Bank of Japan Working Paper Series

Price Setting Behavior and Hazard Functions:

Evidence from Japanese CPI Micro Data

Daisuke Ikeda*

Shinichi Nishioka**

shinichi.nishioka@boj.or.jp

No.07-E-19

August 2007

Bank of Japan

2-1-1 Nihonbashi Hongoku-cho, Chuo-ku, Tokyo 103-8660

* Research and Statistics Department

** Research and Statistics Department

Bank of Japan Working Paper Series No.07-E-19
August 2007

Price Setting Behavior and Hazard Functions:
Evidence from Japanese CPI Micro Data[*]

Daisuke Ikeda[†] and Shinichi Nishioka[‡]

Abstract

We explore the micro price setting behavior using micro data of the consumer price index (CPI) in Japan. We estimate price hazard functions by the finite mixture model with multiple spells allowing the heterogeneity among price setters. Our estimation supports these striking findings. First, unlike much literature, our estimation finds increasing hazard functions and no decreasing hazard function. The estimated hazard functions are classified mainly into four groups: i) the flexible group; ii) the Calvo pricing group with low frequencies; iii) the Taylor pricing group with regular price changes; and iv) the increasing hazard group. Second, a decreasing empirical hazard function, observed in many countries as well as in Japan, is reproduced by estimated hazard functions of those four groups. Third, the increasing hazard group is likely to follow the time-dependent pricing rather than the state-dependent pricing in our sample period 2000-2004.

Keywords: Hazard function, Heterogeneity, Finite mixture model, State-dependent pricing, Time-dependent pricing.

[†] Research and Statistics Department
[‡] Research and Statistics Department *e*-mail: shinichi.nishioka@boj.or.jp

[*] We thank participants in the Research Center for Price Dynamics conference "Inflation Dynamics in Japan, US, and EU" and seminar participants at Hitotsubashi University, and staff members of the Bank of Japan. Any remaining errors are the sole responsibility of our own. The views expressed herein are those of the authors and should not be interpreted as those of the Bank of Japan.

1 Introduction

Price setting behavior is one of the most challenging themes in monetary policy analysis. And it is a starting point in modern micro-founded monetary models[1]. Different price setting behavior carries a different monetary policy implication in monetary models. Albeit many kinds of price setting behavior are developed and used in monetary policy analyses, it remains a controversy over what kinds of price setting behavior are empirically observed.

Motivated by the needs for understanding the actual price setting behavior, this paper explores the micro price setting behavior by estimating price hazard functions.

A price hazard function, defined as a conditional probability of price changes in terms of time, is a key concept in understanding price setting behavior. It is closely related to micro-founded models. For instance, the Calvo (1983)[2] and the Taylor (1999)[3] models have constant hazard functions. The probability of price changes depends only on time regardless of changes in state variables. Hence, the Calvo and the Taylor models are classified as the time-dependent pricing models. On the other hand, the Dotsey, et al. (1999) model has an increasing hazard function which changes depending on state variables. It is one of the most popular state-dependent models in a general equilibrium framework. In addition to those models, Mash (2004) and Coenen, et al. (2006) generalize the Calvo model and allow a hazard function to have any functional forms including decreasing and increasing hazards.

Our estimation of hazard functions supports three striking findings. First, unlike much literature, our estimation finds increasing hazard functions and no decreasing hazard function. Taking the heterogeneity among price setters into account, our estimation identifies eleven types for goods and six types for services. Those types are classified mainly into four groups: i) the flexible group; ii) the Calvo pricing group with low frequencies; iii) the Taylor pricing group; and iv) the increasing hazard group.

Second, a decreasing empirical hazard function, observed in many countries as well as in Japan[4], results from the aggregation bias. The decreasing hazard function of prices has been thought of as a puzzle because it is implausible that the longer the price setters fix

[1] See Woodford (2003) for the comprehensive analysis of monetary policy.

[2] The Calvo model is adopted in many monetary macro models which have been making exciting progress recently. See Christiano, et al. (2005) as a pioneer of monetary macro models.

[3] Taylor (1980) originally developed his price setting model in terms of nominal wage-setting behavior.

[4] See Dhyne, et al. (2006) for the evidence of CPI micro data in European countries reported by the Inflation Persistent Network (IPN) of the European Central Bank. Also, see Klenow and Kryvtsov (2005) for the United States and Higo and Saita (2007) for Japan. For the extensive survey of micro data including producer prices as well as consumer prices, see Alvarez, et al. (2006).

their prices, the less opportunity to change their prices they have. However, an empirical hazard function could be estimated to be decreasing if it consists of several types of hazard functions. Alvarez, et al. (2005) clearly showed that a decreasing empirical hazard function can be reproduced by several Calvo pricing groups with a constant hazard function. Following Alvarez, et al (2005), we show that the empirical decreasing hazard function in Japan is reproduced by our estimated several different hazard functions.

Third, we apply the method developed by Klenow and Kryvtsov (2005) to our sample data and show that the increasing hazard group is more likely to be the time-dependent rather than the state-dependent. Generally, increasing hazard functions do not necessarily mean state-dependent pricing. Our result implies that even though items are classified as the increasing hazard group, they are appropriate to the time-dependent models such as Mash (2004) or Coenen, et al (2006) models rather than the state-dependent model of Dotsey, et al. (1999).

As far as we know, this paper is the first to show the existence of increasing hazards clearly. Our findings sharply contrast with Dhyne, et al. (2006) and, Higo and Saita (2007) who reported that the empirical hazard functions of the CPI micro data in Europe and in Japan are downward sloping respectively. Our estimation shows no decreasing hazard. The sharp contrast arises from the difference between empirical methods not from the difference between data sources. Our estimated empirical hazard function is also downward sloping as Dhyne, et al. (2006) and, Higo and Saita (2007). However, our estimated heterogeneous parametric hazard functions are not downward sloping.

The empirical methods employed so far to estimate hazard functions are mainly classified into the three categories: i) the non-parametric estimation that calculates an empirical hazard function; ii) the semi-parametric estimation with parametric specification of heterogeneity; and iii) the parametric estimation with nonparametric specification of heterogeneity. Much literature uses the first method. Klenow and Kryvtsov (2005) and Higo and Saita (2007) estimated empirical hazard functions and stated that they are downward sloping. Nakamura and Steinsson (2006), using the second method following Meyer (1990), reported that a hazard function is downward sloping for the first few months and then becomes flat for the longer period. Similarly, Fougere, et al. (2005) identified the heterogeneity among items by some state variables and showed that a hazard function is rather flat. Alvarez, et al. (2005), using the third method, assumed several exponential hazard functions and estimated them in order to take the heterogeneity among price setters into account[5]. Due to that assumption, Alvarez, et al. (2005) estimated several

[5] In addition to those estimations, Aucremanne and Dhyne (2005) used a logistic normal regression to estimate hazard functions.

hazard functions with merely two types: the Calvo pricing type and the Taylor pricing type.

Our estimation is categorized as the third method, the parametric estimation with nonparametric specification of heterogeneity, and it is unique in two aspects. First, we extend the finite mixture model used by Alvarez, et al. (2005) in that the Weibull distribution is assumed as hazard functions. Alvarez, et al. (2005) assumed only constant hazard functions and excluded increasing or decreasing hazard functions. On the other hand, our assumption of Weibull distribution enables us to estimate increasing or decreasing hazard functions, in addition to flat hazard functions. Although Alvarez, et al. (2005) is skeptical about increasing hazard functions, we show that increasing hazard functions are estimated significantly.

Second, we use multiple spells of price changes in estimating hazard functions. Generally, in a given sample period, a sequence of prices of an item consists of multiple spells since price changes occur several times. Our sample consists of all the spells in each sequence of prices, while the sample in Alvarez, et al. (2005) consists of one spell selected randomly from a sequence of spells for each item. The random sampling makes each item be characterized by only one randomly selected spell and that may lead to a bias in estimating hazard functions[6]. On the other hand, our method estimates hazard functions accurately because it makes use of all the information in each item. In addition, we use the finite mixture model for our estimation in order to eliminate the aggregation bias. Our estimated heterogeneous hazard functions should represent micro-level price setting behavior, given the assumption of the Weibull distribution.

The rest of the paper is organized as follows. Section 2 reviews the relationship between price setting behavior and hazard functions. Section 3 introduces some definitions related to hazard functions and calculate empirical hazard functions as a benchmark. Section 4 explains the finite mixture model with multiple spells in estimating heterogeneous hazard functions, and Section 5 presents our estimation results. Section 6 discusses some implications derived from the estimation results. Finally Section 7 concludes the paper.

2 Price Setting Behavior and Hazard Functions

In general, micro-founded price setting models are classified into two models: the

[6] If true hazard functions are exponential, random sampling does not have serious problems with a large sample. Since Alvarez, et al. (2005) assume exponential functions, their methodology in estimating hazard functions do not have problems. However, if one assumes non-flat functions like those of ours, one may lose important information with random sampling.

time-dependent model and the state-dependent model. In the time-dependent model, the conditional probability of price changes depends only on the period in which a price is fixed. Then, a hazard function in the time-dependent model has a certain constant shape in terms of the duration of prices. For instance, the Calvo (1983) model has a flat hazard function. It assumes that the opportunity of a price change follows the Poisson process. The assumption means that a price setter has an opportunity of price changes with a constant probability in every period. It is well known that the New Keynesian Phillips curve is derived from the Calvo model with monopolistic competition. Also, the Taylor (1999) model has a constant hazard function which takes a hundred percent in certain durations and zero percent otherwise. It assumes that price setters change their prices only at the beginning of the contract and don't change their prices within the period of contract. Then, its hazard rate takes the value of unity at the beginning of the contract and zero otherwise.

In addition to the Calvo and the Taylor models, Mash (2004) and Coenen, et al. (2006) generalize the Calvo model[7]. They assign different probabilities to different periods, allowing the hazard function to have any functional forms including decreasing and increasing hazards. They show that the Phillips curves derived from their models depend not only on the current GDP gaps and the expected inflation rate in the next period as the New Keynesian Phillips curve, but also on the expected inflation rates in some past and some future periods.

In the state-dependent model, the conditional probability of price changes depends on state variables such as relative prices and inflation rates. Then, a hazard function in the state-dependent model may change its shape in response to real or monetary shocks in a transition, while it has a certain constant shape in a steady state. For instance, Dotsey, et al. (1999) developed the state-dependent pricing model extending the basic menu cost model by Blanchard and Kiyotaki (1987). They assumed that the menu cost follows the random process and varies across price setters. In that case, they showed that a hazard function depends on inflation rates and the distribution of the random process. In addition, they showed that the shape of a hazard function is increasing in a steady state. In a steady state, the longer the price remains fixed, the more the relative price deviates from the optimal relative price due to accumulated productivity shocks. Then, the conditional probability of

[7] Coenen, et al. (2006) estimated hazard functions using indirect inference methods. They set the maximal time interval of fixed prices as four periods to ease the burden of estimation. Let ρ_t be the conditional probability of price changes at time t. Their estimation results showed $[\rho_1, \rho_2, \rho_3, \rho_4] = [0.43, 0.00, 0.14, 1.00]$ in the United States and $[\rho_1, \rho_2, \rho_3, \rho_4] = [0.55, 0.38, 0.21, 1.00]$ in Germany. The hazard functions are regarded as almost decreasing except in the fourth period. The conditional probability of the fourth period must be unity since the maximal time interval of fixed prices is assumed to be four periods.

price changes rises up as a price remains to be fixed.

Bakhshi, et al. (2004) showed that the Phillips curve derived from the Dotsey, et al. (1999) model has the complicated form. It depends not only on the current GDP gaps and the expected inflation rate in the next period, but also on the expected inflation rates in some past and some future periods. Thus, regardless of the time-dependent or state-dependent models, the Phillips curve has the complicated form if a hazard function is not flat as the Calvo model.

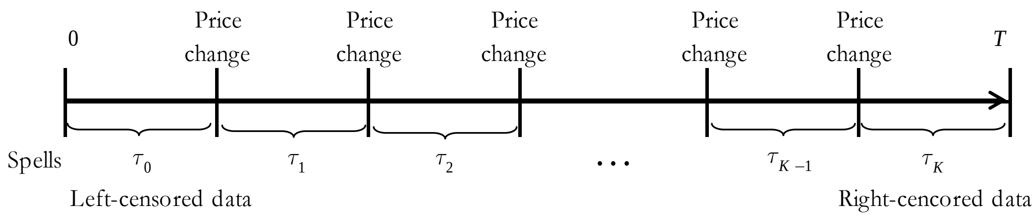
3 Definitions and Empirical Hazard Functions

3.1. Definitions of Hazard Functions

Hazard functions represent the distribution of the length of time that elapses from the beginning of some events until its end. For their usefulness, hazard functions are often used in survival analysis of products, firm's default analysis and biological analysis. In our paper, the hazard function depends on the duration of prices, denoted by τ. The hazard function produces a probability of a price change conditional on the event that a price has fixed for the previous $\tau - 1$ periods. Its output is called the hazard rate. For instance, the hazard rate with $\tau = 5$ means a probability of a price change in period five conditional on the event that a price has fixed for the previous four periods.

We give some definitions of the hazard function. Figure 1 shows a sequence of price changes. The time zero and the time T show the beginning and the end of the sample period respectively. The term *spell* means the duration of prices, that is, the length of time in which the price is fixed. The first spell τ_0 and the final spell τ_K in figure 1 are called the *left-censored data* and the *right-censored data* respectively. The duration of the left-censored data is uncertain since the period of the last price change is unknown. Also the duration of the right-censored data is uncertain since the next price change is unknown. The left-censored data is usually excluded from sample data. But the right-censored data is included as sample data since it is used for calculating the survival probabilities of the final spell. The sequence of spells is called the *trajectory*. In figure 1, the trajectory of prices is $\boldsymbol{\tau} = (\tau_1, \tau_2, ..., \tau_K)$. The left-censored data is excluded from the trajectory.

Figure 1: Sequence of price changes during a sample period



3.2. Empirical Hazard Functions

We calculate empirical hazard functions using the Japanese CPI micro data before estimating the parametric hazard functions. Empirical hazard functions are calculated by the Kaplan-Meier product limit estimator and are widely used because of their simple calculation. Dhyne, et al. (2006), Klenow and Kryvtsov (2005) and Higo and Saita (2007) reported that empirical hazard functions of price changes are downward sloping. Though empirical hazard functions tend to have the aggregation bias, they play an important role as a benchmark of hazard functions.

Table 1 shows the basic statistics used for calculating empirical hazard functions. The price data used here are those of the Ministry of Internal Affairs and Communications' "Retail Price Survey" which is the individual item data of the Consumer Price Index in Japan[8]. The sample period is from January 2000 to December 2004. The output of empirical hazard function, the hazard rate p_t, is obtained from the Kaplan-Meier product limit estimator, which is expressed as:

$$p_t = \frac{d_t}{r_t}$$

where d_t is the number of data whose prices were revised in period t and r_t is the risk set in period t. The risk set r_t is the number of spells at risk in period t. Spells are at risk if they have not yet revised. In other words, the risk set indicates the number of spells in which prices were not revised until period $t-1$.[9] For instance, the risk set in period $t=5$ is the number of spells whose lengths of time are at least 4 periods.[10] Let m_t be the number of data of the right-censored data in period t. Then the risk set in period $t+1$ follows the law of the motion:

$$r_{t+1} = r_t - d_t - m_t$$

The equation indicates that the risk set in period t falls under one of the following three data: (i) d_t, the data whose prices are revised in period t; (ii) m_t, the right-censored data in period t; and (iii) r_{t+1}, the data whose prices are not revised and are not

---

[8] The price data of each item is not purely micro but semi-aggregated. For instance, the price data of major cities is the average of 5 prices of the same item. See Higo and Saita (2007) for definitions of the data in detail.

[9] The empirical hazard in figure 2 does not account for the weights of the CPI. It simply counts the number of samples, and it also excludes left-censored data and samples without a price change. Following Alvarez, et al. (2005), samples are compiled from one spell chosen randomly from each trajectory in each item. Although it is possible to draw empirical hazards using all spells, items with high probability of price changes include a large number of spells with a short duration in their trajectories. Consequently, since the number of samples of d_t with a small t increases substantially, this in turn leads to an upward bias in the hazard. This problem is avoided when one spell per item is used.

[10] The risk set of the 1st term represents the number of all spells included in the samples.

right-censored in period t, which becomes the risk set in period $t+1$.

The law of the motion of the risk set has a distinctive feature. The items with high frequency of price changes are excluded from the risk set in a short period, while those with low frequency of price changes continues to be included in the risk set in a long period. Then, empirical hazard functions have an aggregation bias to be decreasing. And hazard rates are high in a short period and low in a long period. In fact, figure 2 shows actual empirical hazard functions, and confirms a clear decreasing function for goods.

Alvarez, et al. (2005) pointed out the aggregation bias of decreasing empirical hazard functions. They showed that a decreasing empirical hazard function is reproduced by several flat hazard functions, that is, the Calvo type pricing models. Apart from that finding, few papers resolved the decreasing hazard puzzle. And few papers found increasing hazard functions.

From the following section, we extend the model of Alvarez, et al. (2005) and estimate heterogeneous hazard functions. We estimate hazard functions taking into the heterogeneity not only of flat types but also of various patterns including increasing hazards and the Taylor pricing hazards. The Taylor pricing hazards have a high probability of price changes in certain periods. We include the Taylor pricing hazards because services mainly follow the Taylor models as indicated in figure 2.

Table 1: Basic data used for calculating empirical hazard functions

(i) Goods

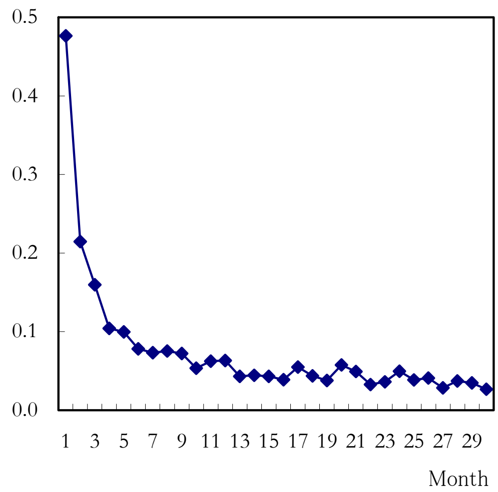
t	Risk Set r_t	Revision d_t	Right-censored m_t	Hazard Rate d_t/r_t
1	21,641	10,310	194	0.476
2	11,137	2,389	125	0.215
3	8,623	1,377	139	0.160
4	7,107	739	136	0.104
5	6,232	621	69	0.100
6	5,542	432	100	0.078
7	5,010	366	96	0.073
8	4,548	342	347	0.075
9	3,859	279	56	0.072
10	3,524	188	142	0.053
11	3,194	200	88	0.063
12	2,906	184	29	0.063
13	2,693	116	32	0.043
14	2,545	113	44	0.044
15	2,388	103	29	0.043
16	2,256	88	56	0.039
17	2,112	116	28	0.055
18	1,968	86	23	0.044
19	1,859	70	112	0.038
20	1,677	97	40	0.058
21	1,540	76	22	0.049
22	1,442	47	33	0.033
23	1,362	49	39	0.036
24	1,274	63	14	0.050
25	1,197	46	25	0.038
26	1,126	46	30	0.041
27	1,050	30	27	0.029
28	993	37	209	0.037
29	747	26	10	0.035
30	711	19	15	0.027

(ii) Services

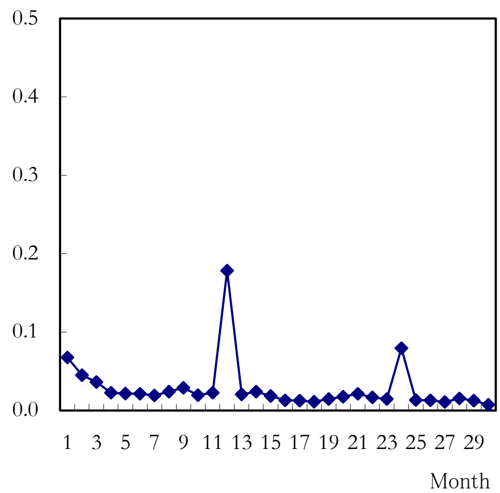
t	Risk Set r_t	Revision d_t	Right-censored m_t	Hazard Rate d_t/r_t
1	4,536	307	40	0.068
2	4,189	188	36	0.045
3	3,965	143	53	0.036
4	3,769	86	20	0.023
5	3,663	79	28	0.022
6	3,556	76	38	0.021
7	3,442	66	53	0.019
8	3,323	80	426	0.024
9	2,817	81	27	0.029
10	2,709	53	33	0.020
11	2,623	59	47	0.023
12	2,517	449	12	0.178
13	2,056	42	15	0.020
14	1,999	48	12	0.024
15	1,939	36	21	0.019
16	1,882	24	50	0.013
17	1,808	23	18	0.013
18	1,767	20	19	0.011
19	1,728	25	28	0.015
20	1,675	29	102	0.017
21	1,544	33	21	0.021
22	1,490	25	16	0.017
23	1,449	21	45	0.015
24	1,383	110	8	0.080
25	1,265	17	8	0.013
26	1,240	16	25	0.013
27	1,199	13	18	0.011
28	1,168	18	35	0.015
29	1,115	14	12	0.013
30	1,089	8	14	0.007

Figure 2: Empirical hazard functions

(i) Goods



(ii) Services



4 Estimation Method Using the Finite Mixture Model

Multiple types of hazard functions are estimated by the finite mixture model.[11] We outline our estimation method in this section.

4.1 Hazard Functions

We focus on shapes of hazard functions. Are hazard functions increasing or decreasing ? To answer the question we assume the Weibull distribution as the distribution function for price changes. The Weibull distribution is commonly used to statistically express a degradation phenomenon or life expectancy. Also it is widely applied to survival analysis in economics. Its hazard function is monotonously increasing or decreasing and cannot be hump-shaped. Nevertheless, the Weibull distribution is useful in two respects. First increasing, flat, and decreasing hazards are expressed by only two parameters. Second the simple formula reduces a computational burden.

The distribution function of the Weibull distribution $F(\tau)$ is defined as:

$$F(\tau) = 1 - \exp\left(-\exp(\lambda)\tau^{\theta}\right) \tag{1}$$

$F(\tau)$ is interpreted as the probability of price changes within period τ. The probability density function $f(\tau)$ and the survival probability function $S(\tau)$ are derived from the equation (1) as follows:

$$f(\tau) = \frac{dF(\tau)}{d\tau} = \exp(\lambda)\theta\tau^{\theta-1}\exp\left(-\exp(\lambda)\tau^{\theta}\right) \tag{2}$$

$$S(\tau) = 1 - F(\tau) = \exp\left(-\exp\left(\lambda\tau^{\theta}\right)\right) \tag{3}$$

The probability density $f(\tau)$ expresses the probability of price changes in period τ and the survival probability $S(\tau)$ expresses the probability of which prices are not revised during τ periods. Furthermore, the hazard function $h(\tau)$ is defined as follows:

$$h(\tau) = \frac{f(\tau)}{S(\tau)} = \exp(\lambda)\theta\tau^{\theta-1} \tag{4}$$

The hazard function $h(\tau)$ produces the probability of price changes in period τ under the condition that prices are not revised during the previous $\tau-1$ periods. In the equation (4) λ is a parameter measuring the level of the probability of price changes and

---

11 See Cameron and Trivedi (2005) for further explanations on the estimation method.

θ is a parameter measuring the slope of the hazard function. The parameter θ and the hazard function $h(\tau)$ have the following relationship:

$$\begin{aligned}
&\text{If } \theta > 1: \ h(\tau) \text{ is increasing} \\
&\text{If } \theta = 1: \ h(\tau) \text{ is constant} \\
&\text{If } \theta < 1: \ h(\tau) \text{ is decreasing}
\end{aligned}$$

Alvarez, et al. (2005) treated the case $\theta = 1$ only. We primarily focus on whether θ is above or below 1.

We also consider the Taylor pricing hazards in addition to flat or increasing hazards. We add dummy variables to the Weibull hazard function to make the function have spikes at the twelfth and twenty-forth month as follows.

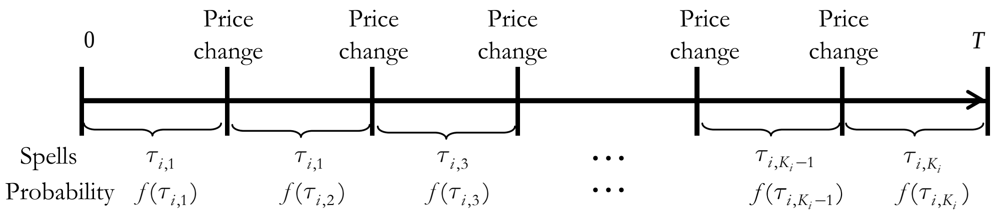
$$h_j(\tau) = \exp\left(\lambda_{0,j} + \lambda_{12,j} dummy_{12} + \lambda_{24,j} dummy_{24}\right) \theta_j \tau^{\theta_j - 1}$$

where $dummy_{12}$ and $dummy_{24}$ are dummy variables which take 1 only when $\tau = 12$ and 24, respectively.[12]

4.2 Likelihood Function

Figure 3 shows the developments of price changes of item i. The subscript i indicates the item's number. The left-censored data are excluded from estimation since the interval of price changes is unobserved. The $K-1$ price changes and one right-censored data are observed. The unconditional probability of price changes of the spell $\tau_{i,k}$ $(k = 1, 2, ..., K-1)$ is defined as $f(\tau_{i,k})$. Also, the survival probability of the right-censored data $\tau_{i,K}$ is defined as $S(\tau_{i,K})$.

Figure 3: Developments in price changes of item i



The joint density function of item i with trajectory $\boldsymbol{\tau}_i$ $(i = 1, 2, ..., N)$ is expressed as:[13]

---

12 Hereafter, we will discuss this paper based on formula (4) which does not include a dummy variable, but the essence of the discussion remains unchanged even with the dummy variable included.

13 In this paper, we examine a case with the right-censored data included. Without the right-censored

$$g\left(\boldsymbol{\tau}_{\mathrm{i}} \mid \theta, \lambda\right) = f\left(\tau_{i,1}\right) \cdots f\left(\tau_{i,K_i-1}\right) S\left(\tau_{i,K_i}\right)$$

The above equation points out our distinctive estimation method. We use trajectories, that is, all spells within a sample period.[14] In our estimation it is a trajectory that characterizes each item's pricing behavior. On the contrary, in the estimation of Alvarez, et al. (2005) it is a spell that characterizes each item's pricing behavior. Alvarez, et al. (2005) randomly chose one spell from each trajectory and formed a sample of their estimation. Then the joint density function in Alvarez, et al. (2005) is written as:

$$g\left(\boldsymbol{\tau}_{\mathrm{i}} \middle| \theta, \lambda\right) = \begin{cases} f\left(\tau_{i,k}\right) & \text{if } k \neq K_i \\ S(\tau_{i,k}) & \text{if } k = K_i \end{cases}$$

Since trajectories have more information than spells, the estimation using trajectories must reveal pricing behavior more accurately.

Taking the heterogeneity among items into consideration, we assume that J types of hazard functions exist and that individual items follow any one of those types. Also we assume that the jth type has a fraction $\pi_j (j = 1, 2, ..., J)$ of the total population. Then the logarithmic likelihood function is expressed as:

$$l(\boldsymbol{\theta}, \boldsymbol{\lambda} \mid \boldsymbol{\tau}) = \sum_{i=1}^{N} \log(\pi_1 g(\boldsymbol{\tau}_{\mathrm{i}} \mid \theta_1, \lambda_1) + ... + \pi_J g(\boldsymbol{\tau}_{\mathrm{i}} \mid \theta_J, \lambda_J))$$

where

$$\sum_{j=1}^{J} \pi_j = 1$$

Using the likelihood function we estimate the unknown parameters to obtain the J heterogeneous hazard functions and their weights. The unknown parameters are $\boldsymbol{\theta} = (\theta_1, ..., \theta_J)$, $\boldsymbol{\lambda} = (\lambda_1, ..., \lambda_J)$ and $\boldsymbol{\pi} = (\pi_1, ..., \pi_J)$.

4.3 EM Algorithm

We estimate the parameters of the hazard functions using the EM algorithm. First, the joint density function of trajectory $\boldsymbol{\tau}_{\mathrm{i}}$ is rewritten as follows:

---

data, the joint density function is $g\left(\boldsymbol{\tau}_{\mathrm{i}} \middle| \theta, \lambda\right) = f\left(\tau_{i,1}\right) \cdots f\left(\tau_{i,K_i}\right)$.

14 Mealli and Pudney (1996) estimated a more complex finite mixture model with multiple spells.

$$
\begin{aligned}
g\left(\boldsymbol{\tau}_{\mathrm{i}}\middle|\theta_j,\lambda_j\right) &= f_j\left(\tau_{i,1}\mid\theta_j,\lambda_j\right)\cdots f_j\left(\tau_{i,k_i-1}\mid\theta_j,\lambda_j\right)S_j\left(\tau_{i,k_i}\mid\theta_j,\lambda_j\right)\\
&= e^{\lambda_j}\theta_j\tau_{i,1}^{\theta_j-1}\exp\left(-e^{\lambda_j}\tau_{i,1}^{\theta_j}\right)\cdots e^{\lambda_j}\theta_j\tau_{i,k_i-1}^{\theta_j-1}\exp\left(-e^{\lambda_j}\tau_{i,k_i-1}^{\theta_j}\right)\cdot\exp\left(-e^{\lambda_j}\tau_{i,k_i}^{\theta_j}\right)\\
&= e^{(k_i-1)\lambda_j}\theta_j^{k_i-1}\left(\prod_{l=1}^{k_i-1}\tau_{i,l}^{\theta_j-1}\right)\exp\left(e^{\lambda_j}\sum_{l=1}^{k_i-1}\tau_{i,l}^{\theta_j}\right),\qquad i=1,\cdots,N,\ j=1,\cdots,J.
\end{aligned}
\tag{5}
$$

Let us introduce the J by 1 vector $\mathbf{d}_{\mathrm{i}}$ to deal with the unobservable types of item i. The vector $\mathbf{d}_{\mathrm{i}}$ is called the latent variable. When the item i belongs to the type j the latent variable vector $\mathbf{d}_{\mathrm{i}}$ takes unity for the jth element and zero for others. Assume that the trajectory $\boldsymbol{\tau}_{\mathrm{i}}$ conditional on $\mathbf{d}_{\mathrm{i}}$ is independently and identically distributed. Then the density is given by

$$
\left(\boldsymbol{\tau}_{\mathrm{i}}\middle|\mathbf{d}_{\mathrm{i}},\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\pi}\right)\sim\ \text{i.i.d.}\ \prod_{j=1}^{J}g\left(\boldsymbol{\tau}_{\mathrm{i}}\middle|\theta_j,\lambda_j\right)^{d_{ij}}
\tag{6}
$$

Also assume that the latent variable $\mathbf{d}_{\mathrm{i}}$ is independently and identically distributed with the multinomial distribution:

$$
(\mathbf{d}_{\mathrm{i}}\mid\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\pi})\ \sim\ \text{i.i.d.}\ \prod_{j=1}^{J}\pi_j^{d_{ij}}
\tag{7}
$$

The above two assumptions imply that the trajectory $\boldsymbol{\tau}_{\mathrm{i}}$ follows

$$
\left(\boldsymbol{\tau}_{\mathrm{i}}\middle|\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\pi}\right)\sim\ \text{i.i.d.}\ \prod_{j=1}^{J}\pi_j^{d_{ij}}g\left(\boldsymbol{\tau}_{\mathrm{i}}\middle|\theta_j,\lambda_j\right)^{d_{ij}}
\tag{8}
$$

From the density (8) the likelihood function is formed as follows.

$$
L\left(\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\pi}\middle|\boldsymbol{\tau},\mathbf{d}\right)=\prod_{i=1}^{N}\prod_{j=1}^{J}\pi_j^{d_{ij}}g\left(\boldsymbol{\tau}_{\mathrm{i}}\middle|\theta_j,\lambda_j\right)^{d_{ij}}
\tag{9}
$$

The log-likelihood is given by

$$
\begin{aligned}
l\left(\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{\pi}\middle|\boldsymbol{\tau},\mathbf{d}\right) &= \sum_{i=1}^{N}\sum_{j=1}^{J}d_{ij}\left[\log\left(g\left(\boldsymbol{\tau}_{\mathrm{i}}\middle|\theta_{j,}\lambda_j\right)\right)+\log\pi_j\right]\\
&= \sum_{i=1}^{N}\sum_{j=1}^{J}d_{ij}\left[(k_i-1)\lambda_j+(k_i-1)\log\theta_j+\left(\theta_j-1\right)\sum_{l=1}^{k_i-1}\log\tau_{i,l}-e^{\lambda_j}\sum_{l=1}^{k_i}\tau_{i,l}^{\theta_j}+\log\pi_j\right]
\end{aligned}
$$

The parameters $(\boldsymbol{\theta}\ \boldsymbol{\lambda}\ \boldsymbol{\pi})$ are estimated by these three steps. First, we calculate the expected log-likelihood. Given the guess of $(\boldsymbol{\theta}\ \boldsymbol{\lambda}\ \boldsymbol{\pi})$, for each item i and each type j, calculate the posterior probability that the trajectory $\boldsymbol{\tau}_{\mathrm{i}}$ belongs to type j as follows.

$$\hat{d}_{ij} = \frac{\pi_j g\left(\boldsymbol{\tau}_{\mathbf{i}} \middle| \theta_j, \lambda_j\right)}{\sum_{j=1}^{J} \pi_j g\left(\boldsymbol{\tau}_{\mathbf{i}} \middle| \theta_j, \lambda_j\right)} \tag{10}$$

The $\hat{d}_{ij}$ in (10) is used as the estimator of the expected value of d_{ij}. Substituting $\hat{d}_{ij}$ s into the log likelihood, we obtain the expected log-likelihood given by

$$El\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\pi} \middle| \boldsymbol{\tau}, \mathbf{d}\right) = \sum_{i=1}^{N} \sum_{j=1}^{J} \hat{d}_{ij} \left[\log g\left(\boldsymbol{\tau}_{\mathbf{i}} \middle| \theta_j, \lambda_j\right) + \log \pi_j\right] \tag{11}$$

This first step is called the Expectation step (E-step). Second, we maximize the expected log-likelihood. The maximization of (11) gives us the following first order conditions.

$$\begin{aligned} \frac{\partial El\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\pi} \middle| \boldsymbol{\tau}, \hat{\mathbf{d}}\right)}{\partial \theta_j} &= \sum_{i=1}^{N} \hat{d}_{ij} \left[\frac{k_i - 1}{\theta_j} + \sum_{l=1}^{k_i - 1} \log \tau_{i,l} - e^{\lambda_j} \sum_{l=1}^{k_i} \tau_{i,l}^{\theta_j} \log \tau_{i,l}\right] = 0 \\ \frac{\partial El\left(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\pi} \middle| \boldsymbol{\tau}, \hat{\mathbf{d}}\right)}{\partial \lambda_j} &= \sum_{i=1}^{N} \hat{d}_{ij} \left[k_i - 1 - e^{\lambda_j} \sum_{l=1}^{k_i} \tau_{i,l}^{\theta_j}\right] = 0 \end{aligned} \tag{12}$$

Solving the equations (12) we obtain the estimated values $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\lambda}}$. Similarly, the estimated value of $\boldsymbol{\pi}$ is obtained from the following first order condition.

$$\hat{\pi}_j = \frac{1}{N} \sum_{i=1}^{N} \hat{d}_{ij} \tag{13}$$

This second step is called the Maximization step (M-step). Third, if $(\hat{\boldsymbol{\theta}}\ \hat{\boldsymbol{\lambda}}\ \hat{\boldsymbol{\pi}})$ converges within a certain small enough criterion value, stop here. The set of converged parameters $(\hat{\boldsymbol{\theta}}\ \hat{\boldsymbol{\lambda}}\ \hat{\boldsymbol{\pi}})$ is the final set of estimated values. Otherwise, with the new guess of $(\boldsymbol{\theta}\ \boldsymbol{\lambda}\ \boldsymbol{\pi})$, return to the first step (E step) and repeat the process until $(\hat{\boldsymbol{\theta}}\ \hat{\boldsymbol{\lambda}}\ \hat{\boldsymbol{\pi}})$ converges. We give appropriate default values of $(\boldsymbol{\theta}\ \boldsymbol{\lambda}\ \boldsymbol{\pi})$ as the first guess and repeat the process until $(\hat{\boldsymbol{\theta}}\ \hat{\boldsymbol{\lambda}}\ \hat{\boldsymbol{\pi}})$ converges.

5 Estimation Results

In this section we report the data source and the estimation results.

5.1 Data

We use data of the "Retail Price Survey" of the Ministry of Internal Affairs and

Communications.[15] The *Retail Price Survey* is the basic data used for compiling the CPI, and data by item of 71 cities are released. Here, for individual items included in the CPI, we use data covering the period from January 2000 to December 2004. Similar to Higo and Saita (2007), however, the following three items are removed from our analysis: (i) items whose price data are not available throughout the year due to its seasonality; (ii) items not appropriate for analyzing the frequency of price changes due to the vast number of data aggregated in released prices; and (iii) items whose sequences of prices are too short.[16] As a result, the number of prices used for analysis adds up to a total of 26,177 made up of 493 items of 71 cities. Table 2 shows that the weight of data in the CPI used for our analysis is 67.8 percent. Among that, 42.8 percent accounts for the weight of goods and 25.0 percent for services.

Table 2: Coverage of data

		Number of items		CPI weight	
		CPI	Our data	CPI	Our data
Total		598	493	100.0%	67.8%
	Goods	456	372	50.5%	42.8%
	Services	142	121	49.5%	25.0%

Note: Weights are on the 2000 base.

5.2 Number of Types

The number of types is estimated in ascending order starting the number of types equal 1. The number of types is judged by the Bayesian Information Criterion (BIC).[17] It is determined when the BIC of the number m $(m = 1, 2, \ldots)$ is larger than that of $m+1$. Otherwise, the new BIC with $m+2$ is calculated and compared with the old one with $m+1$. We repeat the process until the number of types is determined.

Taking into account that the shapes of empirical hazard functions differ largely between goods and services, we estimate them separately. Table 3 shows the BIC by the number of types. For goods, the BIC determined the number of types as 11. For services, the BIC determined the number of types as 6 among which 5 types were without dummies

15 See Higo and Saita (2007) for details on the data and analysis on the frequency of price changes.

16 In principle, items of the CPI which are compiled using price data other than the *Retail Price Survey* are not included here. Nevertheless, it is included for analysis when the price data of the *Retail Price Survey* represents price developments. For details, see Higo and Saita (2007).

17 BIC is defined as $\mathrm{BIC} = l(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\pi} \mid \boldsymbol{\tau}, \mathbf{d}) - (1/2)k\log(n)$ where $l(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{\pi} \mid \boldsymbol{\tau}, \mathbf{d})$ is the logarithmic likelihood, k is the number of parameters and n is the number of samples.

and one type was with dummies.[18]

Table 3: The BIC and the number of types

(1) Goods

No. of Type	BIC
1	-314,520
2	-66,282
3	10,240
4	35,085
5	50,118
6	282,380
7	287,529
8	287,851
9	288,750
10	289,386
11	289,546
12	289,542

(2) Services

		No. of type with dummy	
		1	2
No. of type without dummy	1	-27,438	-26,657
	2	-26,660	-26,464
	3	-26,380	-26,372
	4	-26,336	-26,366
	5	-26,284	-26,354
	6	-26,297	-26,388

5.3 Estimation Results

The estimation results show that 17 types are identified with goods and services. And they are roughly classified into these four groups: (1) the flexible group, in which price changes occur every quarter; (2) the increasing hazard group which has an increasing hazard function; (3) the Calvo group, which has a flat hazard with low probability of price changes; and (4) the Taylor group in which price changes occur on a regular interval.

5.3.1 Goods

Table 4(1) shows the estimation results for goods. The number of types in table 4(1) is affixed to each type in descending order of λ. Table 4(1) shows that the null hypothesis, $\theta < 1$, is rejected significantly for types 1 to 9. Therefore, types 1 to 9 must have $\theta > 1$.

[18] As for the number of types, more than 11 types for goods, and 5 types without dummies and 2 types with dummies for services are not identified. For instance, when assuming the number of type for goods at 12, the weight of the 12th type becomes zero, and also the estimated parameter basically matches with the parameter of any one of the 11 types.

And they are classified as the flexible group or the increasing hazard group. In contrast, the null hypothesis is not rejected for types 10 and 11. That implies that they may still have decreasing hazards. Nevertheless, their value of θ s are close to 1. The estimated value is $\theta = 0.949$ for type 10 and $\theta = 0.968$ for type 11.

Figure 4(1) shows the shape of the estimated hazard functions for each type. According to the figure 4(1), we find that hazard functions of types 1 to 5 belong to the flexible group in which the frequency of price changes is extremely high with an almost a hundred percent probability of price changes after one quarter (after three months). The types 6 to 8 belong to the increasing hazard group. The types 9 to 11 belong to the Calvo group. They have almost flat hazard functions in which the probability of price changes is low.[19] The test of θ shows that the type 9 does not have a decreasing hazard. And the test of θ shows that the types 10 and 11 do not have increasing hazard functions.

The weight shown in parenthesis in figure 4(1) tells the composition of groups for goods. The flexible group (types 1 to 5), which has a high probability of price changes, account for about 20 percent of all items for goods. The increasing hazard group (types 6 to 8), which has a medium frequency of price changes, accounts for almost 30 percent. The Calvo group (types 9 to 11), which has a low frequency of price changes, accounts for almost 50 percent.

5.3.2. Services

Table 4(2) shows the estimation results for services. The type 6 is classified as the Taylor group which has a significantly higher probability of price changes at a specific time, twelfth and twenty- fourth months, than other months. The dummy variables λ_{12} and λ_{24} are both significantly positive. The test of θ >1 supports that the hazard functions of all types except for type 4 are not decreasing.

Figure 4(2) shows the shapes of hazard functions for services. It indicates rather noticeably that the types 1 and 2 belong to the increasing hazard group, while the types 3 to 5 belong to the Calvo group. The values of θ of the Calvo group (types 3 and 4) are close to 1. The probability of price changes of the type 5 is extremely low, that is, the value

---

[19] As regards the number of types and the pattern of hazards, when the number of types increases to a certain number, the overall differences between the two become subtle. For example, it may well be said that for goods, a rough characteristic can be expressed at around type 7 (in fact, the BIC in table 3 shows that there is hardly any difference from type 7 and above). When the number of type is 7, types 2 to 6 are aggregated into two types as is the case when the number of type is 11, and types 10 and 11 are aggregated into one type. As the number of types ascends from 8 onward, the type splits up in descending order of probability of price changes. In the end, when the number of type is 11, types 10 and 11 split up. A new type is not identified from type 12 and onward.

of λ is so small.

The weight of the increasing hazard group (types 1 and 2), which has a relatively high frequency of price changes, accounts for only one to two percent of all services. That percentage is considerably small compared to goods. On the other hand, the weight of the Calvo group (types 3 to 5), which has a low frequency of price changes, accounts for just below 70 percent. Finally, the Taylor group (type 6), in which price changes occur on a regular interval, constitutes about 30 percent of all services.

Table 4: Estimation results

(1) Goods

	λ	θ	π	Test of $\theta>1$
Type 1	192.0 *** (1.022)	174.8 *** (0.942)	0.053	184.5 ***
Type 2	4.799 *** (0.256)	4.709 *** (0.555)	0.038 *** (0.001)	6.680 ***
Type 3	3.750 *** (0.186)	3.867 *** (0.423)	0.043 *** (0.002)	6.781 ***
Type 4	2.675 *** (0.010)	2.975 *** (0.014)	0.032 *** (0.001)	144.3 ***
Type 5	1.880 *** (0.010)	2.367 *** (0.012)	0.039 *** (0.002)	112.4 ***
Type 6	1.280 *** (0.008)	1.924 *** (0.008)	0.060 *** (0.002)	112.1 ***
Type 7	0.739 *** (0.007)	1.564 *** (0.006)	0.094 *** (0.003)	90.97 ***
Type 8	0.250 *** (0.008)	1.285 *** (0.005)	0.130 *** (0.003)	52.95 ***
Type 9	-0.286 *** (0.011)	1.079 *** (0.005)	0.193 *** (0.005)	14.51 ***
Type 10	-0.988 *** (0.026)	0.949 *** (0.008)	0.212 *** (0.007)	-6.066
Type 11	-2.160 *** (0.101)	0.968 *** (0.027)	0.108 *** (0.008)	-1.178

(2) Services

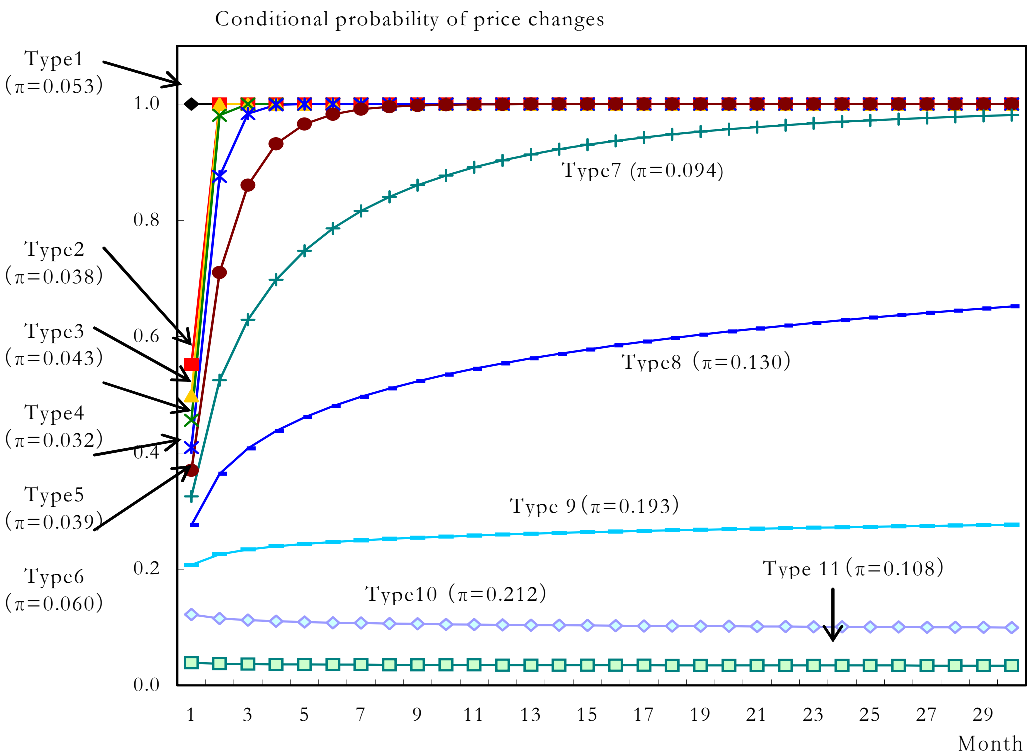
	λ_0	λ_{12}	λ_{24}	θ	π	Test of $\theta>1$
Type 1	1.183 *** (0.039)			1.955 *** (0.077)	0.002	12.49 ***
Type 2	0.050 (0.035)			1.330 *** (0.033)	0.015 *** (0.002)	10.07 ***
Type 3	-0.731 *** (0.033)			1.068 *** (0.018)	0.124 *** (0.010)	3.685 ***
Type 4	-1.741 *** (0.049)			0.972 *** (0.017)	0.444 *** (0.015)	-1.657
Type 5	-11.01 *** (1.931)			3.786 *** (0.687)	0.117 *** (0.017)	4.053 ***
Type 6	-5.609 *** (0.076)	4.612 *** (0.123)	3.478 *** (0.127)	1.746 *** (0.015)	0.298 *** (0.033)	50.09 ***

Notes 1. Figures in parentheses indicate standard errors.

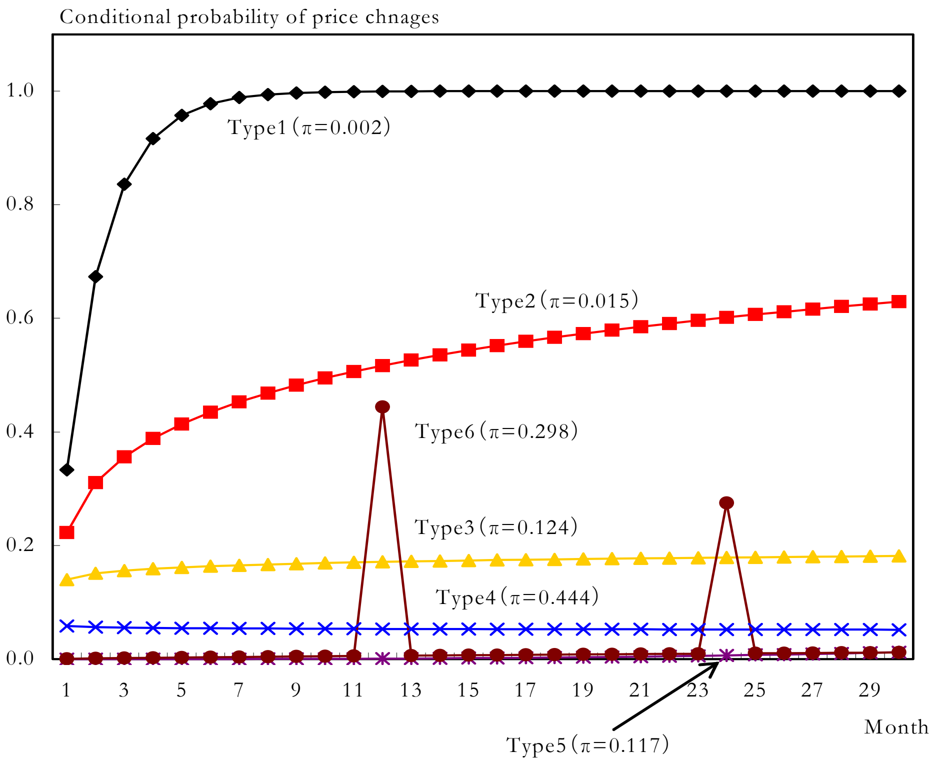
2. *** indicates significance at the 1% level.

Figure 4: Shapes of hazard functions

(1) Goods



(2) Services

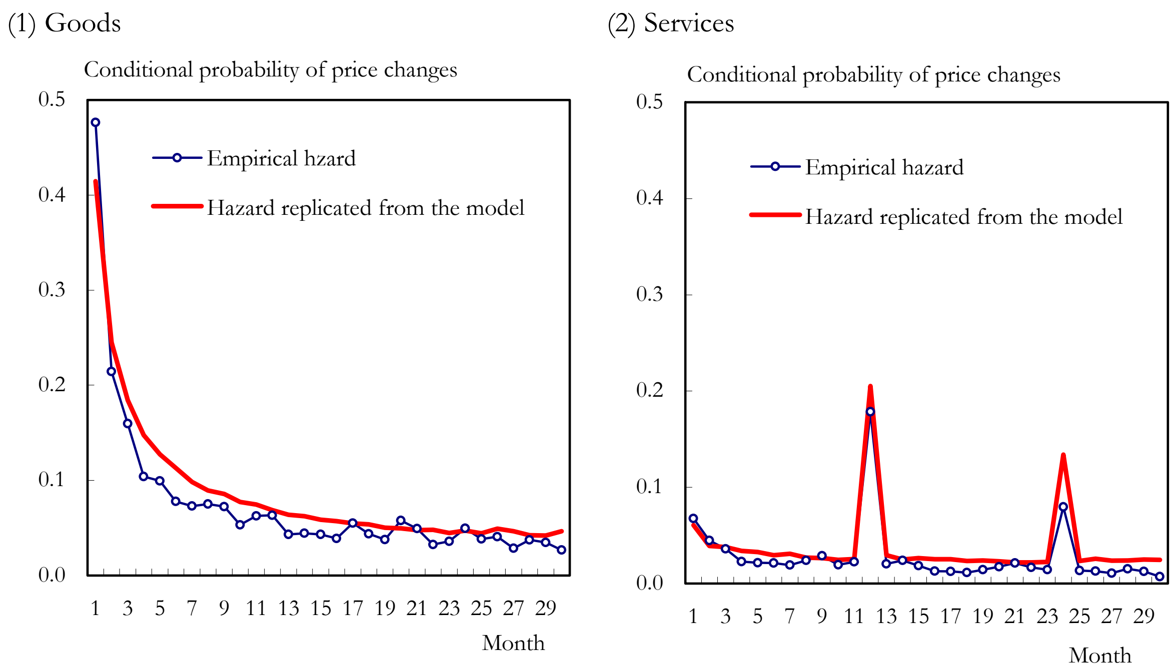


5.4 Empirical Hazard Replicated from the Estimated Model

We examine whether the empirical hazard function replicated from the estimated model fits the original empirical hazard function provided in figure 2. The empirical hazard replicated from the model is calculated as follows. First, we decide a type of each individual item from equation (10), which gives the posterior probability of item i belonging to the j th type. Next, we draw a trajectory of price changes by producing random variables from the estimated Weibull distribution for each item. And we reproduce basic data for calculating the same empirical hazard as in table 1. Finally, we calculate the aggregated empirical hazard function using that table.

Figure 5 shows replicated empirical hazard functions of goods and services. The replicated empirical hazard function of goods is decreasing and matches the original empirical hazard function very well. And that of services has a spike every 12-month and captures the characteristic of services very well. Hence, even without an individual item having a decreasing hazard, the heterogeneity allows the empirical hazard function to be decreasing. In addition, even with an individual item having an increasing hazard, the empirical hazard function could be decreasing.

Figure 5: Empirical hazards replicated from the estimated model



5.5 Relationship between types and items

We explains the relationship between the estimated types and items in detail here. The clear relationship between types and items exits. It helps us understand the characteristic of both types and items.

5.5.1 Goods

Table 5(1) shows the relationship between types and items for goods. For simplicity, the estimated types are classified into these three groups (i) the flexible group (types 1 to 5); (ii) the increasing group (types 6 to 8); and (iii) the Calvo group (types 9 to 11). First, the upper left of table 5(1) shows the contents of items in each group of types. The flexible group, which has high frequencies of price changes, mainly consists of fresh agricultural and aquatic products (35.0%). The increasing group, which has increasing hazards, mainly consists of food products (44.2%). The Calvo group, which has a flat hazard function with low frequencies of price changes, mainly consists of other industrial products (33.3%).

Second, the upper right of table 5(1) shows the contents of groups in each item. If we focus on items with more than 60% of group weights, fresh agricultural and aquatic products is classified as the flexible group and petroleum products is classified as the increasing group. Also textiles, the other industrial products, electricity, gas and water, and publications are classified as the Calvo group.

Third, the lower left of table 5(1) indicates the sharp difference between major large cities and small and middle cities. The flexible group and the increasing group have the highest weights of major cities' items (67.5% and 56.5% respectively). The Calvo group has the highest weights of small and middle cities' items (54.4%). In other words, major cities have more flexible price setters than small and middle cities do. That is because competition among price setters is higher in large cities. Also the sampling of the statistics may affect the difference. The number of surveyed stores is larger in major cities so that semi-aggregated prices in major cities tend to be more flexible than small and middle cities.

Finally, the lower right of table 5(1) indicates a high weight of the Calvo group on each group of cities. For all groups of cities, the Calvo group accounts for more than one-third of all groups of types. Especially, it accounts for more than 90% for the nationwide items. Nationwide items, whose prices are same for all cities, are very sticky.

5.5.2 Services

Table 5(2) shows the relationship between types and items for services. For simplicity, the estimated types are classified into these three groups (i) the increasing group (types 1 to 2); (ii) the Calvo group (types 3 to 5); and (iii) the Taylor group (type 6). In contrast to goods, services do not have the flexible group but have the Taylor group. First, the upper left of table 5(2) shows the contents of items in each group of types. The increasing group mainly consists of eating out (51.5%). The Calvo group also mainly consists of eating out (31.3%). The Taylor group, which has a spike every 12-month, mainly consists of education in general services (42.0%). Albeit all groups do not have the high weight of public services, the flat group has the relatively high weight of public services.

Second, the upper right of table 5(2) shows the contents of groups in each item. If we focus on the items with more than 60% of group weights, no item is classified as the increasing group. The two public services, medical care & welfare and education, and one general service, education, are classified as the Taylor group. The other groups are all classified as the Calvo group. The high number of items classified as the flat group is consistent with the low frequencies of services as shown in figure 5(2).

Third, as same as goods, the increasing group, the group with relatively high frequencies, concentrates in major cities as shown in the lower left of table 5(2). However, as for the Taylor group, no sharp difference between major cities and small and middle cities is observed.

Forth the share of the Calvo group is highest for all groups of cities as shown in the lower right of table 5(2). For items of small and middle cities and nationwide, the share of the increasing group is almost zero. Then, for services, almost no item with high frequencies exits in small and middle cities and nationwide.

Table 5: Relationship between types and items

(1) Goods

① By item

	Ratio of item in each group of types: Group of types 1〜5	6〜8	10〜11	Total	Ratio of group of types in each item: Group of types 1〜5	6〜8	9〜11	Total
Agricultual and aquatic products								
Fresh agricultual and aquatic products	35.6%	1.4%	0.4%	8.1%	92.5%	5.3%	2.3%	100.0%
Other fresh agricultual and aquatic products	12.2%	5.7%	1.8%	5.2%	49.7%	33.2%	17.2%	100.0%
Other agricultual and aquatic products	2.9%	4.2%	1.1%	2.5%	25.2%	52.4%	22.4%	100.0%
Industrical products								
Food products	31.2%	44.2%	25.2%	32.2%	20.4%	41.6%	38.0%	100.0%
Textiles	3.1%	4.8%	6.7%	5.4%	12.1%	27.2%	60.7%	100.0%
Petroleum products	0.8%	14.9%	4.9%	7.1%	2.4%	64.1%	33.4%	100.0%
Other industrial products	13.3%	22.6%	33.3%	25.9%	10.8%	26.6%	62.6%	100.0%
Electricity, gas and water	0.0%	1.8%	22.2%	11.3%	0.0%	4.9%	95.1%	100.0%
Publications	0.9%	0.3%	4.3%	2.4%	7.8%	3.7%	88.5%	100.0%
Total	100.0%	100.0%	100.0%	100.0%	21.1%	30.4%	48.6%	100.0%

② By city

	Ratio of group of cities in each group of types: Group of types 1〜5	6〜8	10〜11	Total	Ratio of group of types in each group of cities: Group of types 1〜5	6〜8	10〜11	Total
Major cities	67.1%	56.5%	39.6%	50.5%	28.0%	34.0%	38.0%	100.0%
Small and middle cities	32.0%	43.2%	54.4%	46.3%	14.6%	28.3%	57.1%	100.0%
Nationwide	0.9%	0.3%	6.1%	3.2%	5.8%	2.7%	91.6%	100.0%
Total	100.0%	100.0%	100.0%	100.0%	21.1%	30.4%	48.6%	100.0%

(2) Services

① By item

	Ratio of item in each group of types: Group of types 1〜2	3〜5	6	Total	Ratio of group of types in each item: Group of types 1〜2	3〜5	6	Total
Public services								
Domestic duties	0.0%	14.2%	5.7%	11.1%	0.0%	86.4%	13.6%	100.0%
Medical care & welfare	0.0%	0.5%	5.3%	1.8%	0.0%	19.9%	80.1%	100.0%
Forwarding & communication	0.0%	5.2%	6.4%	5.2%	0.0%	67.2%	32.8%	100.0%
Education	0.1%	0.9%	7.5%	2.6%	0.3%	22.9%	76.9%	100.0%
Reading & recreation	0.6%	3.5%	0.3%	2.4%	1.5%	95.7%	2.8%	100.0%
General services								
Eating out	51.5%	31.3%	10.6%	27.1%	12.2%	77.4%	10.4%	100.0%
Domestic duties	23.1%	24.0%	13.1%	21.1%	7.0%	76.5%	16.5%	100.0%
Medical care & welfare	0.0%	1.6%	1.2%	1.4%	0.0%	77.6%	22.4%	100.0%
Education	0.0%	4.7%	42.0%	14.3%	0.0%	22.2%	77.8%	100.0%
Reading & recreation	24.7%	14.1%	8.0%	13.1%	12.0%	71.8%	16.1%	100.0%
Total	100.0%	100.0%	100.0%	100.0%	6.4%	67.1%	26.5%	100.0%

② By city

	Ratio of group of cities in each group of types: Group of types 1〜2	3〜5	6	Total	Ratio of group of types in each group of cities: Group of types 1〜2	3〜5	6	Total
Major cities	96.9%	47.3%	47.9%	50.6%	12.2%	62.7%	25.1%	100.0%
Small and middle cities	3.1%	37.9%	46.8%	38.1%	0.5%	66.9%	32.6%	100.0%
Nationwide	0.0%	14.7%	5.4%	11.3%	0.0%	87.4%	12.6%	100.0%
Total	100.0%	100.0%	100.0%	100.0%	6.4%	67.1%	26.5%	100.0%

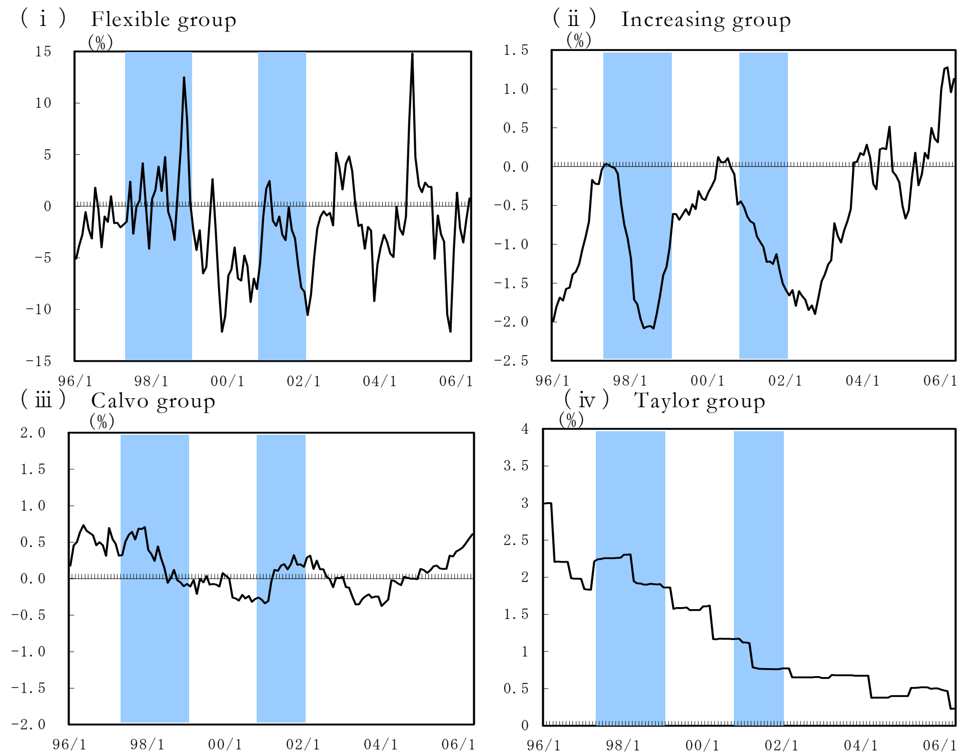
Notes 1: The weight of each price series is obtained by multiplying the weight by item and the weight by municipality in the CPI weight (2000 base).
2. Major cities are designated cities under article 252-19 of the Local Autonomy Law.

5.6 Developments of Price Indices by Group of Types

We examine the developments of price indices for each type. For simplicity the estimated types are classified into four groups: (i) the flexible group; (ii) the increasing group; (iii) the Calvo group; and (iv) the Taylor group. Price indices are calculated for those groups by taking weighted average of the individual price indices using the CPI weight.[20]

Figure 6 shows year-on-year changes of the price indices by group. Albeit the estimation only covers the period from 2000 to 2004, the period is extended backward to 1995 to capture developments in the price indices in the long run.[21] The figure shows the three findings: (i) the flexible group fluctuates substantially without a trend; (ii) the increasing group changes pro-cyclically to business cycles; (iii) the Calvo group changes moderately; and (iv) the Taylor group changes in a lumpy manner. The classification by group of those hazard functions is almost the same as that by the frequency of price changes. It results in extracting the group of items with different frequencies.

Figure 6: Developments of price indices by group of types (year-on-year-changes)



Notes 1: The 2000 base CPI is used for the price indices and weights of items.
2. Shaded areas indicate recessions.

[20] For simplicity, the type of each item is determined by averaging the type among cities, and then the price index of each group is calculated by taking weighted average of the CPIs of items in each group.
[21] Note that the estimation here should be discounted to a certain degree since the possibility that types may change around 2000 is not taken into consideration.

6 Implications

6.1 Does the Increasing Hazard Imply a State-dependent Pricing?

In the previous section, we obtained increasing hazard functions for goods. That, however, does not instantly mean that the items of goods classified as the increasing hazard group fall under the state-dependant model of Dotsey, et al. (1999). In this section, we verify the relative significance between the state-dependent and the time-dependent pricing using the method developed by Klenow and Krystov's (2005).

Let us explain the method of Klenow and Krystov (2005) briefly. First, the price change indicator I_{it} of item i at time t is defined as follows:

$$I_{it} = \begin{cases} 1 & \text{if } p_{it} \neq p_{it-1} \\ 0 & \text{if } p_{it} = p_{it-1} \end{cases} \tag{14}$$

where p_{it} is the logarithm of the price. Using the price change indicator (14), the inflation rate π_t is expressed as

$$\begin{aligned} \pi_t = \sum_{i=1}^{N} w_{it}\left(p_{it} - p_{it-1}\right) = \left(\sum_{i=1}^{N} w_{it} I_{it}\right) \times \left(\frac{\sum_{i=1}^{N} w_{it}\left(p_{it} - p_{it-1}\right)}{\sum_{i=1}^{N} w_{it} I_{it}} \right) \\ \equiv fr_t \times dp_t \end{aligned} \tag{15}$$

where w_{it} is the weight of the price. The equation (15) shows that the inflation is decomposed into two terms: (i) the fraction of items changing prices (fr_t) and (ii) the weighted-average magnitude of price changes (dp_t). The first order Taylor approximation around the average of fr_t and dp_t gives

$$\pi_t \simeq \overline{fr} \cdot \overline{dp} + \overline{dp}\left(fr_t - \overline{fr}\right) + \overline{fr}\left(dp_t - \overline{dp}\right) \tag{16}$$

where $\overline{fr}$ and $\overline{dp}$ are the averages of fr_t and dp_t respectively. The variance of π_t is calculated as:

$$\operatorname{var}\left(\pi_t\right) = \underbrace{\overline{fr}^2 \operatorname{var}\left(dp_t\right)}_{TDP} + \underbrace{\overline{dp}^2 \operatorname{var}\left(fr_t\right) + 2\overline{fr} \cdot \overline{dp} \operatorname{cov}\left(fr_t, dp_t\right)}_{SDP} \tag{17}$$

The above equation indicates that the variance of the inflation rate is decomposed into two terms: time dependent pricing (TDP) term and state dependent pricing (SDP) term. On the one hand, if the price setting behavior follows the time-dependent pricing, the $\mathrm{var}(fr_t)$ and $\mathrm{cov}(fr_t, dp_t)$ are zero since the price change ratio fr_t is constant[22]. Hence, the variance of the inflation rate depends only on the TDP term. On the other hand, if the price setting behavior follows the state-dependent pricing, the SDP term affects the variance of the inflation rate since the fraction fr_t fluctuates in response to shocks.

Klenow and Kryvstov (2005) broke down the variance of the inflation rate using micro data of the U.S. CPI (from 1998 to 2003) as explained above. That breakdown consequently shows that 88-101 percent of the variance of the inflation rate is explained by the *TDP* term. Also, calibrating the state-dependent model of Dotsey, et al. (1999), Klenow and Kryvstov (2005) claim that under standard parameters of the state-dependent model of Dotsey et al. (1999), the *TDP* and *SDP* terms must be explanatory by 20 percent and 80 percent, respectively. Furthermore, they stated that the time-dependent pricing is more plausible than the state dependent pricing. They set the distribution of the menu costs so as to replicate the actual *TDP* and *SDP* terms and showed that the distribution with the state-dependent pricing becomes that with time-dependent pricing approximately.

Table 6 shows the Klenow and Kryvstov's decomposition using our data. It shows that most of the fluctuations in the inflation rate are explained by the *TDP* term. The same decomposition for types 5 to 8 among goods, which have increasing hazard functions, shows that the weight of the TDP term is dominant for those types as well. It displays only a subtle difference from Klenow and Kryvstov (2005).

Those results imply that Japanese prices follow the time-dependent pricing. The implication of the results, however, requires the attention. The sample period used for the analysis is from 2000 to the end of 2004 when inflation rates in Japan were relatively stable.[23]

---

[22] Rigorously the fraction fr_t fluctuates even under the assumption of a time-dependent pricing because the weight w_{it} is different among items. We assume that there are many items and the weights are enough small to have a constant fr_t under a time-dependent pricing.

[23] That implication does not necessarily eliminate the possibility of the other state-dependent models. Golosov and Lucas (2003) focused on the point that while the rate of change in the aggregated price level is small and stable, that in the micro price level is substantial with many items. And many items change their prices frequently. Based on those facts, they constructed the state-dependent model which incorporates idiosyncratic productivity shocks to each firm. In that case, even when the inflation rate is stable, the micro-level prices fluctuate substantially depending on the state of the individual productivity shocks. Furthermore, Gertler and Leahy (2005) simplified Golosov and Lucas (2003) model and derived the New Keynesian Philips curve whose hazard function is flat. In addition, Angeloni, et al. (2005) referred that among micro data analysis of Europe by the INP, there exits evidence indicating the state-dependent pricing including Dotsey, et al. (1999).

Table 6: Decomposition of the variance

	TDP term	SDP term
All types (all goods and services)	102.1%	-2.1%
Goods: type 5	96.5%	3.5%
Goods: type 6	90.2%	9.8%
Goods: type 7	85.9%	14.1%
Goods: type 8	84.0%	16.0%
Klenow and Kryvtsov (2005)	88%～101%	-1～12%

6.2 Heterogeneity

In the previous section, we have demonstrated the multiple types of individual items and strong heterogeneity among those items. It is known that the heterogeneity has influence on the price-setting behavior. The model with the heterogeneity among price setters leads to two different implications from the standard representative agent model. First, the price stickiness estimated from the aggregated price level may possibly underestimate that estimated from the micro data. Carvalho (2006), extending Aoki (2004), constructed the model consisting of multiple Calvo types. His model shows that even when the average frequency of price changes at a macro level is the same, the price stickiness is larger in an economy with heterogeneous agents than in an economy with representative agents. We have confirmed the heterogeneous agents from an almost-perfect flexible agent to a sticky-pricing agent, and thus the intuition of the above discussion applies to our results.

Second, when the heterogeneity exists, the Phillips curve depends on the relative price of each type as shown in Aoki (2004) and Carvalho (2006). In addition, regardless of the heterogeneity and the pricing models, the Phillips curve drawn from agents with increasing hazard functions depends on past inflation rates and future expected inflations of many periods.[24] The heterogeneity obtained from our analysis implies that the Phillips curve has the complicated structure and does not necessarily hold the widely-used simple Calvo model.

7 Conclusion

We estimated the price hazard functions using the Japanese CPI micro data. For empirical analysis, we applied the finite mixture model with multiple spells to deal with the

[24] See Mash (2004) and Coenen, et al. (2006) for the time-dependent model and Bakhshi, et al. (2004) for the state-dependent model.

heterogeneity among items. We assumed the Weibull distribution to estimate various types of hazard functions including decreasing and increasing types. And we used a trajectory, a sequence of spells, in calculating the likelihood of each item to use all the information of the item.

Our estimation demonstrates the strong heterogeneity among items. Specifically, items are classified into these four groups: (i) the flexible group in which price changes occurs every quarter; (ii) the increasing group which has an increasing hazard function; (iii) the Calvo pricing group whose frequency of price changes is low; and (iv) the Taylor pricing group in which price changes occurs on a regular period. Unlike existing literature, the increasing hazard functions are estimated significantly while decreasing hazard functions are not. Second, the deceasing empirical hazard reported in many countries is likely to be caused by aggregation bias. In fact, by reproducing the empirical hazard function from several hazard functions estimated in the finite mixture model, we found that the aggregated empirical hazard function is deceasing, even though it includes no items with a sharp-decreasing hazard. Third, as for our sample period from 2000 to 2004, types with increasing hazards are possibly appropriate to the time-dependent model. The probability of price changes does not change through time.

Albeit we found new results about shapes of individual hazard functions, we are still on the long way to incorporate those findings into monetary models and obtain new implications of monetary policy. The heterogeneity among price setters is obvious in the actual world. So, what kinds of heterogeneity are important to the monetary policy analysis? We focused on hazard functions that have close relationship with the existing micro-founded price setting models. And we estimated several types of hazard functions. However, we do not get clear implications of monetary policy from the results since no model deals with the complicated heterogeneity estimated in this paper. In order to get clear implications of monetary policy, we need simple micro data facts that fit existing models or new models dealing with complicated facts. In this paper we reported the rather complicated facts. Linking models and rich micro data facts is our primary theme in our future research.

References

Alvarez, L. J., P. Burriel and I. Hernando (2005), "Do Decreasing Hazard Functions for Price Changes Make Any Sense?" European Central Bank Working Paper Series, No.461.

Alvarez, L. J., E. Dhyne, M. M. Hoeberichts, C. Kwapil, H. L. Bihan, P. Lunnemann, F. Martins, R. Sabbatini, H. Stahl, P. Vermeulen and J. Vilmunen (2006), "Sticky Prices in the Euro Area: A Summary of New Micro-Evidence", *Journal of the European Economic Association*, Vol.4, pp.575-584.

Angeloni, I., L. Aucremanne, M. Ehrmann, J. Gali, A. T. Levin, and F. Smets (2005), "New Evidence on Inflation Persistence and Price Stickiness in the Euro Area: Implications for Macro Modeling," mimeo.

Aoki, K., (2001), "Optimal Monetary Policy Responses to Relative Price Changes," *Journal of Monetary Economics*, Vol.48, pp.55-80.

Aucremanne, L. and E. Dhyne (2005), "Time-Dependent Versus State-Dependent Pricing: A Panel Data Approach to the Determinants of Belgian Consumer Price Changes," European Central Bank Working Paper Series, No.462.

Bakhshi, H., H. Khan, and B. Rudolf (2004), "The Phillips Curve under State-Dependent Pricing," mimeo.

Blanchard, O. J., and N. Kiyotaki (1987), "Monopolistic Competition and the Effects of Aggregate Demand," American Economic Review, Vol.77, pp.647-666.

Calvo, G. A. (1983), "Staggered Prices in a Utility-Maximizing framework," *Journal of Monetary Economics*, Vol.12, pp.383-398.

Cameron, A. C. and P. K. Trivedi (2005), "Microeconometrics: Methods and Applications," Cambridge University Press

Carvalho, C. (2006), "Heterogeneity in Price Stickiness and the New Keynesian Phillips Curve," *The B.E. Journal of Macroeconomics*, Vol.2: Iss.1 (Frontiers)

Christiano, L. J., M. Eichenbaum and C. L. Evans (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Model," *Journal of Political Economy*, Vol. 113, pp.1-45.

Coenen, G., A. T. Levin and K. Christoffel (2006), "Identifying the Influences of Nominal and Real Rigidities in Aggregate Price-Setting Behavior," mimeo.

Dhyne, E., L. J. Alvarez, H. L. Bihan, G. Veronese, D. Dias, J. Hoffmann, N. Jonker, P. Lunnemann, F. Rumler, and J. Vilmunen (2006), "Price Changes in the Euro Area and the United States: Some Facts from Individual Consumer Price Data," *Journal of Economic Perspectives*, Vol. 20, pp.171-192

Dotsey, M., R. G. King, and A. L. Wolman (1999), "State-Dependent Pricing and the General Equilibrium Dynamics of Money and Output," *Quarterly Journal of Economics*, Vol.114, pp.655-690.

Fougere, D., H. L. Bihan, and P. Sevestre (2005), "Heterogeneity in Consumer Price Stickiness: A Microeconometric Investigation," European Central Bank Working Paper Series, No.536.

Gertler, M. and J. Leahy (2005), "A Phillips Curve with an Ss Foundation," mimeo.

Golosov, M. and R.E. Lucas (2003), "Menu Costs and Phillips Curves," NBER Working Paper Series No.10187.

Higo, Masahiro and Yumi Saita (2007), "Price-setting in Japan: Evidence from CPI Micro Data", mimeo.

Klenow, P. J. and O. Kryvtsov (2005), "State-Dependent or Time-Dependent Pricing: Does It Matter for Recent U.S. Inflation?" mimeo.

Mash, R. (2004), "Optimizing Microfoundations for Inflation Persistence," Oxford University Department of Economics Discussion Paper No.183.

Mealli, F., and S. Pudney (1996), "Occupational Pensions and Job Mobility in Britain: Estimation of a Random-Effects Competing Risks Model," *Journal of Applied Econometrics*, Vol.11, pp.293-320.

Meyer, B. D. (1990), "Unemployment Insurance and Unemployment Spells," *Econometrica*, Vol.58, pp.757-782.

Nakamura, E. and J. Steinsson (2006), "Five Facts About Prices: A Reevaluation of Menu Cost Models," mimeo.

Taylor, J. B., (1980), "Aggregate Dynamics and Staggered Contracts," *Journal of Political Economy*, Vol. 88, pp.1-24.

Taylor, J. B., (1999), "Staggered Price and Wage Setting in a Macroeconomics," Chapter 15 in J. B. Taylor and M. Woodford (eds), *Handbook of Macroeconomics*, Vol.1B, North-Holland.

Woodford, M. (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton University Press.