JSPS Grants-in-Aid for Creative Scientific Research

Understanding Inflation Dynamics of the Japanese Economy

Working Paper Series No.41

Housing Prices and Rents in Tokyo: A Comparison of Repeat-Sales and Hedonic Measures

Chihiro Shimizu
Kiyohiko G. Nishimura
and
Tsutomu Watanabe

May 25, 2009

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/

Housing Prices and Rents in Tokyo: A Comparison of Repeat-Sales and Hedonic Measures

Chihiro Shimizu* Kiyohiko G. Nishimura† Tsutomu Watanabe‡

Preliminary Draft

May 21, 2009

Abstract

Do the indices of house prices and rents behave differently depending on the estimation methods? If so, to what extent? To address these questions, we use a unique dataset that we have compiled from individual listings in a widely circulated real estate advertisement magazine. The dataset contains more than 400 thousand listings of housing prices and about one million listings of housing rents, both from 1986 to 2008, including the period of housing bubble and its burst. We find that there exists a substantial discrepancy in terms of turning points between hedonic and repeat sales indices, even though the hedonic index is adjusted for structural change and the repeat sales index is adjusted in a way Case and Shiller suggested. Specifically, the repeat sales measure tends to exhibit a delayed turn compared with the hedonic measure; for example, the hedonic measure of condominium prices hit bottom at the beginning of 2002, while the corresponding repeat-sales measure exhibits reversal only in the spring of 2004. Such a discrepancy cannot be fully removed even if we adjust the repeat sales index for depreciation (age effects).

JEL Classification Number: C43; C81; C82; R21; R31
Keywords: hedonic price index; repeat sales price index; aggregation bias; depreciation problem

*Correspondence: Chihiro Shimizu, International School of Economics and Business Administration, Reitaku University, 2-1-1 Hikarigaoka, Kashiwa, Chiba 277-8686, Japan. E-mail: cshimizu@reitaku-u.ac.jp. 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.

†Deputy Governor, Bank of Japan.

‡Research Center for Price Dynamics, Hitotsubashi University.

1 Introduction

Fluctuations in real estate prices have substantial impacts on economic activities. In Japan, a sharp rise in real estate prices during the latter half of the 1980s and its decline in the early 1990s has led to a decade-long stagnation of the Japanese economy. More recently, a rapid rise in housing prices and its reversal in the United States have triggered a global financial crisis. In such circumstances, the development of appropriate indices that allow one to capture changes in real estate prices with precision is extremely important not only for policy makers but also for market participants who are looking for the time when housing prices hit bottom.

Research has been conducted intensively on methods of compiling housing price indexes appropriately. The location, history and facilities of each house are different from each other in varying degrees, so there are no two houses that are identical in terms of quality. Even if the location and facilities are the same, the age of the building may differ, in which case the degree of deterioration varies accordingly and the houses are not identical. In other words, houses have "particularity with few equivalents." In addition, the quality of houses changes with market conditions, owing to "depreciation" and "renovation." Such characteristics are particularly evident in the housing market of Japan compared with that of the United States and other countries.

There are two approaches to constructing a housing price index that take into account issues resulting from the aforementioned particularity with few equivalents and changes in quality: the hedonic method and the repeat sales method. In this paper, we compare the hedonic method and repeat sales method for estimating the price and rent indices of the housing market in Tokyo.

On one hand, previous research has identified two major problems for the repeat sales method: (i) there is sample selection bias because houses often traded have some special characteristics (Clapp and Giaccotto 1992); (ii) the assumption that there are no changes in property characteristics and their parameters during the transaction period is unrealistic (Case and Shiller 1987, 1989; Clapp and Giaccotto 1992, 1998, 1999; Goodman and Thibodeau 1998; Case et al. 1991). On the other hand, the hedonic method is said to suffer from the following problems: (iii) there is an omitted variable bias (Case and Quigley 1991; Clapp 2003, Ekeland, Heckman and Nesheim 2004); (iv) the assumption of no structural change during the sample period is unrealistic (Case et al. 1991; Clapp et al. 1991; Clapp and Giaccotto 1992, 1998; Shimizu and Nishimura 2006, 2007, Shimizu, Takatsuji, Ono, and Nishimura 2007).

From the theoretical viewpoints, it is not easy to say which one performs better. However, it is often said that, at least from the practical perspective, the repeat sales method is much easier and less costly to be implemented (Bourassa et al. 2006). However, as far as the Japanese housing market is concerned, there are some reasons to worry more about the problems associated with the repeated sales method. First, the

Japanese housing market is illiquid compared with those in other countries including the U.S. and other countries. Second, there were legal restrictions based on "the national land use plan", prohibiting reselling within a short period until the end of 1990s. Given this circumstance, Shimizu et al (2007) advocates adopting the hedonic method in Japan, and proposes an estimation procedure to take into account structural changes and seasonal sample selection bias.

The rest of this paper is organized as follows. Section 2 presents an overview of the two methods. 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 0.46 million listings of housing prices and about 1.1 million listings of housing rents, both from 1986 to 2008, including the period of housing bubble and its burst. Section 4 presents estimation results. We find some evidence for "depreciation problem" with the repeat sales method and "market structural problem" with the hedonic index. More importantly, we find that there exists a substantial discrepancy in terms of turning points between structural-change-adjusted hedonic and Case-Shiller-adjusted repeat sales price indexes. Specifically, the repeat sales measure tends to exhibit a delayed turn compared with the hedonic measure; for example, the hedonic measure of condominium prices hit bottom at the beginning of 2002, while the corresponding repeat-sales measure exhibits reversal only in the spring of 2004. Such a discrepancy cannot be fully removed even if we consider age-effect adjustment to repeat sales indexes. Section 5 concludes the paper.

2 Five Measures of House Prices

2.1 Traditional hedonic index

Let us begin with a hedonic price index. Suppose that we have the price and property-characteristics data of houses, pooled for all periods $t = 1, 2, \cdots, T$, and that the number of data samples in period t is n_t. Then, a traditional hedonic price index is based on the following house-price estimation model:

$$\ln P_{it} = \ln x_{it}\beta_t + \epsilon_{it} \tag{1}$$

where P_{it} is the price of house i in period t, β_t is a vector of parameters associated with residential property characteristics, x_{it} is a vector of property characteristic for house i in period t, and ϵ_{it} is an error term, consisting of time effect and iid disturbance ($\epsilon_{it} \equiv \alpha + \delta_t + v_{it}$ and $v_{it} \sim N(0, \sigma_v^2)$). The model may be described as the "restricted hedonic model" because the regression coefficient β_t of the house-price determining factor x_{it} is constrained to be constant throughout the entire sample period. The traditional hedonic price index is then constructed from the time dummies.

2.2 Traditional repeat sales index

The traditional repeat-sales method starts with the assumption that the residential property characteristics do not change over time and that their effect on price formation does not change either. The underlying price determination model is not different from equation (1). However, the repeat sales method focuses its attention on houses that appear in different periods in the data set, or houses transacted repeatedly. Suppose that house i is transacted twice, in period s and period t. Then the change in the house price is given by

$$\Delta_{t,s} \ln P_i = (\ln x_{it}\beta_t - \ln x_{is}\beta_s) + (\delta_t - \delta_s) + (v_{it} - v_{is}) = (\delta_t - \delta_s) + (v_{it} - v_{is}) \quad (2)$$

indicating that the price change is solely determined by the difference between the two transaction times, irrespective of residential property characteristics. From equation (2) we have:

$$\Delta_{t,s} \ln P_i = D_i\delta + \nu_{its} \quad (3)$$

where $\nu_{its} \equiv v_{it} - v_{is}$ and D_i is a time dummy variable vector, which takes a value of 1 at the second transaction, -1 at the first transaction, and 0 in the other periods. By estimating (3), we get a repeat sales index as $I^{RS} \equiv \{\exp(0), \exp(\hat{\delta}_2), \cdots, \exp(\hat{\delta}_T)\}$.

2.3 Case-Shiller adjustment to repeat sales index

There are two theoretical problems in the traditional repeat sales method just described. One is heteroscedasticity: error terms are likely to become larger when two transaction dates are further apart. The second is the age effect: a house is not qualitatively the same as time goes by. In this subsection, we consider heteroscedascticity.

When estimating the repeat sales index, Case and Shiller (1987, 1989) have proposed a model in which a GLS estimation is performed taking account of heteroscedasticity. Specifically, we assume:

$$E(\nu_{it} - \nu_{is})^2 = \xi_1(t - s) \quad (4)$$

Then, we use the four-step method. The first step is to calculate equation (3). In the second step, we regress the square of the estimated disturbance term on a constant term, ξ_0, and the transaction period, $t-s$. In the third step, we estimate (3) with GLS, with a weight variable of $[\hat{\xi}_0 + \hat{\xi}_1(t-s)]^{1/2}$. In this way, we get a Case-Shiller-adjusted repeat sales index as $I^{WRS} \equiv \{\exp(0), \exp(\tilde{\delta}_2), \cdots, \exp(\tilde{\delta}_T)\}$.

2.4 Age-adjustment to repeat sales index

Let us now turn to the age effect. Most researchers including Bailey et al. (1963) and Case and Shiller (1987, 1989) estimate price indices under the assumption that housing characteristics do not change over time. However, houses actually deteriorate as time goes by. This means that the utility flow of a house diminishes with time. Thus, it is likely that the quality of a house at the time of selling is lower when the house was bought long time ago than when it was purchased only a few periods ago. The house quality also changes with maintenance expenditure and large-scale renovations. Furthermore, the house quality also changes if there are major changes in the surrounding environment, the convenience of public transport, etc. In research related to estimating housing price indices, these problems are known as aggregation bias. Notably, with respect to the repeat-sales method, Diewert (2007) pointed out a "depreciation problem" based on the number of years since construction and a "renovation problem" based on renovations and so forth. In Japan, since the renovation market is not very developed, the renovation problem is considered to have almost no impact; however, since the number of years for which houses in the market is remarkably short, the depreciation problem is potentially significant.

To take account of the age effect, we assume that the housing characteristics vector x can be divided into capital goods K, which have changing properties, and characteristics Z, which do not change. Per McMillen (2003), the house at the time the building is completed is set as and the house τ years after construction is defined as follows:

$$K_t = K_0 \exp^{c\tau} \tag{5}$$

where $c < 0$ indicates the percentage of depreciation parameter for the house. The rate of change in housing capital goods from period s through period t is $\ln K_t - \ln K_s = c(t-s)$. In this case, the repeat sales regression model is $\ln P_{it}/P_{is} = \theta(t-s) + (\delta_t - \delta_s) + \nu_{its}$. This implies that, if one estimates the repeat sales index without considering the age effect, the estimated time effect has a downward bias.

It is impossible to estimate θ and δ simultaneously because of the presence of multicollinearity. A number of methods have been proposed in order to resolve this problem. Here, we follow the method proposed by Karato, Shimizu, Nakagawa, and Harano (2009). We convert the number of years since construction (i.e., age) into a weekly base. The estimated index for the time effect is in monthly units. Then, we define the capital goods at the second transaction time as

$$K_t = K_0 \exp^{c\tau^{\lambda}} \tag{6}$$

and calculate the following repeat sales regression model

$$\Delta_{t,s} \ln P_i = \theta\tau_i^{\lambda} - \theta[\tau_i - (t-s)]^{\lambda} + (\delta_t - \delta_s) + \nu_{its} \tag{7}$$

This gives the age-adjusted repeat sales index.

2.5 Structural-change adjustment to hedonic index

Finally, consider the traditional hedonic model based on equation (1). It is based on the assumption of no changes in the market structure. However, the assumption of no structural change is not plausible for a long period of time, especially in Japan.

A structural change of the market occurs as a result of various external shocks; and in reality, it takes time for such a change to fully transform the market. Thus, regression coefficients should be regarded as changing successively rather than instantaneously. However, the econometric test of structural changes is based on the assumption of multiple but discrete changes. Thus, the test usually divides observation data into several sub-samples with multiple break points, then tests whether coefficients are the same between sub-periods (for example, Shimizu and Nishimura 2006; Shimizu and Nishimura 2007). Then, it is rather difficult to use such an estimation method of structural change in estimating a gradual, rather than discrete structural change. Instead, it may be more appropriate to estimate regression coefficients on the basis of a process of successive changes by taking a certain length as the estimation "window", by shifting this period in a way of rolling regressions, in essence similar to moving averages. This can be called the overlapping-period hedonic housing index (hereafter, also referred to as the OPHI).

We start by replacing equation (1) by

$$\ln P_{it} = \ln x_{it}\beta_t + \epsilon_{i\psi} \tag{8}$$

where ψ represents the width of the window, and $\epsilon_{i\psi} \equiv \alpha + \delta_\psi + \nu_{i\psi}$. Then we repeatedly estimate this for the period $[1, \psi], [2, \psi+1], [3, \psi+2], \cdots, [T-\psi+1, T]$. The OPHM is a traditional hedonic model with respect to a certain period length ψ. Accordingly, the parameter of the time dummy variable represents the price index of each period with the starting period of length ψ as the reference. Thus, price indices can be obtained directly from the traditional model within the period length. With the OPHM, models for all the periods are estimated by successively shifting the period length ψ by one period.

3 Data Properties

3.1 Variables determining housing prices

The price of a house is determined by information on the land and building. Table 1 shows a list of the explanatory variables. As numerical data representing characteristics of the land and building, we have adopted ground area (GA), floor space (FS), and front road width (RW). The age of the building is the period from its construction to the conclusion of the transaction. To take into account whether the house's windows are south-facing or not (note that Japanese are particularly fond of sunshine), we defined a south-facing dummy (SD). Whether a house has an adjacent private road or not is a private-road dummy (PD). To signify whether a transaction is only for land without a building or not, we have a "land-only" dummy (LD).

Furthermore, the convenience of public transportation from each house location is represented by travel time to the CBD (TT) and time to the nearest station (TS). Travel time to the CBD is measured in the following way. First, we defined the CBD. 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 designated a number of terminal stations as being at the center of major business districts. The terminal stations we have chosen include six on the Yamanote Line (Tokyo, Shinagawa, Shibuya, Shinjuku, Ikebukuro, and Ueno), as well as Otemachi Station, which we have designated as the central station on the Tokyo Metro (Teito Rapid Transit Authority). Then, we investigated average day-time travel times from each station to the seven terminal stations and set the minimum value as the travel time to the CBD for each property.

Regarding the time to the nearest station to the house, times are available for three different means of transportation: foot, bus, and car. However, the data for analysis only include houses within walking distance or bus-transportation distance. Therefore, any difference in the distance between the former and the latter is controlled by a dummy variable (bus dummy: BD). In addition, the walking time (in minutes) is recorded when the house/land is within walking distance, while the walking distance from the house/land to the bus stop and the onboard time from the bus stop to the nearest station (in minutes) are recorded in the case of houses/land in a bus-transportation area. The time to the nearest station (TS) is thus defined as (walking time to nearest station)+(walking time to bus stop)+(onboard time from bus stop to nearest station). Then, for a bus-transportation area, the cross-term of the constant dummy variable with the time to the nearest station (TS) is incorporated in the bus dummy.

These variables are factors attributed to the location or buildings that are single-family houses; regional price differences are expected to exist as well. Therefore, we set a ward dummy (WD) to reflect differences in the various districts' quality of public services and prestige. Furthermore, since most residential ground developments in the Tokyo metropolitan area have been carried out along railway lines, the price structure of houses differs along each railway line; accordingly, we defined a railway line dummy (RD). Moreover, we used a time dummy (TD) to control differences in temporal price changes.

3.2 Statistical properties of samples: hedonic samples versus repeat sales ones

Table 2 shows the descriptive statistics for the major variables in hedonic and repeat sales samples. Note that, by construction, repeat sales samples are a subset of hedonic samples. From this table, we see the sample selection bias may be significant in repeat sales samples. The ratio of repeat sales samples to total hedonic samples is 42.7% in the condominium price data, 26.8% in condominium rent data, and only 6.1% in single family house price data.

The average resale price is 38.62 million yen in Hedonic samples, 44.63 million yen in repeat sales samples for a condominium and 79.50 million yen in Hedonic samples, 76.35 million yen in repeat sales samples for a single family house, and the average rent for a condominium is 0.136 million yen in hedonic samples, 0.156 million yen in repeat sales samples. The data includes a wide range of houses from small properties in the studio apartment class to large properties in the so-called 100-million-yen class. When we compare condominium prices and single-family house prices, we see that single-family house prices are 40 million yen or higher.[1]

In comparison between repeat sales samples and hedonic samples, we find a sharp difference between them in single family home data, and we have already pointed out that the ratio of repeat sales samples to hedonic samples is very low (0.061). Thus, there may be a serious selection bias. In condominium price data, we are unable to identify significant differences in characteristics, "Floor space", "Age of building", "Time to the nearest station", and "Travel time to central business district". However, the average price of repeat sales samples is higher by 6million yen than one of hedonic samples. There are similar tendency in the Condominium rent data. It indicates the possibility of sample selection bias in repeat sales transactions; the prices of the repeat sales samples are higher than standard transactions.

4 Estimation Results

4.1 Standard repeat sales measure vs. standard hedonic measure

We run a standard hedonic regression and obtain the result presented in Table 3. The coefficient of determination adjusted for the degrees of freedom is 0.876 in condominium price model, 0.861 on single-family house price model, and 0.895 in condominium rent model, so that the estimated models have a fairly high explanatory power.

[1]Regarding *FS*, the average value is $58.31m^2$ for owned condominiums, $102.53m^2$ for single family houses, and $40.54m^2$ for rental condominiums in hedonic samples, including all properties from single-person households to large-family households. With respect to Age, the average value is 166.82 months (13.90 years) for owned condominiums, 162.19 months (13.51 years) for single-family houses, and 134.09 months (11.17 years) for rental condominiums.

Among the estimated coefficients, the most interesting one is the age effect of the condominium price model, which is -0.186. It is 16 times higher than that of the single-family house price model. This suggests that condominium price model has a "depreciation problem" in the repeat sales index, as indicated by Diewert (2007) and others. Although in single-family house price model and condominium rent model, "depreciation problem" also occurs in estimating repeat sales index, the magnitude of depreciation (age) effect is very small.

When these indices compared in Figures 1a through 1b, repeat sales index is down more than hedonic index in condominium price and single-family house price index, while condominium rent hedonic index is down more than repeat sales index.

4.2 Case-Shiller and age adjustment to repeat sales measure

As explained in the previous section, traditional repeat sales index or Case-Shiller adjustment index assuming constant quality, are not likely to capture physical depreciations occur between transactions. It is especially troublesome in Japan, where depreciation occurs in large scale than in the US and other advanced countries (Shimizu, Nishimura and Karato(2007)). Taking this in mind, secondly, we estimate age adjustment repeat sales index based on equation (7).

In estimating equation (7), we calibrate λ and θ using maximum likelihood estimation. We can calibrate λ and θ in estimating condominium price model and single-family house price model. However, we cannot get the maximum value for the condominium rent data, so we omit it. Age adjustment repeat sales index improves explanatory power substantially compared with traditional repeat sales index and Case-Shiller adjusted one as we can see Schwarz BIC (Table 4).

When these indices compared in Figures 2a through 2c, traditional repeat sales index and Case-Shiller adjusted index are similar tendency in condominium and single-family house price model. For condominium rent index, Case-Shiller adjustment index shift downward to the traditional repeat sales index.

Next, we investigate the age effect by comparing age-adjusted index and other repeat sales indexes of condominium and single-family house price model. Age-adjusted index of condominium price shifts upward to the two repeat sales indices. In contrast, single-family house price index considering age effect does not shift at all. The difference between the two occurs due to the magnitude of age effect. In the hedonic model (Table 3), the coefficient of age in condominium price is -0.186 and the one of single-family house price is -0.037. The magnitude of age effect in condominium price model is larger than the one of the single-family house price model.

4.3 Structural-change adjustment of hedonic measure

The standard hedonic model is called the structurally restricted hedonic model because it assumes that the regression coefficient of the house-price determining factor is constant throughout all the periods. It is the strong assumption that there are no changes in the property characteristics and their parameters during the estimated period. Especially in Japan, Shimizu and Nishimura (2006, 2007) indicate that significant structural change of real estate market before/after during so called Bubble period. This bias caused by market structural change will be more important when estimated period to be longer.

We estimate OPHM as $\psi = 12$months (i.e., 12-months rolling regression), with the model proposed in section 2.3. Table 5 summarizes estimated hedonic parameters. We can see that the time average of coefficients estimated by 12-rolling regression is approximately equals the estimated coefficient by standard hedonic model. However, there are approximately more than two times differences between the maximum and minimum of coefficients estimated by 12-months rolling regression; especially, there are more than five times differences in "Travel time to central business district" in condominium price model and "Floor space" in single family house price model.

These market structural changes will bring bias in estimating hedonic price index. We can see that the comparison between standard hedonic index and OPHM12 index in Figure 3a to 3c. Regarding to condominium price index and single-family house price index, the differences between the standard hedonic and OPHM12 are larger in the period between rapid-growth from 1987 and decline of bubble from 1991 than in other periods. In condominium rent index, the differences between the two are large in the period in the period from 1991 and after 2000. OPHM12 of condominium price and single-family price index is above standard hedonic index, and OPHM12 of condominium rent index is below standard one.

4.4 How much can the difference be reconciled?

The standard hedonic measure suffers from the problem of omitted variables as well as the problem of shifts in structural parameters. We have solved the latter problem, at least partially, by extending the estimating equation from (1) to (8). On the other hand, the standard repeat sales measure faces the problem of non-random sampling and the problem of changes in attributes of a house, such as its aging. We have removed a part of the latter problem by extending the estimating equation from (2) to (7). We now proceed to ask how much the difference between the hedonic and repeat-sales measures has been reconciled through these adjustments.

Figure 4 compares the five indices in terms of the quarterly growth rates for the case of selling prices of condominiums. For example, the horizontal axis in the upper left panel represents the growth rate of the standard repeat-sales index, while the vertical axis represents the growth rate of the Case-Shiller type repeat-sales index. One can clearly see that almost all dots are exactly on the 45 degree line, implying that these two indices are closely correlated with each other. In fact, the coefficient of correlation is 0.990 at the quarterly frequency, and 0.953 at the monthly frequency. Similarly, the lower left panel compares the growth rate of the standard repeat-sales index and the age-adjusted repeat-sales index. Again, almost all dots are on the 45 degree line, indicating a high correlation between the two indices (the coefficient of correlation is 0.989 at the quarterly frequency and 0.953 at the monthly frequency). The two panels suggest that these two adjustments to the standard repeat-sales index are of little quantitative importance, as far as the Japanese housing data is concerned.

Turning to the upper right panel, which compares the standard hedonic index and the standard repeat-sales index, the dots are now scattered along the 45 degree line but not exactly on it, indicating a much lower correlation than before (0.957 at the quarterly frequency and 0.584 at the monthly frequency). Finally, the lower right panel compares the standard repeat-sales index and the rolling hedonic index, showing that the two indices are correlated even more weakly (0.910 at the quarterly frequency and 0.513 at the monthly frequency). These two panels suggest that the role of rolling regression in eliminating the discrepancy between the hedonic and the repeat-selaes indices may not be so large.

To investigate dynamic relationships between the five indices, we conduct pairwise Granger causality tests. The results for condominiums and single family houses are presented, respectively, in the upper and lower panels of Table 6. The number in each cell represents the p-value associated with the null hypothesis that the index on the column does not Granger-cause the index on the row. For example, the number in the cell of the third column and the second row, 0.2018, represents the p-value associated with the null hypothesis that the Case-Shiller type repeat-sales index does not cause the standard repeat-sales index. Cells shaded by blue color indicate that the corresponding p-value is smaller than one percent. The panel for condominiums shows that one can easily reject the null that the standard hedonic index does not cause the other four indices. On the other hand, one cannot reject the null that each of the other four indices does not cause the standard hedonic index. These two results indicate that fluctuations in the standard hedonic index tend to precede those in the other four indices. The same property is observed for the selling prices of single family houses, as well as the housing rents of condominiums.

To illustrate such lead-lag relationships between the five indices, we compare them in terms of the timing in which each index hit bottom after the bursting of the housing bubble in the early 1990s. The result is presented in Figure 5. We see that all of the three repeat-sales indices hit bottom simultaneously on the first quarter of 2004. In contrast, the two hedonic indices hit bottom on the first quarter of 2002, indicating that a turn in the hedonic indices preceded the one in the repeat-sales indices by two years.

One may wonder where such lead-lag relationships between the hedonic and repeat-sales indices comes from. The first possibility is related to the omitted variable problem in the hedonic indices. The variables omitted in hedonic regressions might move only with some lags relative to the other variables. If this is the case, the hedonic indices, which ignore those variables, could precede the repeat-sales indices. The second possibility is related to the sample selection problem in the repeat-sales indices. As we have seen in Table 2, the fraction of the sample employed in producing the repeat-sales indices is very limited, and more importantly, it might be biased in the sense that the employed sample consists of housing units whose prices and rents exhibit a delayed response to various shocks.

How can we learn more about mechanism behind the lead-lag relationships? One way to identify a factor behind the relationships is to apply hedonic regression to the repeat-sales sample (i.e. the sample employed in producing the repeat-sales indices). The new hedonic index produced in this way differs from the standard one in terms of the employed sample, while they are identical in terms of the list of explanatory variables, so that they commonly suffer from the problem of omitted variables. Therefore, any difference between the new and standard hedonic indices can be regarded as stemming from the difference in employed samples. For example, if we still observe a lead-lag relationship between the new and standard hedonic indices, it implies that the relationship comes from the sample selection bias in repeat-sales indices. This is our future task.

5 Conclusion

In this study, for considering aggregation bias in estimating house price indexes, we estimated three types of repeat sales index and two types of hedonic index for both condominium and single family house prices. We also calculated the four types of rent index except for the repeat sales rent index reflecting the age effect.

One may say that a rolling hedonic regression solves the problem of structural shifts in parameters, and thus yielding a price index which is better than the traditional hedonic index. At the same time, we can see the age-adjusted cum Case-Shiller adjusted repeat sales index is better than other repeat sales indices, i.e. traditional repeat sales index and the original Case-Shiller index.

However, the rolling hedonic index is still not free from the bias coming from omitted variables bias, which is not shared with the repeat sales methodology. On the other hand, the repeat sales model is not immune to the issue of structural change as well as the sampling selection bias. So neither approach is perfect, which means that we should be aware of the defects of each approach especially when we put the indices in the policy context.

First, it is possible that there would be a significant lag in identifying when the housing market starts to recover. For example, the effect caused by depreciation (age effect) is not reflected in the repeat sales index used in the U.S. and other various countries. We find significant lags in oft-cited Case-Shiller-adjusted repeat sales index compared to the hedonic type indices.

Second, the choice of a particular index affects economic policy because the housing rent is a major determinant of the consumer price index. In Japan, the house rent is surveyed through fixed-point observation of the same properties. Its methodology is similar to the repeat sales methodology in terms of the survey methodology and hence it has the bias (depreciation problem) pointed out in this study.

Finally, there is an issue of data collection. We cannot apply the hedonic methodology unless we could collect the characteristic data of each housing transactions even if the methodology could overcome the depreciation problem. It would take more time and cost if they would collect data and maintain the database. Consequently we have to consider the balance of the costs and the accuracy of the indices in the methodology selections.

References

[1] Bailey, M.J., R.F. Muth, and H.O. Nourse (1963), "A Regression Model for Real Estate Price Index Construction," Journal of the American Statistical Association **58**, 933-942.

[2] Bourassa, S.T, Martin Hoesli and Jian Sun (2006), "A Simple adternative house price index method," Journal of Housing Economics **15**, 80-97.

[3] Cannaday, R.E., H.J. Munneke, and T.T. Yang (2005), "A Multivariate Repeat-Sales Model for Estimating House Price Indices," Journal of Urban Economics **57**, 320-342.

[4] Case, B., H.O. Pollakowski, and S.M. Wachter (1991), "On Choosing among House Price Index Methodologies," AREUEA Journal **19**, 286-307.

[5] Case, B., and J.M. Quigley (1991), "The Dynamics of Real Estate Prices," Review of Economics and Statistics **22**, 50-58.

[6] Case, K.E., and R.J. Shiller (1987), "Prices of Single Family Homes since 1970: New Indexes for Four Cities," New England Economic Review Sept./Oct., 45-56.

[7] Case, K.E., and R.J. Shiller (1989), "The Efficiency of the Market for Single Family Homes," The American Economic Review **79**, 125-137.

[8] Chau, K.W., S.K. Wong, and C.U. Yiu (2005), "Adjusting for Non-Linear Age Effects in the Repeat Sales Index," Journal of Real Estate Finance and Economics **31**, 137-153.

[9] Clapp, J.M., and C. Giaccotto (1992), "Estimating Price Trends for Residential Property: A Comparison of Repeat Sales and Assessed Value Methods," Journal of Real Estate Finance and Economics **5**, 357-374.

[10] Clapp, J.M., and C. Giaccotto (1998), "Price Indices Based on the Hedonic Repeat-Sales Method: Application to the Housing Market," Journal of Real Estate Finance and Economics **16**, 5-26.

[11] Court, A.T. (1939), "Hedonic Price Indexes with Automotive Examples," in The Dynamics of Automobile Demand, General Motors, New York.

[12] Coulson, N.E., and D.P. McMillen (2008), "Estimating Time, Age and Vintage Effects in Housing Prices," Journal of Housing Economics **17**, 138-151.

[13] Crone, T.M., and R.P. Voith (1992), "Estimating House Price Appreciation: A Comparison of Methods," Journal of Housing Economics **2**, 324-338.

[14] Diewert, W. E. (2007), "The Paris OECD-IMF Workshop on Real Estate Price Indexes: Conclusions and Future Directions," The University of British Columbia, Department of Economics, Discussion Paper 07-01.

[15] Dombrow, J., J.R. Knight, and C.F. Sirmans (1997), "Aggregation Bias in Repeat-Sales Indices," Journal of Real Estate Finance and Economics **14**, 75-88.

[16] Ekeland, I., J.J. Heckman, and L. Nesheim (2004), "Identification and Estimation of Hedonic Models," Journal of Political Economy **112**, 60-109.

[17] Goodman, A.C. and T.G.Thibodeau (1998), "Dwelling Age Heteroskedasticity in Repeat Sales House Price Equations," Real Estate Economics **26**, 151-171.

[18] Hill, R.C., J.R. Knight, and C.F. Sirmans (1997), "Estimating Capital Asset Price Indexes," Review of Economics and Statistics **79**, 226-233.

[19] McMillen, D.P. (2003), "The Return of Centralization to Chicago: Using Repeat Sales to Identify Changes in House Price Distance Gradients," Regional Science and Urban Economics **33**, 287-304.

[20] Shimizu, C. (2009), "Estimation of Hedonic Single-Family House Price Function Considering Neighborhood Effect Variables," CSIS Discussion Paper: The University of Tokyo, No.93.

[21] Shimizu, C., and K.G. Nishimura (2006), "Biases in Appraisal Land Price Information: The Case of Japan," Journal of Property Investment and Finance **26**, 150-175.

[22] Shimizu, C., and K.G. Nishimura (2007), "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 **35**, 475-496.

[23] Shimizu, C., K.G.Nishimura and K.Karato (2007), "Nonlinearity of Housing Price Structure-Secondhand Condominium Market in Tokyo Metropolitan Area-," CSIS Discussion Paper, University of Tokyo, No.86. Available at http://www.csis.u-tokyo.ac.jp/dp/86.pdf

[24] Shimizu, C., K.G.Nishimura, and T.Watanabe (2008), "Residential Rents and Price Rigidity: Micro Structure and Macro Consequences," Research Center for Price Dynamics Working Paper Series, No.29. Available at http://www.ier.hit-u.ac.jp/ ifd/doc/IFD_WP29.pdf

[25] Shimizu, C., H. Takatsuji, H. Ono, and K.G. Nishimura (2007), "Change in House Price Structure with Time and Housing Price Index," RIPESS (Reitaku Institute of Political Economics and Social Studies) Working Paper, No.25. Available at http://ripess.reitaku-u.ac.jp/report/doc/ 2009020715164991_0.pdf

Table 1: List of variables

Symbol	Variable	Contents	Unit
GA	Ground area / square meters	Ground area	m^2
FS	Floor space / square meters	Floor space	m^2
RW	Front road width	Front road width	10cm
Age	Number of years since construction	Period between the date when the data are deleted from the magazine and the date of construction of the building	months
TS	Time to nearest station	Time to nearest station (by foot or bus)	minutes
TT	Travel time to central business district	Minimum day-time train travel time to nearest of 7 terminal stations in 2000	minutes
BD	Bus dummy	The time includes bus travel time 1, does not include bus travel time 0.	(0,1)
FD	First floor dummy	The property is on the ground floor 1, on another floor 0.	(0,1)
SD	South-facing dummy	Windows facing south 1, Facing other directions 0.	(0,1)
LDk (k=0,…,K)	Location (ward) dummy	*k*-th administrative district 1, other district 0.	(0,1)
TDm (m=0,…,M)	Time dummy (monthly)	*m*-th month 1, other month 0.	(0,1)

Table 2: Summary statistics of housing prices and rents

Variables	Condominium(Asset Price)		Single Family(Asset Price)		Condominium(Rent Price)	
	Hedonic	Repeat Sales	Hedonic	Repeat Sales	Hedonic	Repeat Sales
Price of Condominium/Single Family (10,000 Yen) , Rent of Condominium(Yen/month)	3,862.26 (3,190.83)	4,463.43 (4284.10)	7,950.65 (8275.04)	7,635.24 (7055.96)	136,229.50 (116,436)	156,260.13 (122366.20)
FS : Floor space (m²)	58.31 (21.47)	59.54 (24.09)	102.53 (43.47)	105.82 (45.60)	40.54 (26.63)	44.97 (26.84)
GA : Ground Area (m²)	- -	- -	108.20 (71.19)	101.41 (63.17)	- -	- -
Age :Age of Building(months)	166.82 (101.17)	180.20 (101.35)	162.19 (102.66)	63.79 (99.39)	134.09 (89.27)	105.35 (80.86)
TS : Time to the nearest station: (minutes)	7.96 (4.43)	7.77 (4.28)	9.85 (4.54)	9.60 (4.37)	7.28 (4.03)	6.92 (3.92)
TT : Travel Time to Central Business District (minutes)	12.58 (7.09)	10.73 (6.88)	13.23 (6.34)	11.89 (6.18)	10.20 (6.48)	9.28 (6.28)
1986/01-2008/12	n=157,627	n=67,436	n=315,791	n=19,428	n=1,139,043	n=305,557

Table 3: Results of hedonic regressions

Estimation Method OLS
Dependent Variable Logarithm of Price or Rent per ㎡
Independent Variables

Variables	Condominium(Price)		Single Family(Price)		Condominium(Rent)	
	coefficient	t-value	coefficient	t-value	coefficient	t-value
Constant	4.470	358.778	4.615	378.620	8.951	826.498
FS : Floor space (㎡)	0.029	25.340	0.002	125.046	-0.191	-762.560
GA : Ground Area (㎡)	-	-	-0.002	-213.860	-	-
Age :Age of Building(months)	-0.186	-351.585	-0.011	-190.559	-0.037	-466.013
TS : Time to the nearest station: (minutes)	-0.069	-92.747	-0.013	-137.989	-0.052	-230.110
Bus:Bus Dummy	-0.137	-6.603	-0.198	-24.595	-0.010	-3.350
Bus×TS	0.007	0.815	0.002	4.300	0.018	13.690
TT : Travel Time to Central Business District	-0.068	-68.028	-0.009	-114.091	-0.077	-261.313
Top:Top of Building	0.022	5.390	-	-	-	-
BeforeConstruction Standard	-0.090	-80.770	-	-	-0.122	-256.050
Steel Dummy	0.010	10.650	-	-	0.082	200.050
Balcony Area	0.022	32.950	-	-	-	-
Road Width	-	-	0.207	154.500	-	-
Private Road	-	-	-0.003	-9.840	-	-
Land only Dummy	-	-	-0.109	-63.180	-	-
Old house	-	-	-0.086	-36.020	-	-
New Construction	-	-	-0.121	-69.330	-	-
1986/01-2008/12	n=157,627		n=315,791		n=1,139,043	
Adjusted R-square=	0.876		0.861		0.895	

Table 4: Age adjustment to repeat sales measure

	standard error of reg.	adjusted R-square	S.B.I.C
(Condominium Price)			
Traditional_RS	0.1752	0.7494	-20896.9
Case&Shiller	0.1914	0.7587	-13583.5
Age-adjustment	0.0075	-	-28817.2
(Single Family House Price)			
Traditional_RS	0.2115	0.4756	-2755.8
Case&Shiller	0.2190	0.5093	-1783.2
Age-adjustment	0.0244	-	-3574.8
(Condominium Rent)			
Traditional_RS	0.0627	0.1385	-190929.0
Case&Shiller	0.0684	0.1454	-176623.0
Age-adjustment	-	-	-

	θ	λ
Condominium Price		
coef.	-0.0272	0.8944
s.e.	0.0015	0.0113
P-value	[.000]	[.000]
Single Family		
coef.	-0.0093	1.1041
s.e.	0.0031	0.0269
P-value	[.003]	[.000]

Table 5: Hedonic estimates of key parameters

	Constant	*FS* : Floor space	*Age* :Age of Building	*TS* : Time to the nearest station:	*TT* : Travel Time to Central Business District
Condonimium Price					
Standard Hedonic mode	4.470	0.029	-0.186	-0.069	-0.068
12-months rolling regression					
Average	4.852	0.047	-0.182	-0.072	-0.072
Standard Deviation	0.629	0.078	0.029	0.010	0.031
Minimum	4.193	-0.124	-0.237	-0.098	-0.130
Maximum	6.171	0.133	-0.108	-0.050	-0.022
Single Family Price					
Standard Hedonic mode	4.615	0.002	-0.011	-0.013	-0.009
12-months rolling regression					
Average	4.912	0.002	-0.012	-0.013	-0.009
Standard Deviation	0.261	0.001	0.001	0.002	0.002
Minimum	4.596	0.001	-0.015	-0.019	-0.012
Maximum	5.425	0.003	-0.009	-0.009	-0.004
Condonimium Rent					
Standard Hedonic mode	8.951	-0.191	-0.037	-0.052	-0.077
12-months rolling regression					
Average	9.132	-0.178	-0.042	-0.059	-0.081
Standard Deviation	0.117	0.037	0.015	0.016	0.014
Minimum	8.884	-0.224	-0.071	-0.090	-0.111
Maximum	9.312	-0.092	-0.018	-0.028	-0.054

number of models= 265

Table 6: Pairwise Granger-causality tests

Condominium	Standard repeat sales	Case-Shiller repeat sales	Age-adjusted Repeat sales	Standard hedonic	Rolling hedonic
Standard repeat sales		0.0120	0.0019	0.0037	0.0000
Case-Shiller RS	0.2018		n.a.	0.0411	0.0000
Age-adjusted RS	0.0568	n.a.		0.1067	0.0000
Standard hedonic	0.0005	0.0001	0.0000		0.0000
Rolling hedonic	0.0067	0.0095	0.0025	0.2209	

Single family house	Standard repeat sales	Case-Shiller repeat sales	Age-adjusted Repeat sales	Standard hedonic	Rolling hedonic
Standard repeat sales		0.2726	0.4345	0.2119	0.0040
Case-Shiller RS	0.2397		n.a.	0.1714	0.0098
Age-adjusted RS	0.3275	n.a.		0.1622	0.0078
Standard hedonic	0.0028	0.0025	0.0023		0.0018
Rolling hedonic	0.0705	0.0642	0.0709	0.1642	

Note: The number in each cell represents the p-value associated with the null hypothesis that the variable on the column does not Granger-cause the variable on the row. Cells shaded by blue color indicate that the p-value is smaller than 0.01, and thus the null hypothesis is rejected.

Figure 1a: Standard Hedonic *vs*. Traditional Repeat Sales (Condominium price)

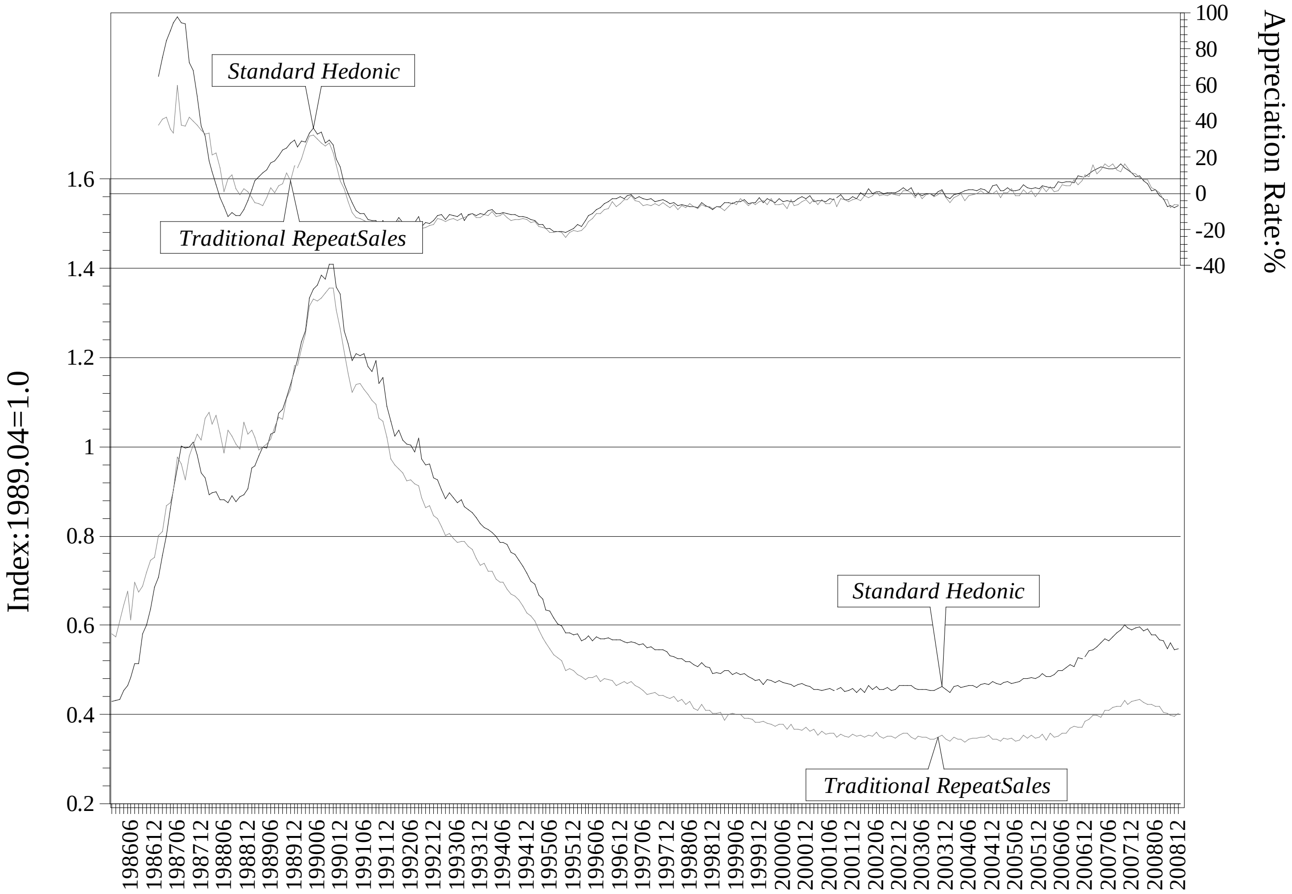

Figure 1b: Standard Hedonic *vs*. Traditional Repeat Sales (Single family house price)

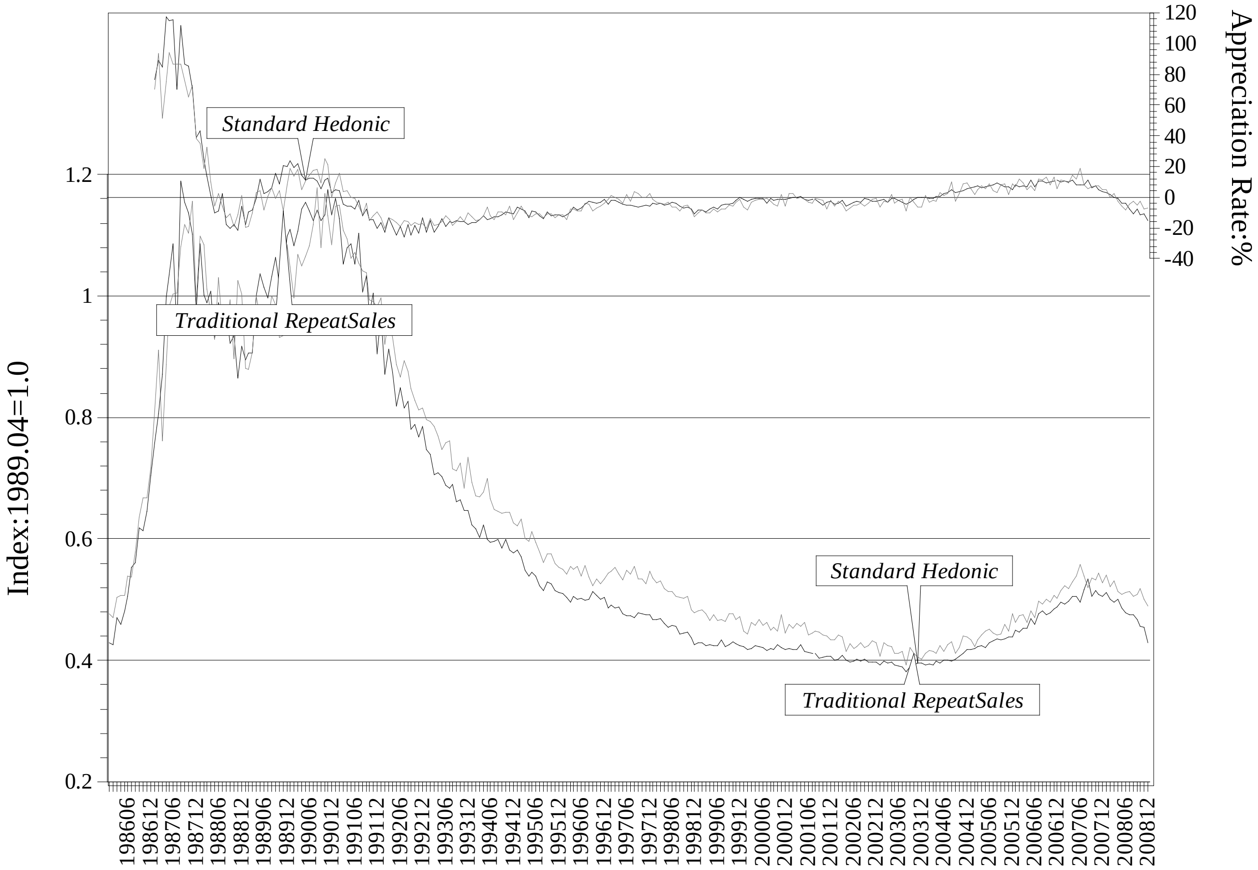

Figure 1c: Standard Hedonic *vs*. Traditional Repeat Sales (Condominium rent)

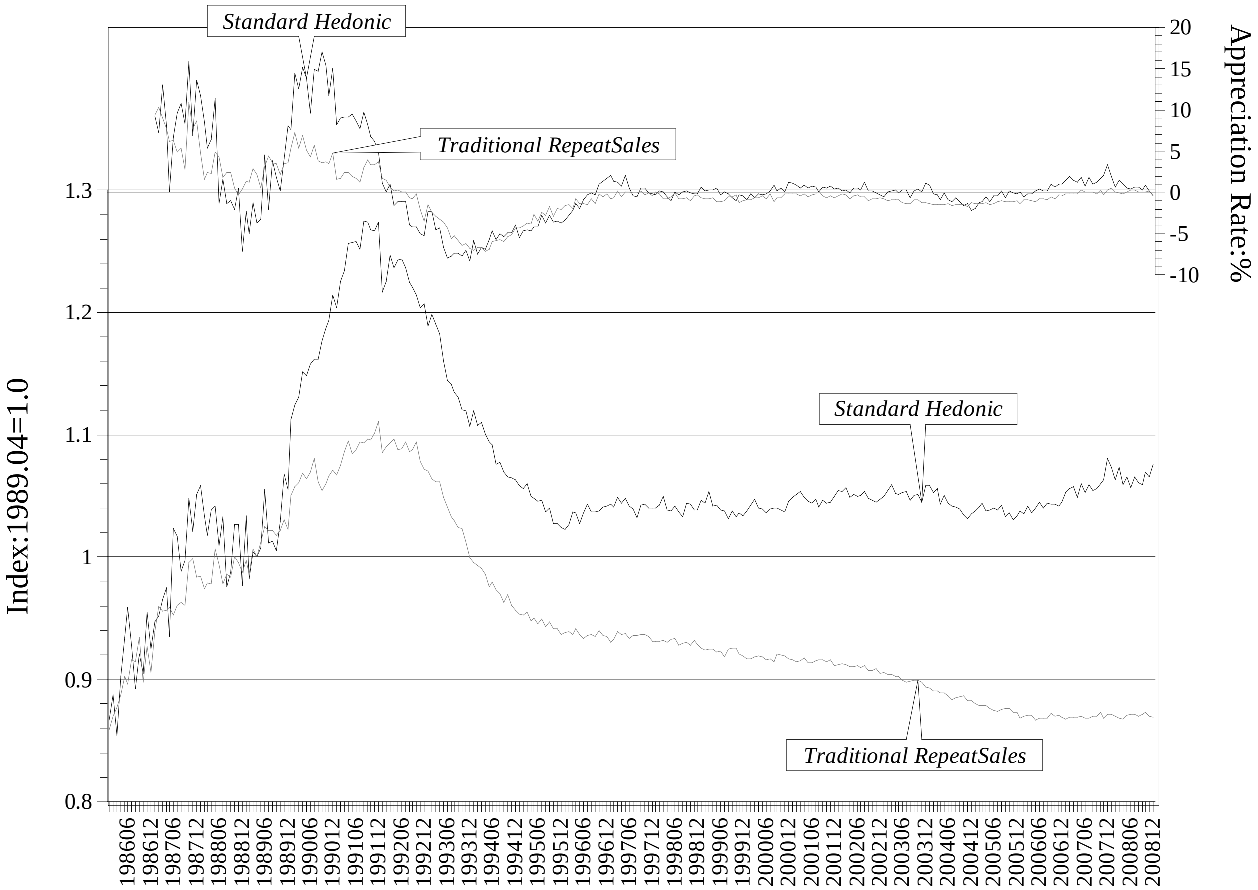

Figure 2a: Case-Shiller adjustment *vs*. Age adjustment Repeat Sales (Condominium price)

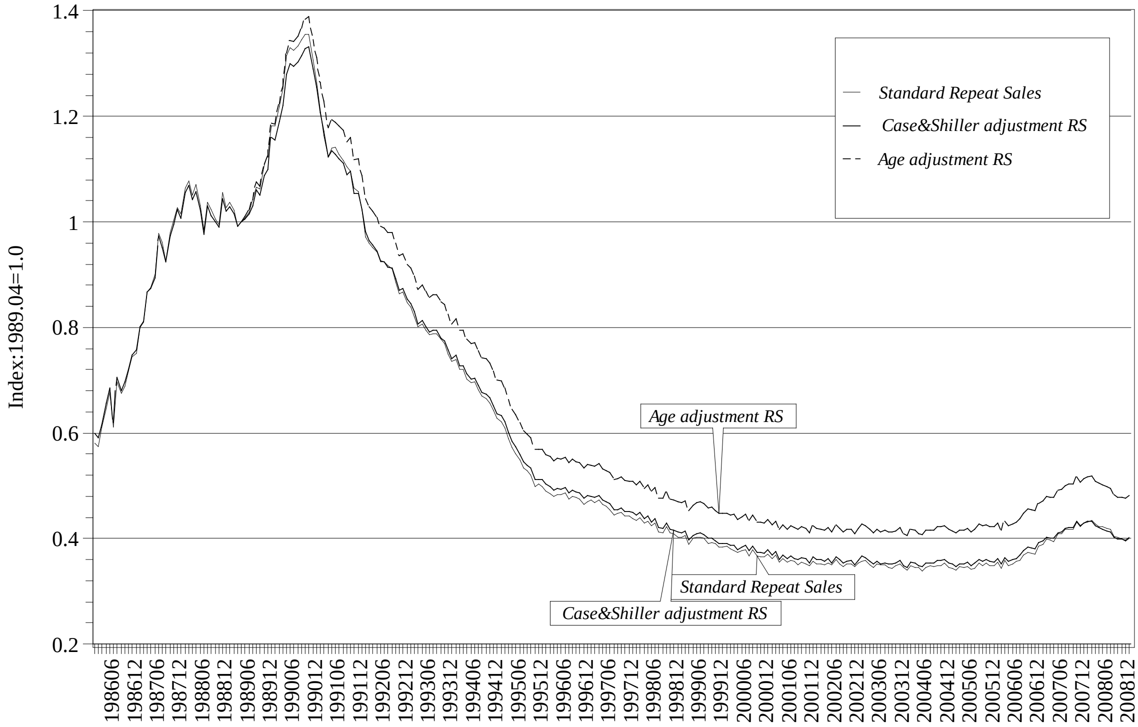

Figure 2b: Case-Shiller adjustment *vs*. Age adjustment Repeat Sales (Single family house price)

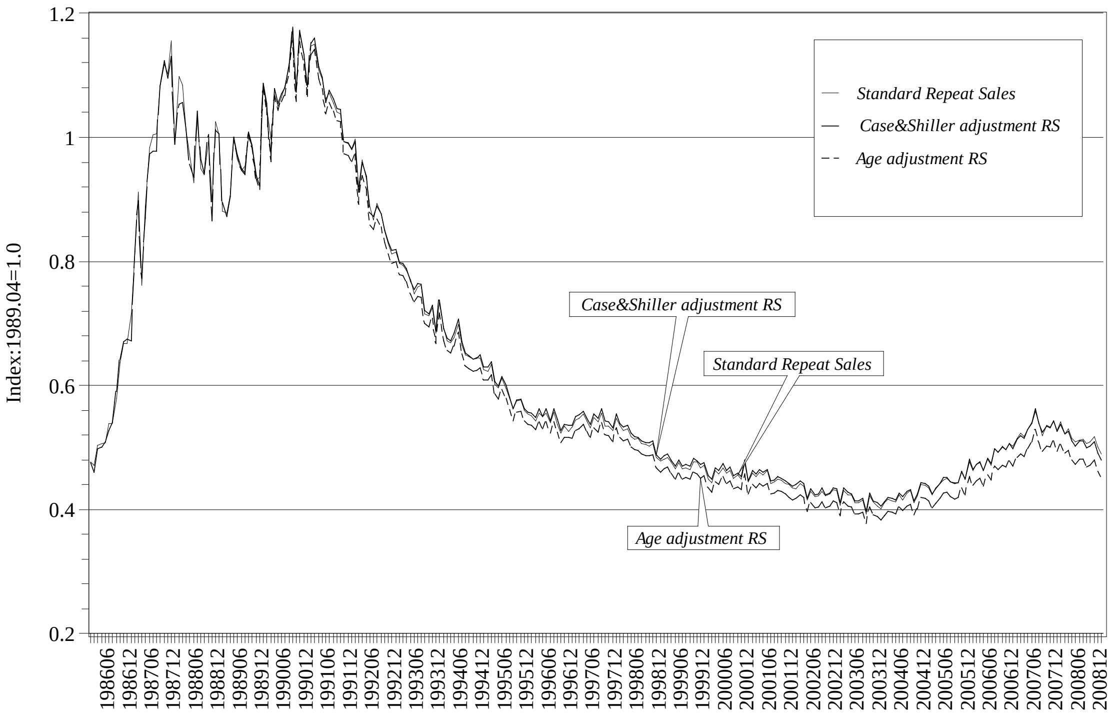

Figure 2c: Case-Shiller adjustment (Condominium rent)

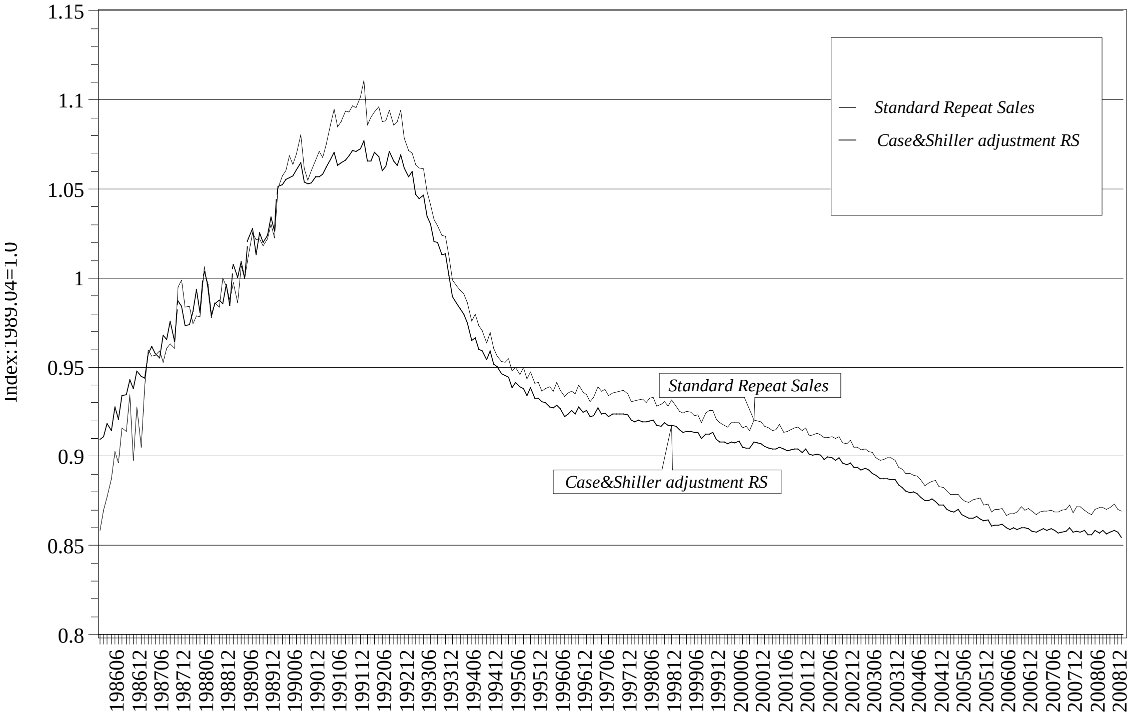

Figure 3a: Standard Hedonic *vs*. 12-months rolling regression (OPHM) (Condominium price)

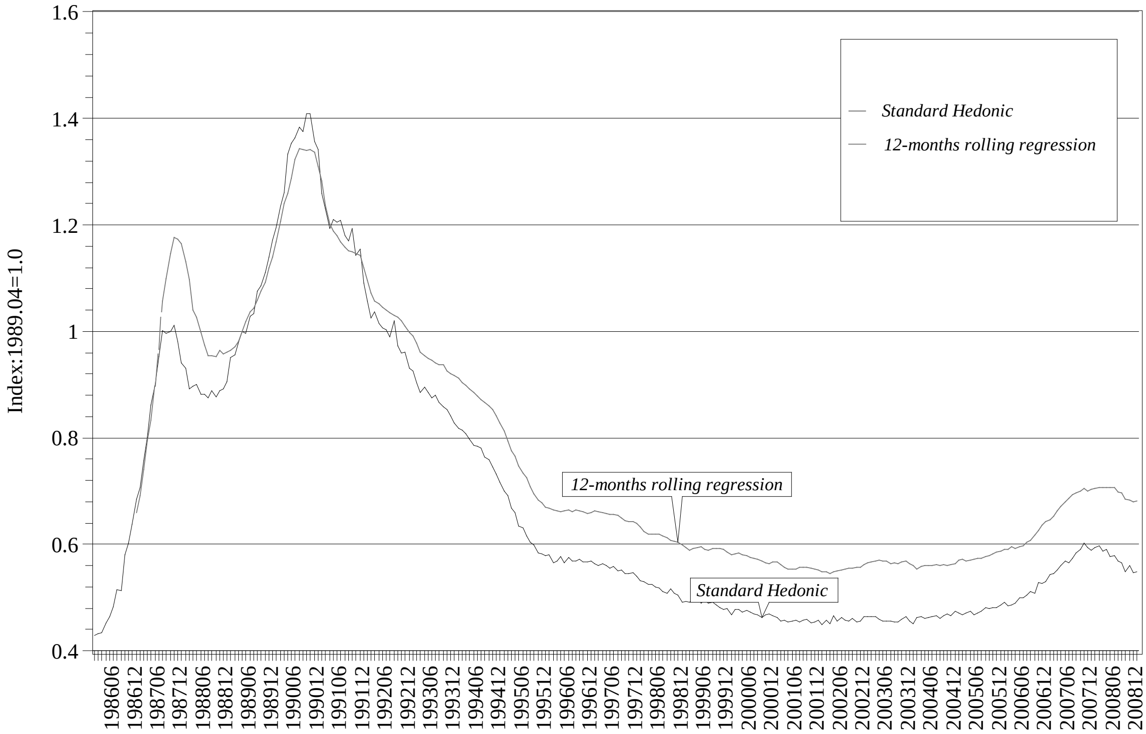

Figure 3b: Standard Hedonic *vs*. 12-months rolling regression (OPHM)
(Single family house price)

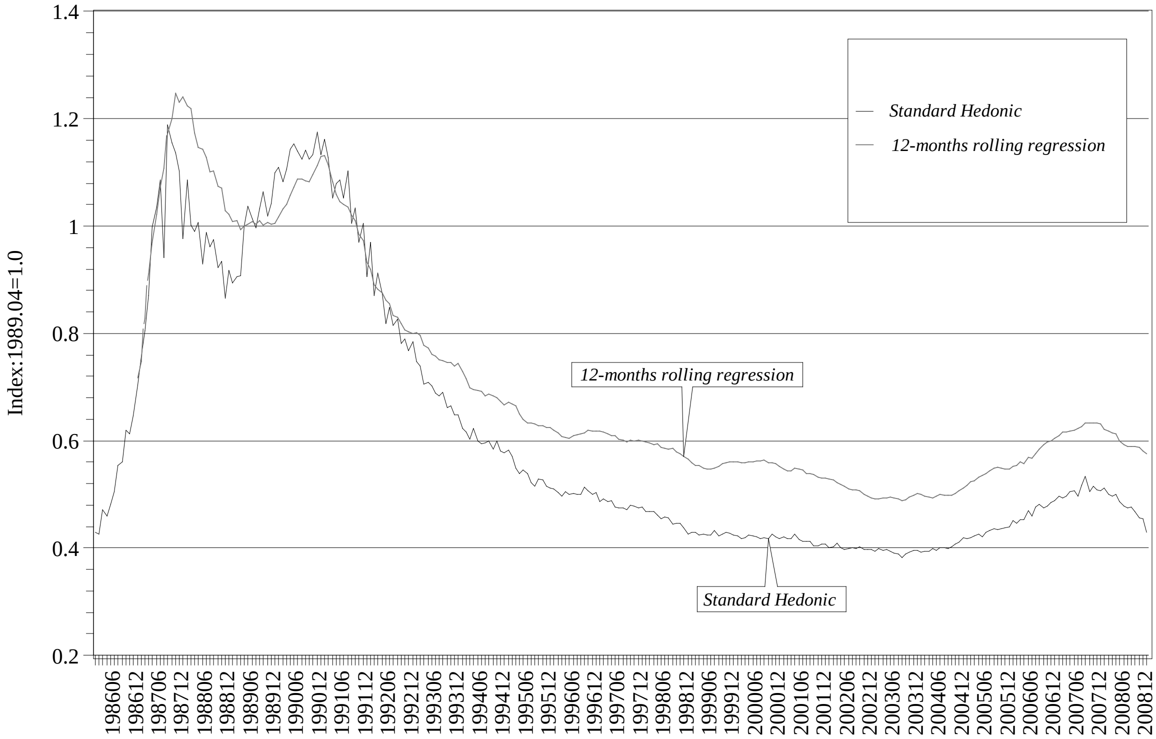

Figure 3c: Standard Hedonic *vs*. 12-months rolling regression (OPHM) (Condominium rent)

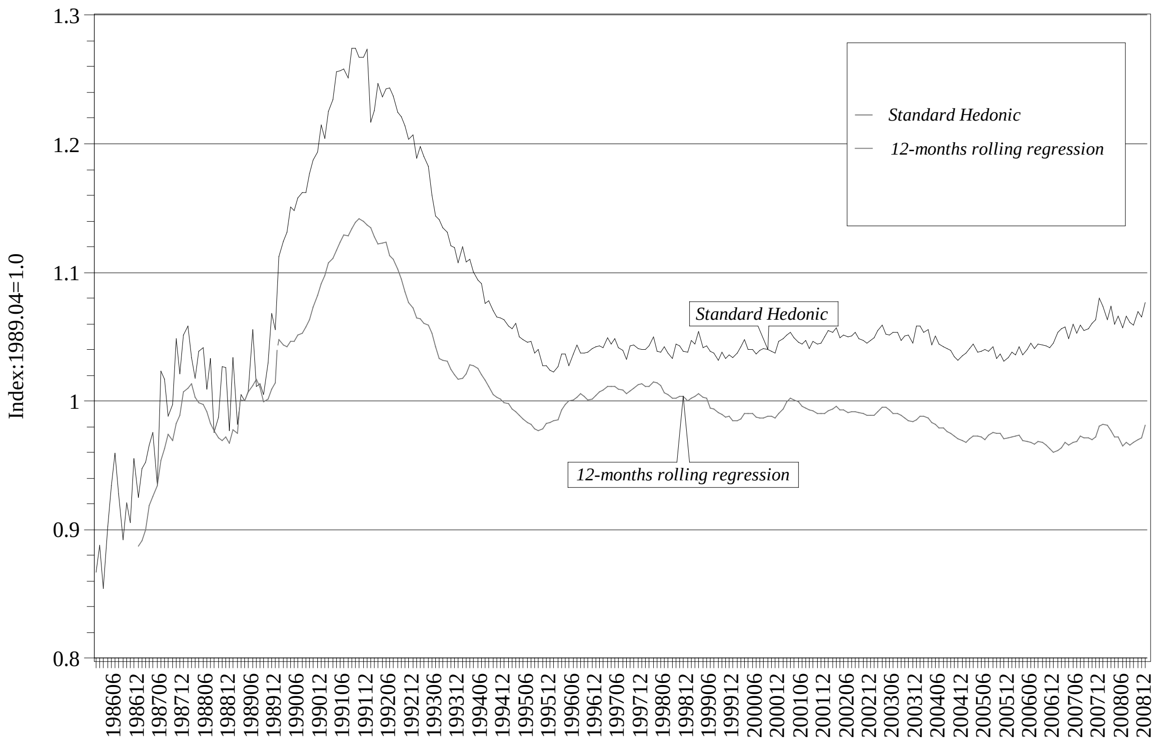

Figure 4: Comparison of the five indices in terms of the quarterly growth rate

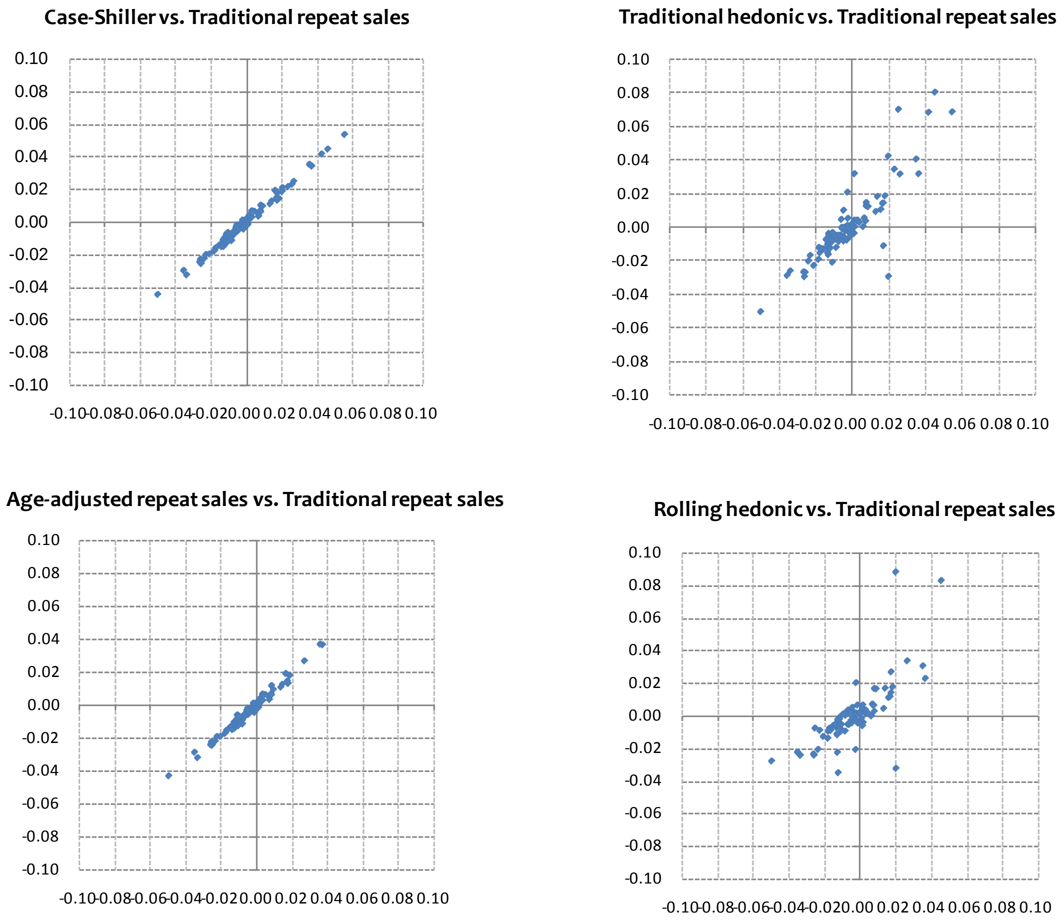

Figure 5: When did the condominium price hit bottom?