JSPS Grants-in-Aid for Creative Scientific Research
Understanding Inflation Dynamics of the Japanese Economy
Working Paper Series No.52

Structural and Temporal Changes in the Housing Market and Hedonic Housing Price Indices

- A Case of the Previously Owned Condominium Market in the Tokyo Metropolitan Area -

Chihiro SHIMIZU
Hideoki TAKATSUJI
Hiroya ONO
and
Kiyohiko G. NISHIMURA
May 12, 2010

Research Center for Price Dynamics
Institute of Economic Research, Hitotsubashi University
Naka 2-1, Kunitachi-city, Tokyo 186-8603, JAPAN
Tel/Fax: +81-42-580-9138
E-mail: sousei-sec@ier.hit-u.ac.jp
http://www.ier.hit-u.ac.jp/~ifd/

Structural and Temporal Changes in the Housing Market and Hedonic Housing Price Indices[*]

- A Case of the Previously Owned Condominium Market in the Tokyo Metropolitan Area -

Chihiro SHIMIZU[†] Hideoki TAKATSUJI[‡] Hiroya ONO[§] Kiyohiko G. NISHIMURA[**]

May 1, 2010

Abstract

An economic indicator faces two requirements. It should be timely reported and should not significantly be altered afterward to avoid erroneous messages. At the same time they should reflect changing market conditions constantly and appropriately. These requirements are particularly challenging for housing price indices, since housing markets are subject to large temporal/seasonal changes and occasional structural changes. In this study we estimate a hedonic price index of previously-owned condominiums of Tokyo 23 Wards from 1986 through 2006, taking account of seasonal sample selection biases and structural changes in a way it enables us to report the indexes timely which are *not* subject to change after reporting. Specifically, we propose an overlapping-period hedonic model (OPHM), in which a hedonic price index is calculated every month based on data in the "window" of a year ending this month (this month and previous eleven months). We also estimate hedonic housing price indexes under alternative assumptions: (i) no structural change ("structurally restricted") and (ii) different structure for every month ("structurally unrestricted"). Results suggest that the structure of the housing market, including seasonality, changes over time, and these changes occur continuously over time. It is also demonstrated that structurally restricted indices that do not account for structural changes involve a large time lag compared with indices that do account for structural changes during periods with significant price fluctuations.

***JEL Classification Number*:** C43; C81; R21; R31
***Key Words*:** structural change, seasonal sample selection bias, (un)restricted hedonic model, overlapping-period hedonic model

---

[*] This research is a part of the project entitled: Understanding Inflation Dynamics of the Japanese Economy, funded by JSPS Grant-in-Aid for Creative Scientific Research (18GS0101), headed by Tsutomu Watanabe. During the study, we received relevant comments from Tsutomu Watanabe, Erwin Diewert(UBC), David Fenwick(Statistics Office, UK), Mick Silver(IMF), Koji Karato and Masayuki Nakagawa. We sincerely thank them for their help. This version will publish in *International Journal of Housing Markets and Analysis.*

[†] Associate Professor (PhD), The International School of Economics and Business Administration, Reitaku University, 2-1-1 Hikarigaoka, Kashiwa-Shi, Chiba 277-8686 Japan,e-mail: cshimizu@reitaku-u.ac.jp.

[‡] Professor (PhD), The International School of Economics and Business Administration, Reitaku University.

[§] Professor (PhD), The International School of Economics and Business Administration, Reitaku University.

[**] Deputy Governor (PhD), Bank of Japan.

1. Objectives of the study

Japan, the United States, and most advanced nations have experienced housing bubbles and subsequent collapses of the bubbles in succession. Recently, much attention has been focused on housing price indices. In macroeconomic policy, housing price indices are considered to be a possible candidate of “early warning signals” of sometimes devastating financial bubbles. In microeconomic spheres, there are growing needs for hedging against volatility in housing markets, and housing price indices may be used as a means of index trades.

Economic indicators including housing price indices face two requirements. They should be reported in a timely manner and should not be significantly altered afterward to avoid erroneous messages. At the same time they should reflect changing market conditions constantly and appropriately. These requirements are particularly challenging for housing price indexes, since housing markets are subject to large temporal/seasonal changes and occasional structural changes.

When we construct the housing price indices, we should consider the characteristics of the housing market. The specifications and facilities of each house differ in varying degrees, so there are no two houses of identical quality. Even when the specifications and facilities are identical, if the age of the building differs, the degree of deterioration differs accordingly, so that the houses are not identical. In other words, houses have “particularity with few equivalents”. In addition to such a problem, the quality of houses (in particular, condominiums) changes with time owing to fairly rapid technological progress. Such characteristics are particularly evident in the housing market of Japan compared with the United States and other countries (Shimizu, Nishimura and Asami, 2004).

To date, a number of methods for estimating housing price indices have been investigated at the research level. Among these, the methods used the most are the hedonic method proposed by Court (1939) and established theoretically by Rosen (1974) and the repeat sales (RS) method elaborated by Bailey et al. (1963) and Case and Shiller (1989).[1]

In 2006, the OECD and IMF held a first international conference related to the construction of housing price indices. After that meeting, further discussions took place in May 2009 at the U.N.'s Ottawa meeting and then at a BIS conference in November 2009.[6] They focused on the establishment of a housing price information-gathering system and on calculation methods. The two main methods of calculating housing price indices are the Repeat Sales Method and the Hedonic Method. Between them, it has been decided that international guidelines should be developed in a unified manner using the Hedonic Method.[7] Then, if the Hedonic Method is going to be developed as a housing price index, what points need to be paid attention to?

In this paper, we proposed a new estimation methodology for hedonic price indices. The rest of this paper is organized as follows. Section 2 presents an overview of the methodology of house price indices. Section 3 describes properties of housing in the dataset. The dataset we use in this paper is compiled from individual listings in a widely circulated real estate advertisement magazine. The dataset contains more than 0.2 million listings of housing prices, from 1986 to 2006, including a housing bubble period and its

---

[6] The next meeting will be held at the United Nations (United Nations Economic Commission for Europe (UNECE), International Labour Office (ILO), Meeting of Experts on Consumer Price Indices, Geneva, Switzerland,10-12.May.2010.

[7] Diewert (2007) pointed out that repeat sales methods give rise to two problems: the depreciation problem and the renovation problem. In the U.S., the repeat sales method in estimating housing indices is used in general and which was proposed by Case and Shiller (1989). It should be noted that an adjustment to cope with the depreciation problem and the renovation problem is conducted in constructing the official S&P/Case-Shiller home price index. Standard & Poor's (2008) states that "Sales pairs are also weighted based on the time interval between the first and second sales. If a sales pair interval is longer, then it is more likely that a house may have experienced physical changes. Sales pairs with longer intervals are, therefore, given less weight than sales pairs with shorter intervals." Moreover, Shimizu, Nishimura, and Watanabe (2010) pointed out that the repeat sales index lagged behind the hedonic index.

burst. Section 4 presents estimation results. We find that standard hedonic measures are biased because of changes in parameters over time. More importantly, we find that there exists a substantial discrepancy in terms of turning points between new structural change adjusted hedonic price indices and standard hedonic price indexes. Specifically, the standard hedonic price indices tend to exhibit a delayed turn compared with the new structural change adjusted hedonic measure. Section 5 concludes the paper.

2. Changes in market structure and housing price indices

2.1. Hedonic Models

The hedonic method treats a given property price as an aggregate (a bundle of characteristics) of the various values of that property's characteristics (qualities and features) and estimates the value of each characteristic using regression analysis. If pooled data are used, it is possible to calculate quality-adjusted price indices.

The estimation of housing price indices by the hedonic price model has the following two major problems: (i) omitted variables bias, the occurrence of bias in housing price indices because of the difficulty in collecting all the variables required for the estimation of functions as well as because of the presence of unobservable factors such as environmental variables (see Case and Quigley, 1991; Clapp, 2003), and (ii) a structural change issue, that is, the necessity of accommodating changes in the house price structure, because the housing market is examined over a long period of time (Case, Pollakowski and Wachter, 1991; Clapp, Giaccotto and Tirtiroglu, 1991; Clapp and Giaccotto, 1992, 1998).

Regarding problem (i), problems relating to the control of unobservable environmental variables can be avoided in the repeat-sales method (Case and Quigley, 1991; Case and Shiller, 1989). However, because the fluidity of housing markets is considerably lower (i.e., the market is thinner) in Japan than in the United States and other nations, and because institutional restrictions strongly suppress reselling within a short period, in accordance with the law based on the National Land Use Plan, problems with repeat-sales sample selection bias unique to Japan will still occur. If the repeat-sales method is to be applied in Japan, such a sample selection bias will be an extremely large problem, and in addition to that, the estimation of housing prices with high renewal frequency is impossible because of the small number of samples, and the estimation of indices in a limited area is difficult; consequently, the repeat-sales method is not a very practical method (Shimizu, Nishimura and Watanabe (2009)).

Under such circumstances, the importance of estimating hedonic housing price indices with high accuracy while solving the above problems inherent in the hedonic price model is extremely high in Japan. Therefore, in this study, we focus on the greatest problem involved in the hedonic price model, which relates to changes in the market structure. The omitted variables problem (i) will be discussed in another report (according to Clapp 2003, unobservable variables are handled by adding coordinate data).

For solving the structural changes problem in the estimation of a structurally restricted hedonic model (hereafter, also referred to as the RHM) under the assumption of no changes in the market structure, several researches (Case, Pollakowski, and Wachter, 1991; Clapp, Giaccotto and Tirtiroglu, 1991) propose a structurally unrestricted hedonic model (hereafter, also referred to as the URHM) under the assumption that the structure changes in each period.

The URHM assumes that the market structure changes each period, successively.

Such a structural change of the market occurs as a result of various external shocks. It is considered that there is, in reality, a certain adjustment period before such a change penetrates into the market. Accordingly, regression coefficients should be regarded as changing successively rather than instantaneously. However, generally, the estimation of a model with structural changes is performed by dividing observation data into several periods with break points, then using the divided data of each period (for example, Shimizu and Nishimura, 2007). Namely, the continuity of the observation data is disconnected at the break points. Therefore, it is rather difficult to use such an estimation method, under the assumption of the occurrence of successive structural changes, to determine regression coefficients allowing for successive changes.

Instead, it may be more natural and desirable to estimate regression coefficients on the basis of a process of successive change by taking a certain period length τ as the estimation period and shifting this period, similar to the process of obtaining moving averages.

We propose a new hedonic price conducting method, the Overlapping Period Hedonic Model (hereafter, also referred to as the OPHM)). What is distinctive about the OPHM is that it is not necessary to identify when a structural change occurred, and, moreover, if the target period length for the estimate is taken as τ, it is estimated while overlapping the data for the period $\tau-1$, and it is therefore possible to continue using the information from the previous period.

The various estimation methods are shown below.

2.2. Equations of hedonic models

2.2.1. Structurally restricted hedonic housing price index: RHM

Assume that we have data for housing prices and residential property characteristics, which are pooled for all the periods $t = 1, 2, \ldots, T$, and that the number of data samples in each period is n_t. A housing price estimation model that can be used to obtain a structurally restricted price index is given as follows.

$$\ln P_{it} = \sum_{k=1}^{K} \beta_k X_{ikt} + \sum_{s=1}^{\tau} \delta_s D_s + \varepsilon_{it} \qquad (1)$$

$t = 1, 2, \ldots, T$.
$i = 1, 2, \ldots, n_t$ (designates ith data among the n_t data samples in period t).
P_{it} = price of housing i in period t (designates ith data among the data in period t, instead of designating the same housing i over each of the t periods).
β_k = parameter of residential property characteristic k.
X_{ikt} = value of property characteristic k of housing i in period t.
δ_s = parameter of the time dummy variable in period s.
D_s: when $s = 1$, this takes a constant value of 1 (constant term). When $2 \leq s \leq T$, this is a time dummy variable, and it takes a value of 1 when $s = t$ and a value of 0 otherwise.
ε_{it} = random disturbance term.

This model is called the structurally restricted hedonic model (RHM) because it assumes that the regression coefficient β_k of the housing-price determining factor X_{ikt} is constant throughout all the periods. From this, the RHM is obtained as follows. The estimated price $\hat{P}_t$ of housing with residential property characteristic values $\{X_k\}$ ($k = 1, 2, \ldots, K$) in period

t ($t = 1, 2, \ldots, T$) is given as follows.

$$\ln \hat{P}_t = \sum_{k=1}^{K} \hat{\beta}_k X_k + \hat{\delta}_1 + \hat{\delta}_t \tag{2}$$

$$\ln \hat{P}_1 = \sum_{k=1}^{K} \hat{\beta}_k X_k + \hat{\delta}_1 \tag{3}$$

Here, $\hat{\beta}_k$, $\hat{\delta}_1$, and $\hat{\delta}_t$ are estimated values of the parameters. Accordingly, the housing price index $\hat{P}_t / \hat{P}_1$ in period t, where the housing price in period $t = 1$ is used as the reference, is obtained as follows.

$$\ln(\hat{P}_t / \hat{P}_1) = \hat{\delta}_t \tag{4}$$

In addition, the change in the price index from period $t-1$ to period t can be expressed as follows.

$$\ln(\hat{P}_t / \hat{P}_{t-1}) = \hat{\delta}_t - \hat{\delta}_{t-1} \tag{5}$$

In this case, the price index is obtained under the assumption of specific residential property characteristic values $\{X_k\}$. However, as we can see in the above process, price indices are expressed using only time dummy variables without involving residential property characteristic values in the RHM. Seasonality and its changes over time are also a part of the latter problem.

2.2.2. Structurally unrestricted hedonic housing price index: URHM

Using similar data to those described above, a housing price estimation model that can be used to obtain a URHM is given as follows.

$$\ln P_{it} = \sum_{k=1}^{K} \beta_{kt} X_{ikt} + \delta_t + \varepsilon_{it} \tag{6}$$

Here, no time dummy variables are used, and instead, the parameter β_{kt} and a constant term δ_t related to residential property characteristics are assumed to change in each period. Namely, because the model does not assume the restriction of constant parameters, it is called the URHM. From equation (6), the URHM is obtained as follows. The estimated price $\hat{P}_t$ of housing with residential property characteristic values $\{X_k\}$ (k = 1, 2, …, K) in period t (t = 1, 2, …, T) is given as follows.

$$\ln \hat{P}_t = \sum_{k=1}^{K} \hat{\beta}_{kt} X_k + \hat{\delta}_t \tag{7}$$

$$\ln \hat{P}_1 = \sum_{k=1}^{K} \hat{\beta}_{k1} X_k + \hat{\delta}_1 \tag{8}$$

Therefore, the housing price index $\hat{P}_t / \hat{P}_1$ in period t, where the housing price in period t = 1 is used as the reference, is obtained as follows.

$$\ln(\hat{P}_t / \hat{P}_1) = \sum_{k=1}^{K} (\hat{\beta}_{kt} - \hat{\beta}_{k1}) X_k + (\hat{\delta}_t - \hat{\delta}_1) \tag{9}$$

In addition, the change in the price index from period $t-1$ to period t can be expressed as follows.

$$\ln(\hat{P}_t / \hat{P}_{t-1}) = \sum_{k=1}^{K} (\hat{\beta}_{kt} - \hat{\beta}_{k,t-1}) X_k + (\hat{\delta}_t - \hat{\delta}_{t-1}) \tag{10}$$

Thus, for the URHM, price indices are obtained for specific residential property characteristics. When specific residential property characteristics change, the price index changes accordingly.

2.2.3. Hedonic model considering market structural changes

Assuming that we have pooled data over the periods 1, 2, …T, with respect to some of these periods, i.e., a period length τ, (which might be called a "window"), we assume the following basic model.

$$\ln P_{it} = \sum_{k=1}^{K} \beta_k X_{ikt} + \sum_{s=1}^{\tau} \delta_s D_s + \varepsilon_{it}, \tag{11}$$

where t = 1, 2, …, τ (taking part of the entire pooled data consisting of 1, 2, …T periods, namely, taking a certain period length τ, the periods within this range are numbered from 1 to τ).

In addition, we express a period with length τ starting from period r among periods 1, 2, …T, as $[r, r + \tau - 1]$. Then, our estimation method is obtained by applying the above basic model to periods [1, τ], [2, τ + 1], …, $[r, r + \tau - 1]$, …,$[T - \tau + 1, T]$ successively. From this, successive changes of the market structure can be reflected in changes in the parameters. We call this model the OPHM, and the period length τ is the overlapped estimate period length.

The OPHM is an RHM with respect to a certain period length τ. Accordingly, the parameter of the time dummy variable represents the price index of each period with the starting period of length τ as the reference. Thus, price indices can be obtained directly from the basic model within the period length τ. With the OPHM, models for all the periods are estimated by successively shifting the period length τ by one period. Here, the problem remains of how to connect the price indices obtained by the estimation in each period length τ to construct the price index for all periods. Our method is as follows.

We designate the housing price index throughout all the periods as q_r. This represents the price index of period r among periods 1, 2, …T. We designate the reference period as period 1 and assume that $q_1 = 0$. We also designate a parameter of the time dummy variable obtained by applying the basic model to the data for the period length τ starting from period r among periods 1, 2, …T, i.e., $[r, r + \tau - 1]$ is designated as $\hat{\delta}_1^{(r)}, \hat{\delta}_2^{(r)}, \cdots, \hat{\delta}_\tau^{(r)}$ by explicitly expressing the period r.

3. Data

3.1. Price data

The subject of the analysis is the 23 wards of the Tokyo metropolitan area (621 square kilometers), and the analysis period is approximately 20 years between January 1986 and September 2006.

As the main information source, we used the prices of previously-owned condominiums published in Residential Information Weekly (or *Shukan Jyutaku Joho* in Japanese) published by RECRUIT, Co. This magazine provides information on the characteristics and asking prices of listed properties on a weekly basis. Moreover, *Shukan Jutaku Joho* provides time-series data on housing prices from the week they were first posted until the week they were removed as a result of successful transactions.[8] We only use the price in the final week because this can be safely regarded as sufficiently close to the contract price.[9]

3.2. Data regarding housing characteristics

Table 1 shows a list of housing characteristics. These include the floor space (*FS*) and the *Age* as key housing characteristics. The convenience of public transportation from each housing location is represented by travel time to the central business district (CBD),[10] which is denoted by *TT*, and time to the nearest station,[11] which is denoted by *TS*. We use a ward dummy, *LD*, to indicate differences in the quality of public services available in each district, and a railway line dummy, *RD*, to indicate along which railway/subway line a housing property is located.

---

[8] There are two reasons for the listing of a unit being removed from the magazine: a successful deal or a withdrawal (i.e. the seller gives up looking for a buyer and thus withdraws the listing). We were allowed access to information regarding which of the two reasons were applicable for individual cases and discarded those where the seller withdrew the listing.

[9] Recruit Co., Ltd. provided us with information on contract prices for about 24 percent of all the listings. Using this information, we were able to confirm that prices in the final week were almost always identical with the contract prices (i.e., they differed at a probability of less than 0.1 percent).

[10] Travel time to the CBD is measured as follows. The metropolitan area of Tokyo is composed of 23 wards centering on the Tokyo Station area and containing a dense railway network. Within this area, we choose seven railway/subway stations as the central stations, specifically Tokyo, Shinagawa, Shibuya, Shinjuku, Ikebukuro, Ueno, and Otemachi. Then, we define travel time to the CBD by the minutes needed to commute to the nearest of the seven stations in the daytime.

[11] The time to the nearest station, *TS*, is defined as walking time to the nearest station if a housing property is located within walking distance from a station, and the sum of the walking time to a bus stop and onboard time from the bus stop to the nearest station if a housing property is located in a bus transportation area within walking distance from a station. We use a bus dummy, *BD*, to indicate whether a housing property is located in an area within walking distance from a station or in a bus transportation area.

The observation data consist of 211,179 samples collected between January 1986 and September 2006.

Table 1 also shows descriptive statistics of the major variables. The average resale price of a condominium is 39.04 million yen, with a fairly large standard deviation of 23.48 million yen. Regarding the *FS*, the average is 56.57 m^2[a] and as for the *Age*, the average value is 165 months (13.75 years). Because the history of condominiums in Japan is short, it is expected that this index will increase over time.

Regarding the *TS*, the average value is 7.60 minutes. On average, while many properties are conveniently located, some are beyond walking distance. This indicates that, in general, convenience is emphasized in the construction of condominiums because of the required characteristics of condominiums. Regarding the *TT* to the CBD, the average is 15 minutes, indicating that most condominiums are concentrated in areas of high convenience.

4. Estimation results

4.1. Estimation of RHM: Restricted Hedonic Index

The estimated RHM for the 23 wards of Tokyo is as follows.

$$\begin{array}{llllll}
\log RP/FS = 4.631 & +0.0126\cdot \log FS & -0.189\cdot \log Age & -0.078\cdot \log TS & -0.117\cdot \log TT & +0.019\cdot \log NU \\
\quad (498.23) & (+10.81) & (-337.38) & (-99.69) & (-36.21) & (40.90)
\end{array}$$

$$\begin{array}{llllll}
-0.276\cdot BD & +0.058\cdot(BD\times \log WT) & -0.026\cdot FF & +0.018\cdot HF & -0.097 FD & +0.0093\cdot SD \\
(-13.140) & (6.970) & (-19.210) & (8.000) & (-10.150) & (10.790)
\end{array}$$

$$+\beta_{1h}\sum_h LD_h + \beta_{2i}\sum_i RD_i + \beta_{3j}\sum_j TD_j$$

（ ）: t-value

Adjusted R-square: 0.837

Number of observations: 211,178

Because the coefficient of determination adjusted for the degree of freedom is 0.837, the estimated model has a fairly high explanatory power (refer to Table 2 for details).

Because the data was pooled for sales between 1986 and 2006, we corrected the time point by forcibly introducing the *TD* (time dummy) and the *RD* (railway line dummy) so that the structure of the previously-owned condominium price was estimated using property characteristics specific to condominiums. Among the property characteristics specific to condominiums, *FS* (floor space), *BS* (balcony space), and *NU* (number of units) have positive values, and *Age* (age of building), *TS* (time to nearest station), and *TT* (travel time to CBD) are estimated with negative values.

First, regarding *FS*, the unit price was shown to increase with increasing floor space. A similar tendency was observed for *BS* and *NU*. This indicates that consumers show a strong preference for the floor space of each property as well as the floor space of the entire condominium.

As *Age* increases, we expect not only functional deterioration but also economic deterioration because of the improvement of facilities in newer condominiums. The results obtained showed that as *TS* and *TT* to the CBD increase, the convenience decreases because of the greater distance from populated areas, resulting in a decrease in the price.

Furthermore, the level of public service differs for each administrative ward, and there are broad differences in residential environments depending on administrative cities and wards or railway line areas which cannot be taken into consideration in our estimated function. Therefore, these differences were estimated using the dummy variables.

4.2. Estimation of URHM: Un-Restricted Hedonic Index

Next, we estimated the URHM. In accordance with the definition of equation (6), we divided the data into *t* periods (here, monthly) and estimated the structure of housing prices. Regarding the price index, we estimated the prices of previously-owned condominiums in each period by substituting the specific residential property characteristics common to all periods into the explanatory variables, and obtained the structurally unrestricted hedonic housing price indices relative to the reference period based on the estimated prices.

Figure 1 shows changes in the number of samples and the coefficient of determination adjusted for the degrees of freedom.

The adjusted coefficient of determination decreased from 1986 through 1995, and then increased from 1996. However, on the whole, it maintained an average of around 0.75, showing fairly good results.

The number of samples was approximately 500 per month from 1986 through 1989, which then increased significantly to an average value of 844. However, there is more than a threefold difference depending on the month. In each year, transactions are concentrated from January to March, which is the end of the fiscal year, when there are large movements of people in Japan, and the number of transactions significantly decreases around July and August, thus showing seasonal changes. However, there is no apparent correlation between the number of samples and the adjusted coefficient of determination.

Next, we focused on the regression coefficients of the estimated model. Table 3 shows descriptive statistical values for the regression coefficients over 250 periods. Figures 2 and 3 show changes in the regression coefficients over time. All of the regression coefficients show major fluctuations in each period or every several periods. However, the fluctuations show a certain tendency, though not a gradual change, over time. In addition, we can see that all the variables are around (higher or lower than) the regression coefficients estimated using the RHM.

Thus, structural changes are expected to occur as a background to changes in the regression coefficients. For example, with respect to *FS* (Figure 2), the regression coefficient is estimated as positive in the basic model, while it is estimated as negative from 1987 to the end of 1995. This indicates that the structure apparently differs for this period (1987–1995) compared with other periods.

For *Age* (Figure 3), the trend is a decrease until the mid-1990s, followed by an increase thereafter. While such changes are observed over all periods, there are major fluctuations within short periods beyond the range of such changes (beyond the difference in the regression coefficient between the beginning and ending periods). These changes in the short periods cannot be considered to be structural changes.

On the basis of these findings, we consider that part of the change in the regression coefficients over time is caused by structural changes. It is also expected that some bias in the observation data group in each period may have a certain effect.

4.3. Estimation of OPHM

4.3.1. Estimation of model and changes in price structure

Next, an OPHM is estimated on the basis of the procedure described in Section 2.2.3. In the OPHM, to absorb changes during a short period, the presence of which became apparent in the estimation of the URHM, data over a certain period are pooled for the estimation. Using such an estimation method, we expect to absorb changes during a short period; however, upon estimation of the OPHM, setting an overlapped estimation period length (τ) is the key.

In the estimation of the URHM, it has been clarified that there is a seasonal characteristic of changes in the thickness (more precisely, the number of transactions) of the housing market, and that the number of transactions increases at the end of the fiscal year between January and March, and the number decreases in July and August. When we attempt to absorb such a seasonally fluctuating characteristic in the market, the estimation period should be set to exceed one year. When the period is set to be longer, parameters are expected to be stable. However, it becomes difficult to accurately represent changes in the market structure. In this study, on the basis of such assumptions, we examine the effects of varying τ between 12 months and 36 months on the price index and on the regression coefficients that represent the price-forming structure of the major variables.

First, estimation results using τ of 12 months are shown in Figure 4. The coefficient of determination decreases from 1986 through 1995, similar to that in the URHM, and then it increases from the second half of 1996. On the whole, the coefficient of determination maintains an average of approximately 0.75, indicating a good result similar to that of the URHM.

Next, we focused on the regression coefficients of the estimated model. Table 3 shows descriptive statistical values of the regression coefficients for 238 periods. Figures 5 and 6 show temporal changes in the regression coefficients.

When we look at the changes in the regression coefficients over time, the wide fluctuations observed in the URHM are eliminated and smooth changes are shown, making it easier to clarify overall tendencies. The absolute values of the regression coefficients of the *TS* and *TT* to the CBD tend to decrease over time.

The coefficients of variation were obtained as follows (Table3): *FS* = 2.424 (2.428 in the URHM), *Age* = −0.163 (−0.179 in the URHM), *TS* = −0.178 (−0.232 in the URHM), and *TT* to CBD = −0.554 (−0.779 in the URHM). On the whole, the values are smaller than those in the URHM. However, the variations of the coefficients of *FS* and *Age* are not significantly different from those of the URHM, indicating that the regression coefficients changed not because of temporal changes caused by sample bias, but because of large structural changes. The coefficients of variance of *TS* and *TT* to the CBD were significantly lower than those in the URHM. This is speculated to be because of large temporal changes in the regression coefficients caused by sample bias existing in a unit of time.

Next, in Figures 5 and 6, we observed changes in the regression coefficients when τ was varied between 12 months and 36 months. Compared with the estimation results of the URHM, with regard to the temporal changes in the regression coefficients of each variable, we can see the presence of a time lag in the changes in regression coefficients as τ becomes longer. This tendency is commonly observed in all the variables. Such a time lag is also expected to affect the price indices.

4.3.2. Evaluation of indices by different τ

We set τ in order to avoid the effects of specific bias in the observation data when the data were divided monthly. The bias, if present, is reflected in the regression coefficients and may generate effects that are difficult to differentiate from structural changes. The background of the bias is not clear at this stage of the study, but from the monthly samples, seasonal changes in the timing of when properties are put on the market were confirmed. Therefore, if we wish to avoid the seasonal fluctuation characteristic, periods of at least 12 months are necessary. Namely, the data division should accommodate the bias resulting from seasons with or without a large movement of people. However, when τ was extended, the presence of a time lag relative to changes in the market was also observed.

Therefore, we varied τ between 12 months and 36 months, and observed the changes in the price indices, as shown in Figure 7. We can see that the time lag in the regression coefficients also affects the price indices. As the connection period length increases, while the price change excluding the effects of regression coefficients is delayed, the increase in the price index occurs at a time lag. This effect is particularly strong in the reversal period of 2002 and after, when prices hit bottom.

These findings indicate that τ should be short for the estimation of the price index. Then, because it is necessary to overlap at least 12 months to exclude the seasonal change in the number of transactions, we determined the optimal τ to be 12 months.

4.4. Evaluation of OPHM

In the analysis described in the previous section, the OPHM with $\tau = 12$ months was shown to be the most accurate in representing the market trends. Here, we compare the RHM, the URHM, and the OPHM for $\tau = 12$ months (Figure 8).

When these indices are compared, large fluctuations in the URHM are noted. The magnitude of these fluctuations seems to be different from our actual experience of price changes, because we did not experience any large increase or decrease in the prices of condominiums with specific qualities during the periods when the indices showed a large increase or decrease. In particular, because the fluidity of the condominium market is low, instantaneous changes in the price are not expected.

The magnitude of fluctuation in an index is not an a priori point of evaluation of the quality of the index. However, such excessive fluctuations cannot be justified. In particular, while large changes were observed in the regression coefficient, it is difficult to expect that consumers in the market randomly and significantly changed their preferences. Therefore, we speculate that the regression coefficient largely changed because of bias in the data, and that, as a result, fluctuations in the price index occurred.

When we compare the RHM with the OPHM, there are large differences between the two from 1986 through 1990. When we further compare the URHM with the above two indices, the URHM showed random fluctuations centered on the OPHM. While the URHM enables us to represent market changes most sensitively, it fluctuates considerably because of the large bias in samples. Considering these characteristics the URHM can represent changes in the market structure more appropriately than the RHM, and is expected to be more accurate.

The results of the above analysis demonstrated that among the RHM, URHM, and OPHM (τ = 12 months), the OPHM (τ = 12 months) is superior, because the RHM is not able to accurately estimate the market trends under the circumstance of large changes in the price level and the price structure during the years from 1986 through 1990, and because the URHM fluctuates considerably due to the seasonal fluctuation of transactions as well as other biases.

5. Conclusions

After international housing price index guidelines are developed in 2011, the governments of various countries will begin to develop housing price indices based on the hedonic method. In so doing, they will face three significant problems. The first problem is the question of which model to use for estimating the index. Various models exist for the hedonic method, as we have shown in this paper. The second problem is the issue of implementation. It is not sufficient to estimate economic indicators only once – they must continue to be estimated periodically (monthly) and perpetually. In this case, price indices must be estimated according to market conditions; in other words, they must respond to market structural changes. The third problem is that it is also necessary to address the issue of seasonal sample selection bias, since indices need to be estimated on a monthly basis.

In response to these problems, this paper has proposed a new hedonic model that is able to address short-term seasonal sample selection bias and long-term structural changes to the market.

As a result of the analyses, we clarified the following.

First, when using the URHM model, which assumes changes in the market structure each period, the regression coefficients widely fluctuate in each period or every couple of periods. Although this fluctuation is observed within short periods, a specific trend is observed in the long term. The fluctuation is speculated to be because of not only the occurrence of structural changes in the condominium market, but also the occurrence of bias in the transaction samples in each period. More precisely, we find that there is some seasonal fluctuation in the condominium market, which leads to active or sluggish transactions depending on the time of year, and that such changes affect the regression coefficients.

To eliminate the bias existing in transaction samples in each period, which are mainly exhibited as the seasonal fluctuation, and to eliminate changes in the market structure, we proposed the OPHM. Here, to eliminate the bias in the number of transactions and in samples affected by the seasonal change in the number of transactions, we set τ from 12 months to 36 months for the estimation. The results depicting the changes in major regression coefficients over time showed gradual changes instead of the wide fluctuations observed in the URHM. Thus, it became easier to understand the trend.

When τ between 12 months and 36 months were compared, the existence of a time lag in the regression coefficients was observed. This time lag in the regression coefficients brought about a time lag in the price index. From these findings, τ should be short. In addition, considering the aim of resolving the bias as a result of seasonal changes in the number of transactions, the most suitable τ was determined to be 12 months.

The comparison among the RHM, URHM, and OPHM ($\tau = 12$ months) showed that there are large fluctuations in the URHM, and that the magnitude of the fluctuations differs significantly from that of actual prices.

The comparison between the URHM and OPHM showed the presence of large differences between the two indices from 1986 through 1990 (the Bubble period in Japan). When the URHM is added to this comparison, we can see random changes in the URHM centered on the values of the OPHM. From the results of the comparison of these three price indices, it was speculated that, because there were large changes in the price structure from 1986 through 1990, the RHM was unable to properly describe such changes, resulting in large differences observed between the URHM and URHM/OPHM (τ = 12 months).

The results of the above series of analyses indicated the superiority of estimation by the OPHM (τ = 12 months) in the previously-owned condominium market in Japan, when structural changes in the market are to be properly taken into consideration.

The results collected in this paper were obtained through ongoing monthly estimation of actual housing price indices in Japan over a period of ten years. We hope that in the future our findings will serve as a reference for those countries that need to develop housing price indices based on the hedonic method.

References;

Bailey MJ, Muth RF, Nourse HO. A Regression Model for Real Estate Price Index Construction, Journal of American Statistical Association 1963; 58; 933−942.

Bourassa, ST, Hoesli M, Sun J. A simple alternative house price index method. Journal of Housing Economics 2006; 15; 80–97.

Case B, Quigley JM. The dynamics of real estate prices. Review of Economics and Statistics 1991; 22; 50–58.

Case KE, Shiller RJ. The efficiency of the market for single-family homes. The American Economic Review 1989; 79; 125–137.

Case B, Pollakowski HO, Wachter SM. On choosing among house price index methodologies. AREUEA Journal 1991; 19; 286–307.

Clapp JM. A semiparametric method for valuing residential locations: application to automated valuation. Journal of Real Estate Finance and Economics 2003; 27; 303–320.

Clapp JM, Giaccotto C. Estimating price trends for residential property: a comparison of repeat sales and assessed value methods. Journal of Real Estate Finance and Economics 1992; 5; 357–374.

Clapp JM, Giaccotto C. Price indices based on the hedonic repeat-sales method: application to the housing market. Journal of Real Estate Finance and Economics 1998; 16; 5–26.

Clapp JM, Giaccotto C, Tirtiroglu D. Housing price indices based on all transactions compared to repeat subsamples. AREUEA Journal 1991; 19; 270–285.

Court AT, Hedonic price indexes with automotive examples, in: The Dynamics of Automobile Demand, General Motors, New York.1939.

Diewert E, The Paris OECD-IMF Workshop on Real Estate Price Indexes: Conclusions and Future Directions. University of British Columbia Department of Economics, Discussion Paper 2007; 07-01.

Rosen S. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition, Journal of Political Economy 1974; 82; 34−55.

Shimizu C, Nishimura KG. Biases in appraisal land price information: the case of Japan. Journal of Property Investment and Finance 2006; 26; 150–175.

Shimizu C, Nishimura KG. Pricing structure in Tokyo metropolitan land markets and its structural changes: pre-bubble, bubble, and post-bubble periods, Journal of Real Estate Finance and Economics 2007; 35: 476-495.

Shimizu, C, Nishimura KG, Asami Y. Search and Vacancy Costs in the Tokyo housing market: Attempt to measure social costs of imperfect information. Regional and Urban Development Studies 2004; 16; 210-230.

Shimizu, C, Nishimura KG, Watanabe T, House Prices and Rents in Tokyo - A Comparison of Repeat-sales and Hedonic measures-, Research Center for Price Dynamics Institute of Economic Research (Hitotsubashi University), Working Paper 2009 ; No.48.

Table 1. List of analyzed data

Symbols	**Variables**	**Unit**	**Average**	**Standard Deviation**
RP	Resale Price of Condominium	10000yen	3,904.66	2,348.54
FS	Floor space/ square meters	m^2	56.57	19.40
AGE	Number of years since construction	year	165.74	91.98
TS	Time to nearest station	minute	7.60	4.27
TT	Travel Time to central business district(terminal stations)	minute	15.32	5.30
BS	Balcony space/ square meters	m^2	8.14	5.96
NU	Number of units	unit	100.03	131.05
RT	Market reservation time	week	11.58	10.62
FD	First floor dummy	(0,1)	-	-
HF	Highest floor dummy	(0,1)	-	-
SD	South-facing dummy	(0,1)	-	-
FD	Ferroconcrete dummy	(0,1)	-	-
LDj (j=0,…,J)	Location (Ward) dummy	(0,1)	-	-
RDk (k=0,…,K)	Railway line dummy	(0,1)	-	-
TDl (l=0,…,L)	Time dummy (monthly)	(0,1)	-	-

*Terminal Staions : Tokyo,Shinagawa,Shibuya,Shinjuku,Ikebukuro,Ueno, and Ootemachi

Table 2. Estimation results of the RHM: 23 wards of Tokyo

Method of Estimation		
OLS		
Dependent Variable		
RP : Resale Price of Condominiums (in log)		
Independent Variables		
Property Characteristics (in log)	**Coefficient**	**t-value**
Constant	4.631	498.230
FS : Floor space	0.013	10.810
Age : Age of building	-0.190	-337.380
TS : Time to the nearest station	-0.078	-99.690
TT : Travel Time to CBD	-0.040	-36.210
NU : Number of units	0.019	40.900
RT : Market reservation time	0.014	32.530
Property Characteristics (dummy variables)	**Coefficient**	**t-value**
BD : Bus Dummy	-0.276	-13.140
TS × *BD*	0.059	6.970
FF : First Floor Dummy	-0.026	-19.210
HF :Highest floor dummy	0.018	8.000
FD :*Ferroconcrete dummy*	-0.010	-10.150
SD :*South-facing dummy*	0.009	10.790
Location (Ward) Dummy: *LDj* (*j* =0,…,*J*)	Yes	
Railway Line Dummy: *LDk* (*k* =0,…,*K*)	Yes	
Time Dummy:*TDl* (*l* =0,…,*L*)	Yes	
Adjusted R square=	0.837	
Number of Observations=	211,179	

Table 3. Statistical values of major regression coefficients (URHM and OPHM)

Pricipal Independent Variables	***RHI* :1986.01 - 2006.09**	***NRHI***		***OPHM***	
		Average	Standard deviation	Average	Standard deviation
FS :Floor space/square meters	0.013	0.033	0.081	0.033	0.079
Age :Age of building	-0.190	-0.185	0.033	-0.185	0.030
TS :Distance to nearest station	-0.078	-0.082	0.019	-0.082	0.015
TT :Travel Time to central business district	-0.040	-0.041	0.032	-0.042	0.023
Adjusted-R Square	0.837	0.741	0.054	0.738	0.049
Number of Samples	211,179	844.720	282.977	10,178.252	2,709.339
		1986.01 - 2006.09		1986.12 - 2006.09	

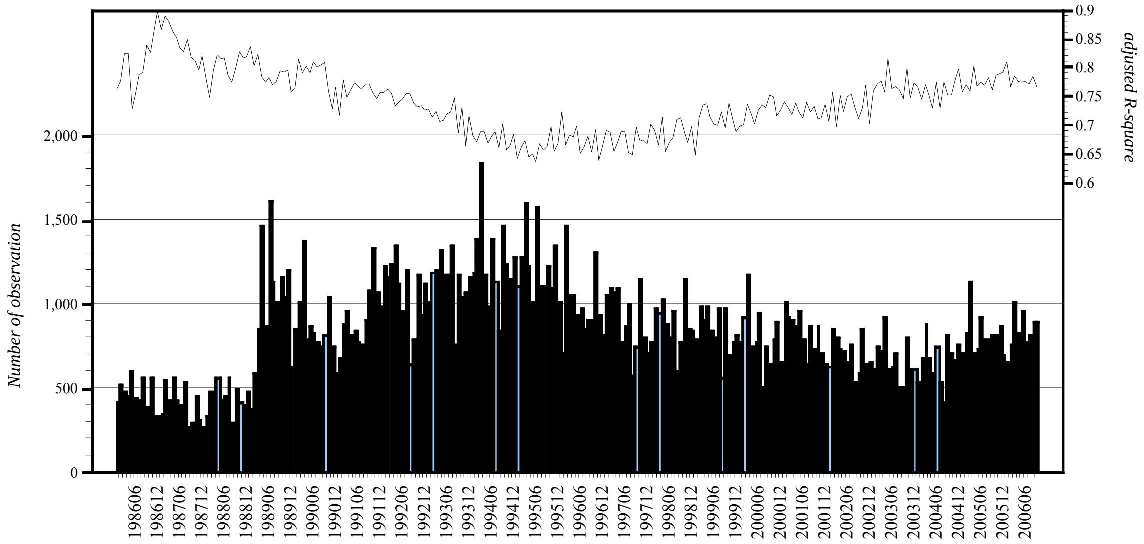

Figure 1. Estimation accuracy of the URHM: between 1986/01 and 2006/09

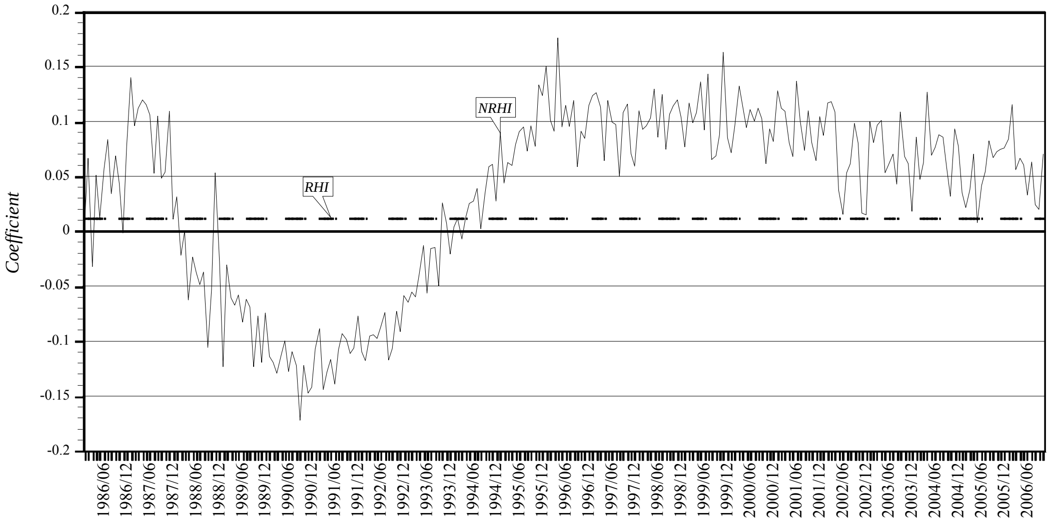

Figure 2. Time profile of regression coefficient of the URHM, floor space *FS*: 1986/01–2006/09

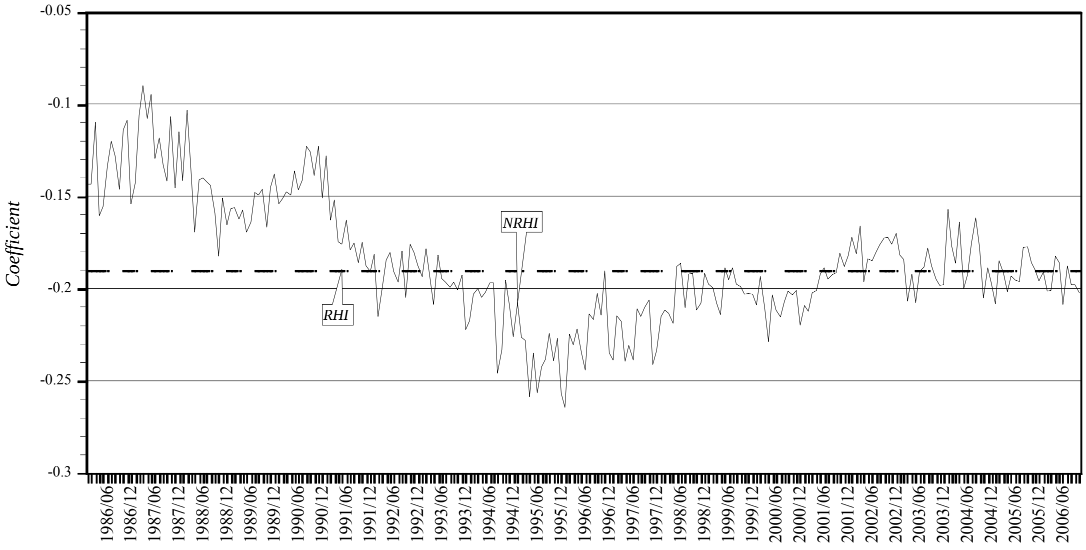

Figure 3.Time profile of regression coefficient of the URHM, age of building *Age*: 1986/01–2006/09.

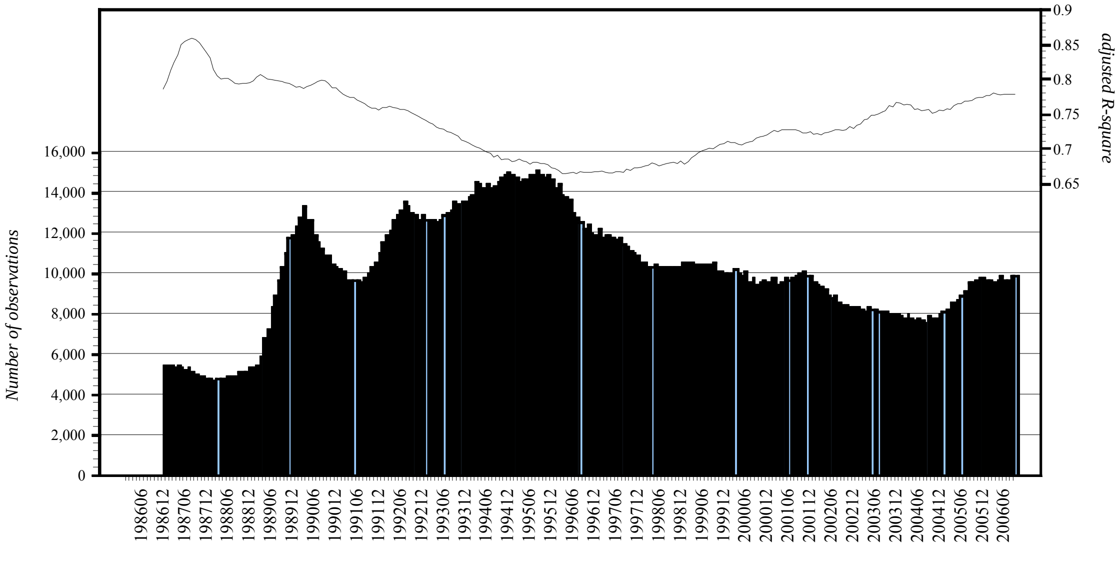

Figure 4. Estimation accuracy of the OPHM: between 1986/01 and 2006/09

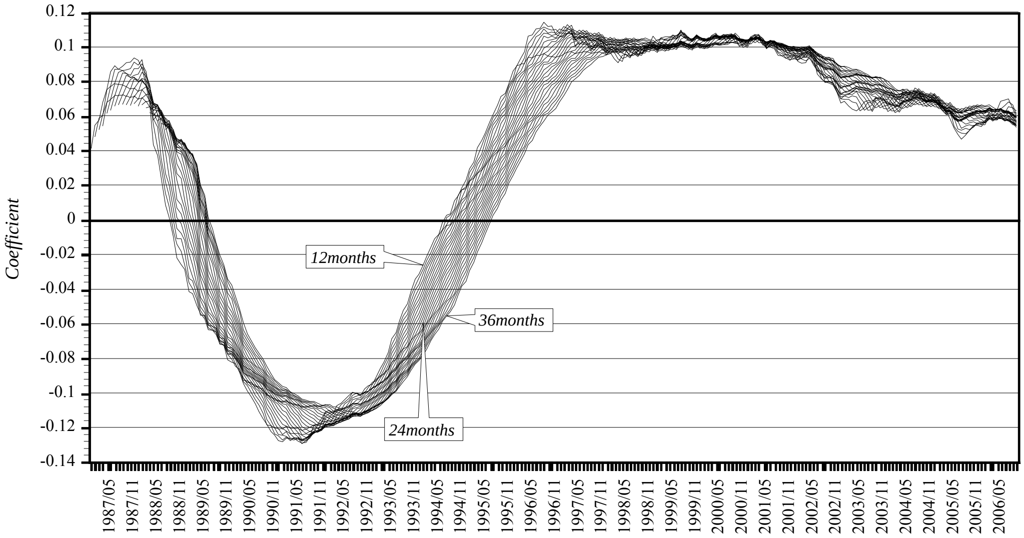

Figure 5. Time profile of regression coefficient of the OPHM, floor space *FS*: 1986/01–2006/09

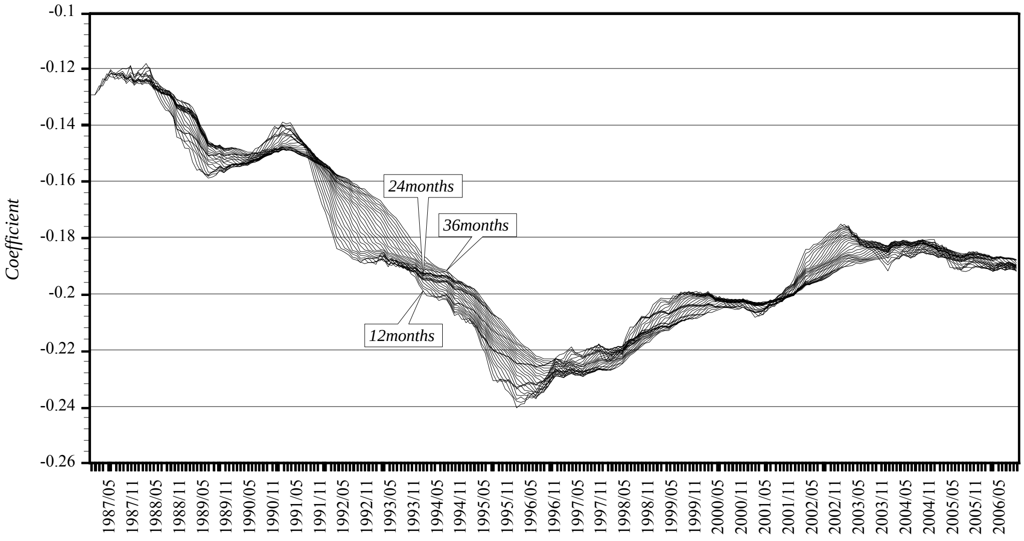

Figure 6. Time profile of regression coefficient of the OPHM, age of building *Age*: 1986/01–2006/09

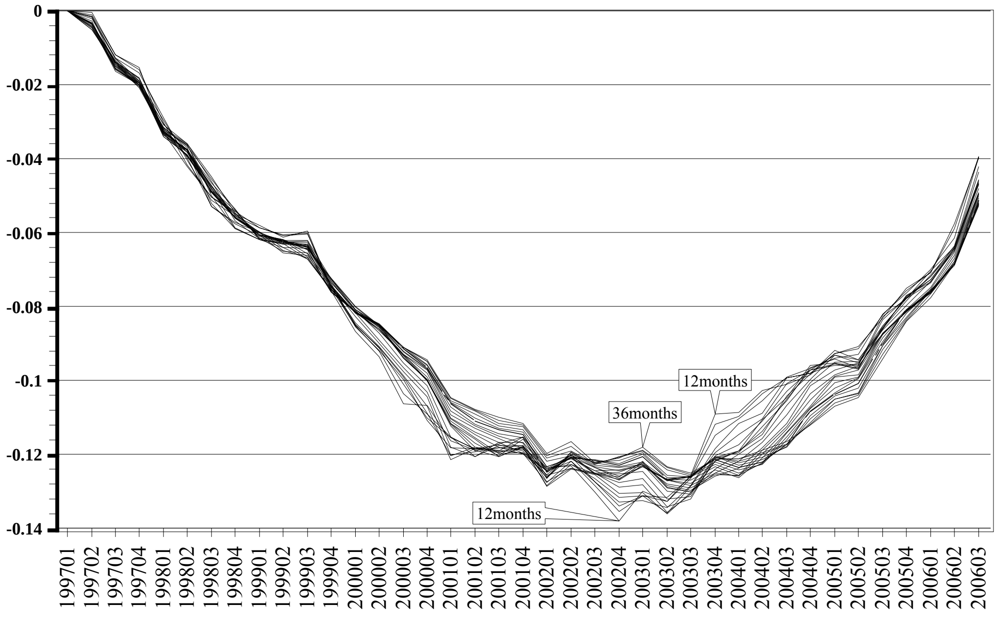

Figure 7. Time profile of the OPHM: 1997/01–2006/09

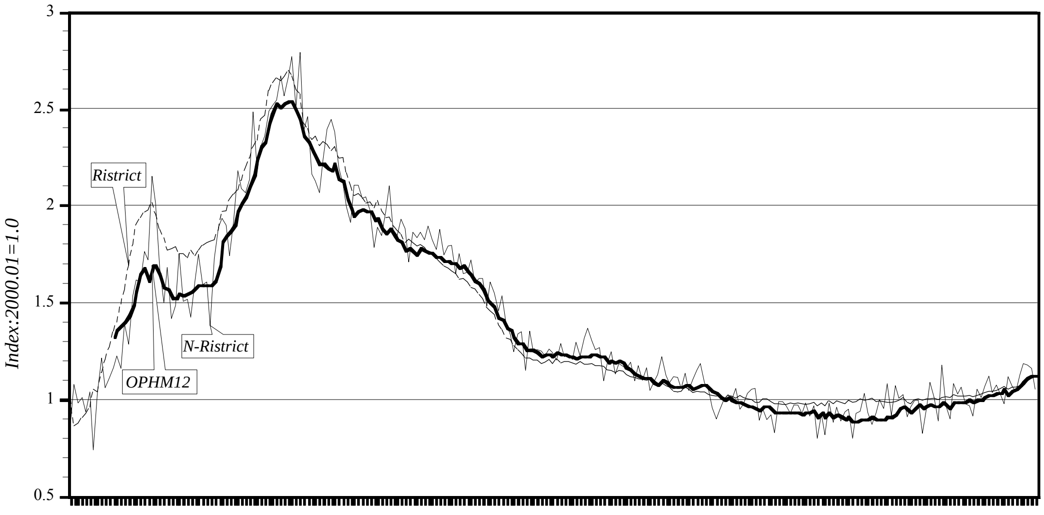

Figure 8. Comparison of indices between the URHM, URHM, and OPHM (τ =12): 1986/01–2006/10