, but shows a slow convergence. Additionally, $J = 8$ is incorporated as well. This value is recommended by Hanson and Johnson (2002) as a rule of thumb,

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 the econometric model in this study relative to the previous statistical papers.

B. Details of prediction of the monthly fee

This appendix describes the method of predicting the monthly fees under the counterfactual to complement Section 4.2. To predict p_{h_m}, the following predictive mean is considered:

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

A dual-loop Monte Carlo integral is conducted to calculate the double integral numerically. Furthermore, p_{h_m} must be integrated on the marginal distribution, which corresponds to the reduced form. As a result 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 representing the iterations for the 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 θ. The random numbers θ^l, $l = 1, 2, ..., L$, are generated from the posterior distribution of $\theta|$Data. The posterior samples of the MCMC estimation can be adopted as the random numbers in this step.

The inner loop implements the integral for p_{h_m}. Given θ^l, $p_{h_m}^{rl}$, $r = 1, 2, ..., R$, are generated from the distribution of $p_{h_m}^{lr} | f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \theta^l, \text{Data}$. To conduct the numerical solution, the MCMC sampling is implemented for the inner loop, which requires closed forms for the conditional predictive distributions for $q_{h_m}^{new} | \mathbf{p}_m^{new}, \mathbf{q}_{(-h_m)}^{new}, f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \theta, \text{Data}$ and $p_{h_m}^{new} | \mathbf{q}_m^{new}, \mathbf{p}_{(-h_m)}^{new}, f_{h_m}^{new} = 0, T_{h_m}^{new} = 0, \theta, \text{Data}$, for $h_m = 1, 2, ..., H_m$. Therefore, the random samples are iteratively drawn from these conditional distributions, given the previous draws. After appropriate burn-in periods, R', we have random draws from the marginal predictive distributions, which can serve as p_{new}^{rl}.

The conditional predictive distributions are derived as follows. First, in a similar manner to the derivation of the likelihood function, the joint predictive density for the dependent variables is obtained using a change of variable as

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

where

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

and the support conditions yield

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

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

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

where

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

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

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

where

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

After running the outer and inner loops, the double integral is approximated 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 can be conducted separately for each market. Owing to this separability, the empirical analysis focuses a specific market.

C. Estimated posterior densities and sample paths

Figures 5, 6, 7 and 8 are here

D. Tables and Figures

Prefecture	# Homes	# Elders	Prefecture	# Homes	# 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 1: Numbers of homes and elders in prefectures, excluding prefectures 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	$\mathbf{x}$, $\mathbf{w}$	2.523	0.476
Years since opening	$\mathbf{x}$, $\mathbf{w}$	7.378	5.986
Chain dummy	$\mathbf{x}$, $\mathbf{w}$	0.315	0.465
Occupancy rate	$\mathbf{x}$	0.914	0.146
Local average rent	$\mathbf{w}$	6051	1813
Local average wage (1,000 yen)	$\mathbf{w}$	1025	120
Sample size	H	1265	

Table 2: Descriptive statistics

Variable		Mean	95% Interval	IF
β_d	Constant	-4.311***(0.238)	[-4.772, -3.839]	631.52
	# Residents per worker	-0.675***(0.057)	[-0.793, -0.570]	444.66
	Chain dummy	0.379***(0.053)	[0.278, 0.485]	38.971
	Years since opening	0.018(0.030)	[-0.043, 0.073]	24.017
	α_F	0.038(0.032)	[-0.025, 0.102]	67.079
	α_f	0.001(0.003)	[-0.005, 0.008]	85.607
	Occupancy rate†	0.445***(0.165)	[0.123, 0.777]	339.17
β_s	Constant	2.914***(0.068)	[2.765, 3.044]	1414
	# Residents per worker	-0.409***(0.028)	[-0.459, -0.345]	1410.8
	Chain dummy	0.213***(0.025)	[0.155, 0.255]	947.94
	Years since opening	-0.094***(0.015)	[-0.124, -0.069]	749.74
	Local average rent†	0.191***(0.016)	[0.165, 0.228]	716.26
	Local average wage†	-0.010(0.016)	[-0.049, 0.012]	682.02
α		0.139***(0.001)	[0.137, 0.141]	260.15
γ	γ_0	0.282(2.986)	[-5.903, 5.717]	17.022
	γ_{T1}	-2.460(2.222)	[-7.365, 1.297]	24.909
	γ_{T2}	-2.196(2.324)	[-7.266, 1.549]	14.281
	γ_{f1}	-0.767(2.977)	[-6.735, 4.883]	16.901
	γ_{f2}	-2.488(1.995)	[-7.191, 0.174]	13.609
	γ_{fT}	-0.512(3.132)	[-6.673, 5.591]	8.3553
τ_ω		2.924***(0.342)	[2.262, 3.408]	619.66
τ_ξ		0.925***(0.080)	[0.772, 1.084]	50.701
	Sample size			1265

Table 3: Estimation result for real data

Standard deviations in parentheses. $***$, $**$ and $*$ indicate that 99%, 95% and 90% credible intervals do not include zero, respectively. Variables with † are exclusion restrictions.

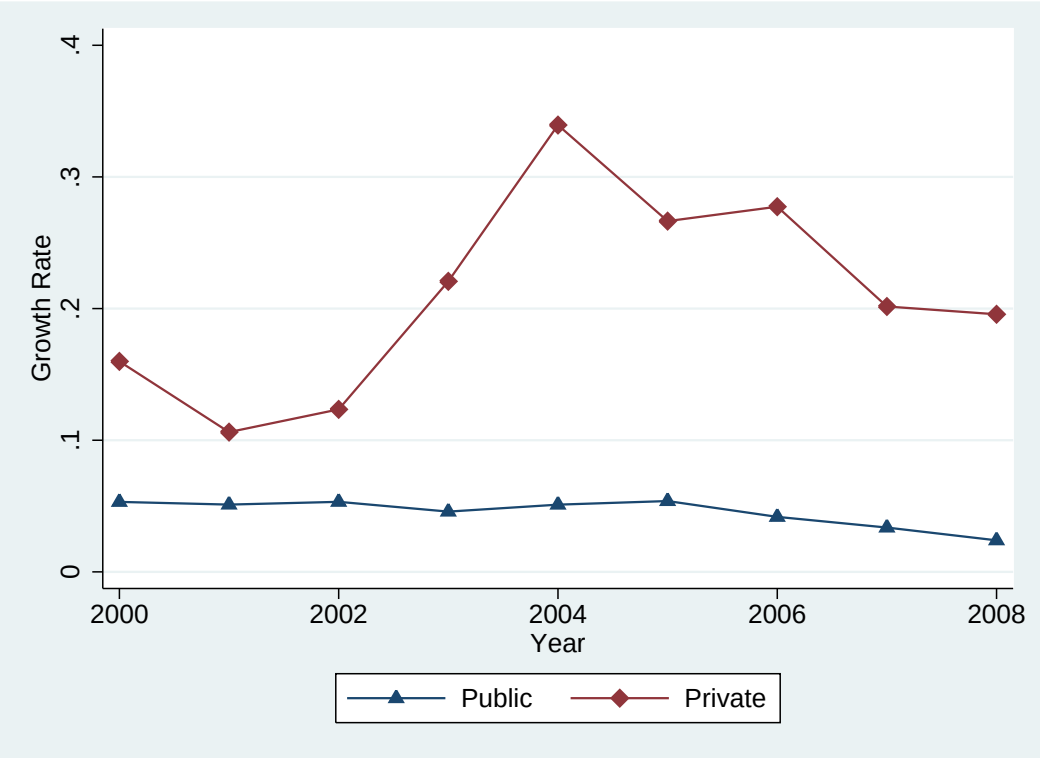

Figure 1: Annual growth rates of capacities of public and private homes

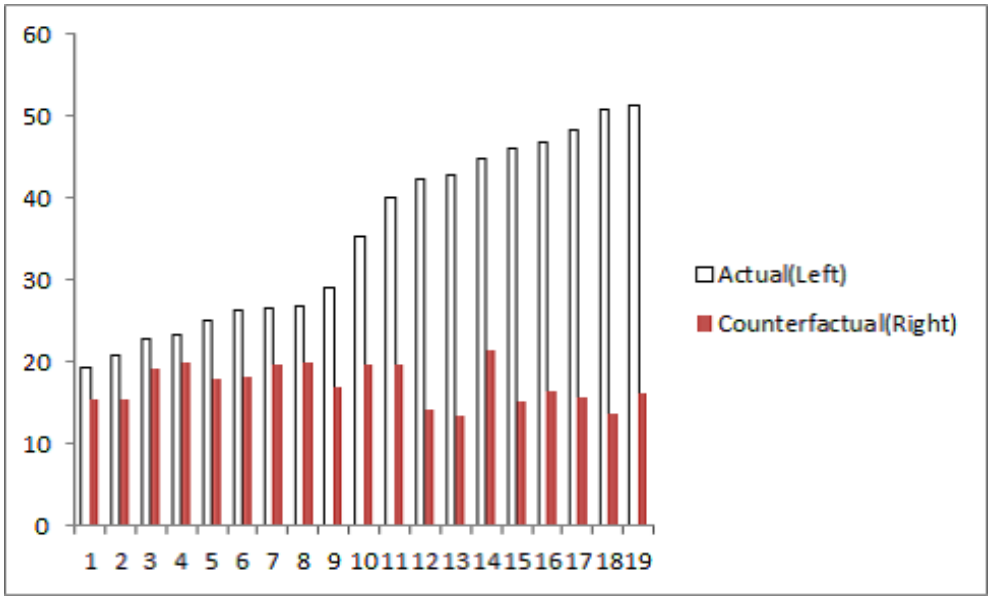

Figure 2: Prediction result for short-lived consumers ($\tau_i \leq T_h$):
X axis indexes homes and Y axis measures monthly fees in 10,000 yen

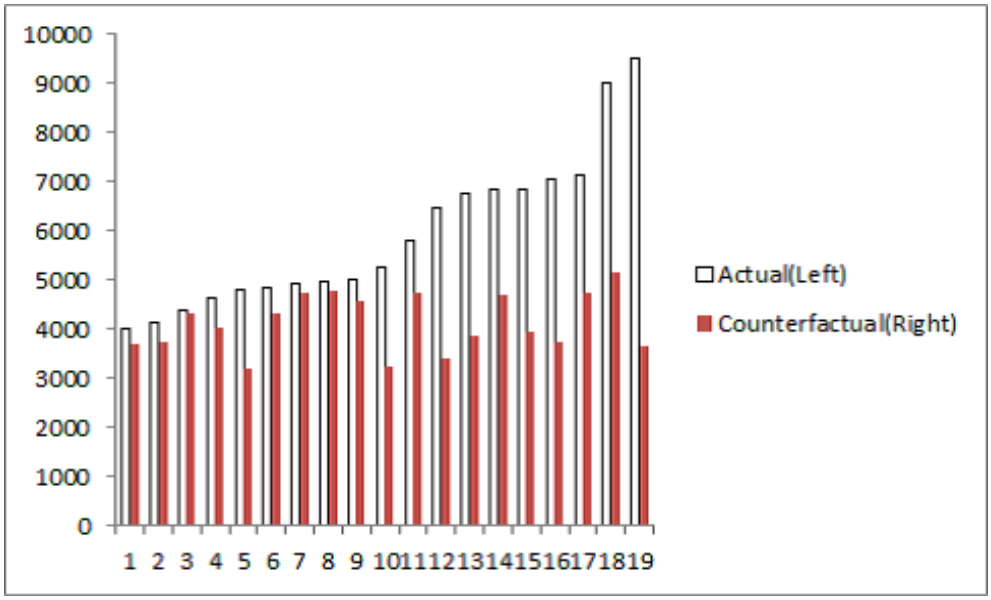

Figure 3: Prediction results for long-lived consumers: 20 year lifetime ($\tau_i > T_h$): $\tau_i = 240$
X axis indexes homes and Y axis measures lifetime payments in 10,000 yen

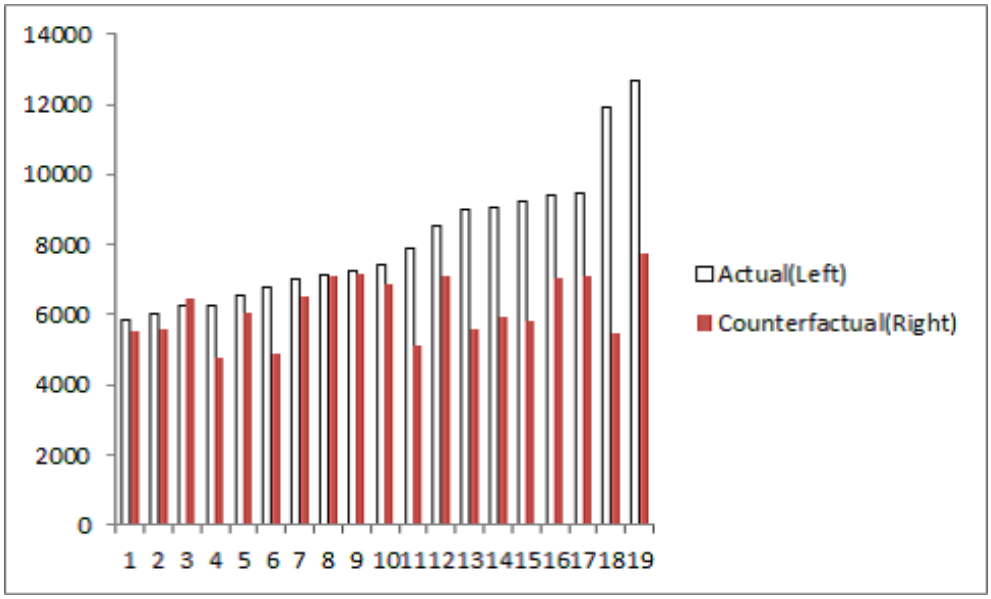

Figure 4: Prediction results for long-lived consumers: 30 year lifetime ($\tau_i > T_h$): $\tau_i = 360$
X axis indexes homes and Y axis measures lifetime payments in 10,000 yen

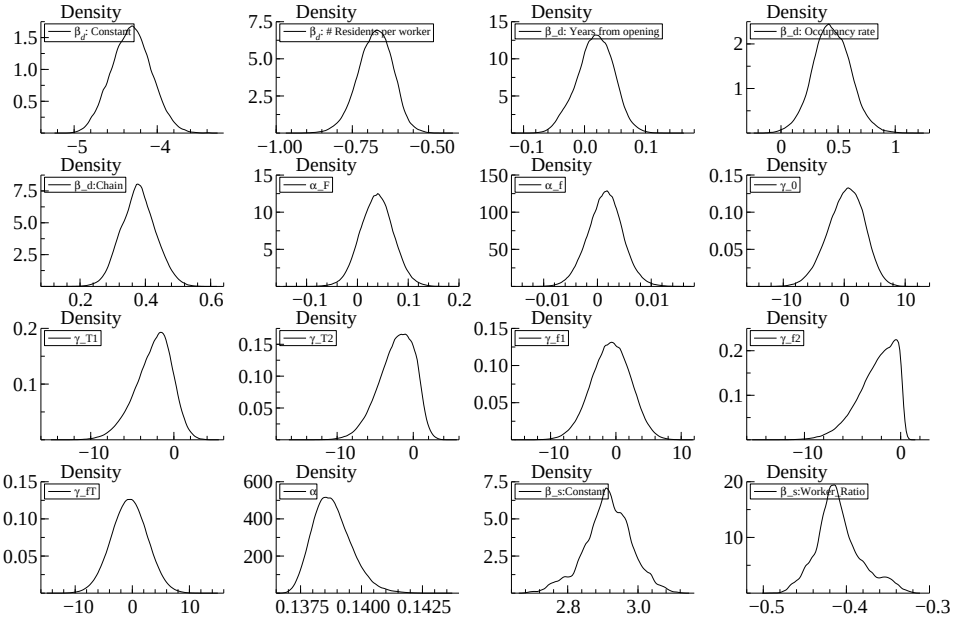

Figure 5: Posterior densities for the MCMC sampler, 1

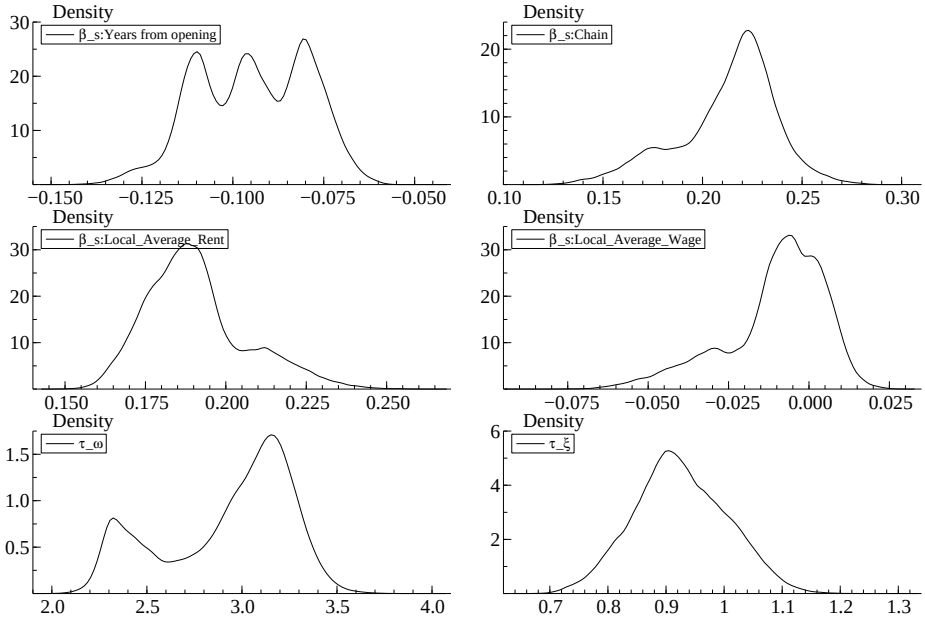

Figure 6: Posterior densities for the MCMC sampler, 2

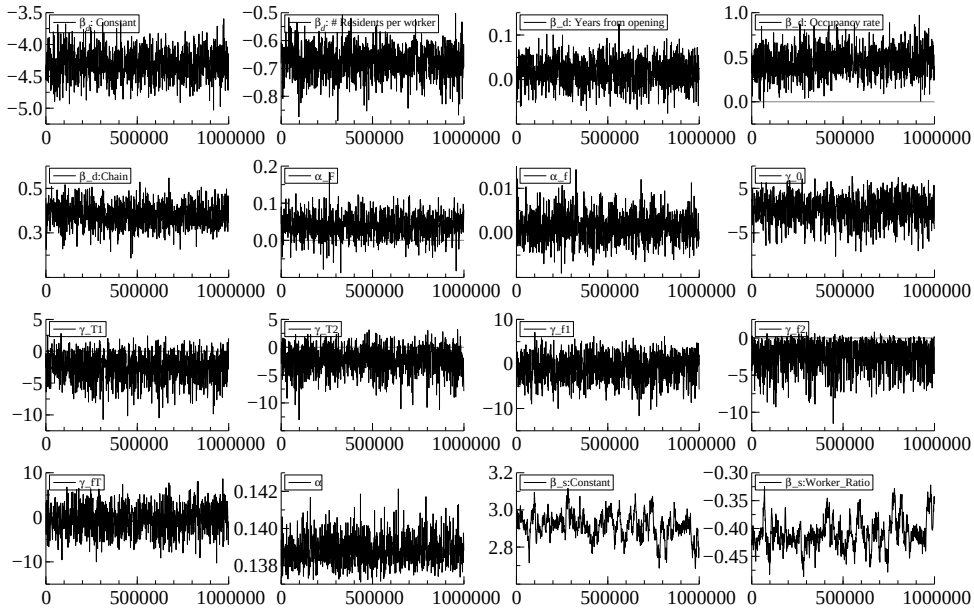

Figure 7: Sample paths for the MCMC sampler, 1

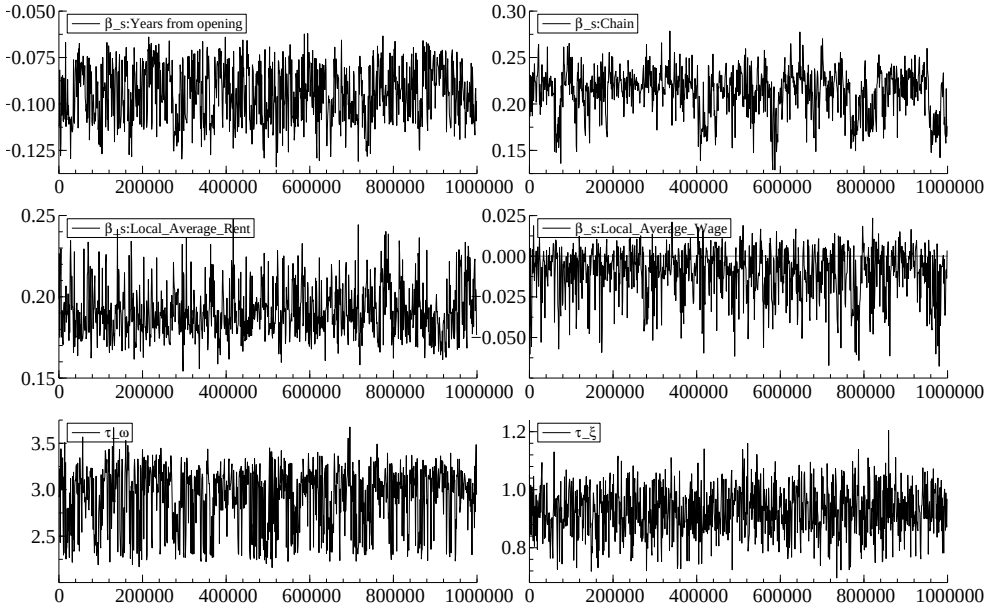

Figure 8: Sample paths for the MCMC sampler, 2