CARF Working Paper

CARF-F-313

Detecting Real Estate Bubbles: A New Approach Based on the Cross-Sectional Dispersion of Property Prices

Takaaki Ohnishi
The Canon Institute of Global Studies
The University of Tokyo
Takayuki Mizuno
National Institute of Informatics
Chihiro Shimizu
Reitaku University
Tsutomu Watanabe
The University of Tokyo

April 2013

CARF is presently supported by Bank of Tokyo-Mitsubishi UFJ, Ltd., Dai-ichi Mutual Life Insurance Company, Meiji Yasuda Life Insurance Company, Nomura Holdings, Inc. and Sumitomo Mitsui Banking Corporation (in alphabetical order). This financial support enables us to issue CARF Working Papers.

Detecting Real Estate Bubbles: A New Approach Based on the Cross-Sectional Dispersion of Property Prices

Takaaki Ohnishi* Takayuki Mizuno† Chihiro Shimizu‡ Tsutomu Watanabe§

First draft: December 31, 2012
This version: April 8, 2013

Abstract

We investigate the cross-sectional distribution of house prices in the Greater Tokyo Area for the period 1986 to 2009. We find that size-adjusted house prices follow a lognormal distribution except for the period of the housing bubble and its collapse in Tokyo, for which the price distribution has a substantially heavier right tail than that of a lognormal distribution. We also find that, during the bubble era, sharp price movements were concentrated in particular areas, and this spatial heterogeneity is the source of the fat upper tail. These findings suggest that, during a bubble period, prices go up prominently for particular properties, but not so much for other properties, and as a result, price inequality across properties increases. In other words, the defining property of real estate bubbles is not the rapid price hike itself but an increase in price dispersion. We argue that the shape of cross sectional house price distributions may contain information useful for the detection of housing bubbles.

JEL Classification Numbers: R10; C16
Keywords: house price indexes; lognormal distributions; power-law distributions; fat tails; hedonic regression; housing bubbles; market segmentation

*Canon Institute for Global Studies and University of Tokyo. E-mail: ohnishi.takaaki@canon-igs.org

†National Institute of Informatics. E-mail: mizuno@nii.ac.jp

‡Reitaku University. E-mail: cshimizu@reitaku-u.ac.jp

§Correspondence: Tsutomu Watanabe, Graduate School of Economics, University of Tokyo. E-mail: watanabe@e.u-tokyo.ac.jp. We would like to thank Kosuke Aoki, Donald Haurin, Christian Hilber, Tomoyuki Nakajima, Kiyohiko G. Nishimura, Misako Takayasu, and Hiroshi Yoshikawa for helpful comments and suggestions on an earlier version of this paper. This study forms part of the project on "Understanding Persistent Deflation in Japan" funded by a JSPS Grant-in-Aid for Scientific Research (24223003) and has further been supported by the program titled "Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures" as well as the program titled "High Performance Computing Infrastructure." Ohnishi gratefully acknowledges a grant from the Ministry of Education, Culture, Sports, Science, and Technology of Japan (Grant-in-Aid for Young Scientists (A) No. 23686019).

1 Introduction

Property market developments are of increasing importance to practitioners and policymakers. The financial crises of the past two decades have illustrated just how critical the health of this sector can be for achieving financial stability. For example, the recent financial crisis in the United States in its early stages reared its head in the form of the subprime loan problem. Similarly, the financial crises in Japan and Scandinavia in the 1990s were all triggered by the collapse of bubbles in the real estate market. More recently, the rapid rise in real estate prices - often supported by a strong expansion in bank lending - in a number of emerging market economies has become a concern for policymakers. Given these experiences, it is critically important to analyze the relationship between property markets, finance, and financial crisis.

One of the most urgent tasks in this respect is the development of methods to detect bubbles in property markets, which is key to making reliable forecasts about future financial crises and/or to mitigating them. Consider the case of residential properties. In real estate economics, the fundamental value of a house is determined as the present discounted value of current and future income flows resulting from renting the house to someone (see, for example, Himmelberg et al. (2005)). In a normal situation, due to price arbitrage, the price of a house remains close to its fundamental value. However, in some cases, arbitrage forces are not present, and therefore prices deviate substantially from the fundamental value. This is what is called a housing bubble. To detect housing bubbles defined in this way, we must have a reliable estimate of the fundamental value. This requires knowing about market participants' expectations on rental prices in the coming years. However, such expectations are not observable, and therefore the fundamental value is quite difficult to estimate. Thus, it is next to impossible for researchers or policymakers to tell, when they observe a price hike, whether it comes from a rise in fundamental values or something else.

In this paper, we propose an alternative approach to detecting real estate bubbles. We propose making use of information on the cross-sectional dispersion of real estate prices. It is often believed that *all* prices rise equally during a bubble period. However, this is not the case. What happens instead is that prices go up prominently for particular properties, but not for other properties, and as a result, price inequality across properties increases during a bubble period. In other words, the defining characteristic of real estate bubbles is not the rapid price hike itself but an increase in price dispersion.

Given these considerations, the present paper addresses the following empirical questions. First, we examine whether the price distribution is close to a normal distribution, as is often assumed in empirical studies on house price indexes, or whether it has fatter tails than a Gaussian distribution. Second, we are interested in how the shape of the price distribution is affected by house attributes, including the size and location of a house. Third, we would like to know how the shape of the distribution changes over time. In order to examine these three questions, we focus on the housing bubble Japan experienced in the late 1980s and its burst in the early 1990s.

Recent studies on the cross-sectional distribution of house prices include Gyourko et al. (2006), McMillen (2008), Van Nieuwerburgh and Weill (2010), and Maattanen and Tervio (2010). The main interest of Gyourko et al. (2006), Van Nieuwerburgh and Weill (2010), and Maattanen and Tervio (2010) is the relationship between the house price distribution and the income distribution. For example, Maattanen and Tervio (2010) ask whether the recent increases in income inequality in the United States have had any impact on the distribution of house prices. On the other hand, McMillen (2008) focuses on the change in the house price distribution over time and asks whether the change in the price distribution comes from a change in the distribution of house characteristics such as size, location, age, and so on, or from a change in the implicit prices associated with those characteristics. The focus of our paper is closely related to the issues discussed in these papers, but differs from them in some important respects. First, this paper is the first attempt to specify the shape of the house price distribution, paying particular attention to the tail part of the distribution. Second, this paper examines the effect of a housing bubble on the cross-sectional price distribution. While steep increases in the *mean* of house prices in various countries in recent decades have received a lot of attention in the literature, the change in the shape of the cross-sectional price distribution has received much less attention. In this paper, we seek to fill this gap.

Our main findings are as follows. First, the cross-sectional distribution of house prices has a fat upper tail and the tail part is close to a power-law distribution. This is confirmed by the goodness-of-fit test recently proposed by Malevergne et al. (2011). On the other hand, the cross-sectional distribution of house sizes, as measured by the floor space, has an upper tail that is less fat than that of the price distribution and is close to an exponential distribution. These two findings suggest a particular functional form of hedonic regression to identify the size effect. We construct size-adjusted prices by subtracting the house size (multiplied by a positive coefficient) from the log price and find that the size-adjusted price follows a lognormal distribution for most of the observation period. An important exception is the period of the housing bubble and its collapse in 1987-1995, during which the price distribution in each year has a power-law tail even after controlling for the size effect.

Second, we divide the area covered by our sample (Greater Tokyo) into small pixels and find that size-adjusted prices almost follow a lognormal distribution *within* each of these pixels even during the bubble period, but the mean and variance of each distribution are highly dispersed *across* different pixels. This finding implies that the sharp price hike during the bubble period was concentrated in particular areas, and this spatial heterogeneity is the source of the fat upper tail observed for the bubble period.[1] We interpret this as evidence for market segmentation during a bubble period.

The rest of the paper is organized as follows. Section 2 provides a brief overview of the Japanese housing bubble in the late 1980s. Section 3 then explains the dataset and the empirical strategy we employ. Next, Sections 4 and 5 present our size- and location-adjustments to house prices. Finally, Section 6 concludes the paper.

2 Overview of the Japanese Bubble in the Late 1980s

In this section, we provide a brief overview of what happened during the Japanese real estate bubble in the late 1980s and its collapse in the 1990s, as well as how the government and the central bank responded.

2.1 The real estate bubble in the late 1980s and its collapse in the 1990s

Figure 1 shows changes in the mean of the cross-sectional house price distribution in the upper panel, the standard deviation in the middle panel, and the transaction volume in the lower panel. The data is compiled from individual listings in a real estate advertisement magazine, which is published on a weekly basis. (More details on the dataset used in the paper are provided below). We see that the mean price exhibits a sharp increase between the beginning of 1987 and the beginning of 1988. Previous studies refer to this as the first phase of the housing bubble in Tokyo. After a short break in 1988, prices started to rise again in 1989 and continued to do so until the fall of 1990. This is the second phase of the housing bubble. Soon after the fall of 1990, prices started to turn down, followed by a slow but persistent decline for more than a decade until prices bottomed out in 2002, when the mean price reached the level before the bubble started in 1987. Prices finally began to rise again in 2003 and continued to rise until registering a sharp decline in 2008 due to the recent global financial crisis.

Turning to the standard deviation shown in the middle panel, this exhibits a sharp rise during the first phase of the bubble and stayed high during the second phase.[2] Finally, the bottom panel, which shows the transaction volume, indicates that the number of transactions

[1] Cochrane (2002) argues that an important feature of the tech stock bubble in the late 1990s is that it was concentrated in stocks related to internet business. Cochrane (2002: 17) states that "if there was a 'bubble,' or some behavioral overenthusiasm for stocks, it was concentrated on Nasdaq stocks, and Nasdaq tech and internet stocks in particular."

[2] We also see a secular increase in price dispersion since 1993. We are not quite sure why this is the case, but recent studies, including Van Nieuwerburgh and Weill (2010) and Gyourko et al. (2006) find some evidence that the recent rise in house price dispersion across regions in the United States is related to the change in income distribution across regions. For example, Van Nieuwerburgh and Weill (2010) find that the cross-sectional coefficient of variation of house prices across 330 metropolitan statistical areas in the United States increased from 0.15 in 1975 to 0.53 in 2007. Through a counterfactual simulation, they show that this increase in the dispersion of house prices is accounted for mostly by the increase in income inequality.

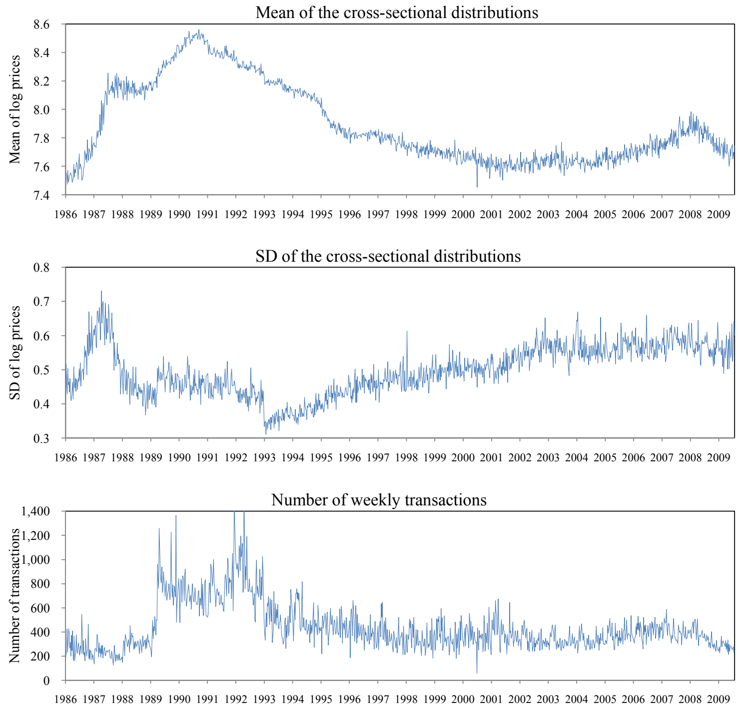


Figure 1: Weekly fluctuations in prices and transaction volume

exhibits a sharp increase at the beginning of 1989, exactly when the mean price started to rise, although the transaction volume remained practically unchanged during the first phase of the bubble. Somewhat interestingly, the transaction volume remained at a high level even in 1991 and 1992, when the mean price had already started to decline.

Okina et al. (2001) identify three characteristics of the Japanese economy during the bubble period. The first characteristic is a rapid and substantial rise in asset prices, including stock and real estate prices. Stock prices exhibited a rapid rise during the initial stage of the bubble; the Nikkei 225 began to accelerate in 1986 and the index hit a peak of 38,915 yen at the end of 1989, when it was 3.1 times higher than at the time of the Plaza Agreement in September 1985. Land prices followed stock prices with a lag of about one year, increasing sharply in 1987 and peaking in late 1990, as mentioned above, with increases spreading from Tokyo to the other major cities such as Osaka and Nagoya, and then to other cities of smaller size. The second characteristic of this period is very high economic growth. The business cycle hit a trough in November 1986 and the economy then expanded for 51 months until 1991. Real GDP grew at an average annual rate of 5.5 percent during this period, driven by business fixed investment, housing investment, and expenditure on consumer durables. The third characteristic is a rapid increase in money supply and credit. The annual growth rate of money supply reached more than 10 percent in April-June 1987 as a result of monetary easing by the Bank of Japan as well as financial deregulation. Also, bank borrowing and financing from capital markets substantially increased in 1988 and 1989.

These three characteristics of the Japanese economy during the bubble period were closely related with each other. In particular, land prices, bank borrowing, and business investment were tightly linked through the credit cycle mechanism highlighted by Kiyotaki and Moore (1997). Shimizu and Watanabe (2013), using information from land registry data, examine empirically whether such a link indeed exists by looking at the market value of land owned by firms and the amount of bank lending to those firms. They show that rapid price rises in the late 1980s raised the value of land as collateral, making it possible for banks to extend larger loans.

2.2 Causes of the real estate bubble in the 1980s

Previous studies on the real estate bubble in the 1980s identify the following as factors behind the emergence and expansion of the bubble: (1) aggressive behavior of financial institutions; (2) financial deregulation; (3) inadequate risk management by financial institutions; (4) the introduction of capital adequacy requirements for banks; (5) protracted monetary easing by the Bank of Japan; (6) taxation and regulation; (7) overconfidence and euphoria; (8) demographic changes; (9) over-concentration of economic functions on Tokyo, and Tokyo becoming an international financial center.[3] Obviously these factors are not mutually independent. On the demand side, demographic changes, over-concentration, and euphoria are three important factors, creating excess demand in the real estate market. On the other hand, taxation and regulation were critically important on the supply side, because land taxation and land-related regulations made land supply price-inelastic, thereby making it impossible to eliminate excess demand without raising real estate prices. On the monetary side, expansionary monetary policy and banks' loose lending behavior are two important factors that made it possible for firms and individuals to have easy access to liquidity.[4] In the rest of this section, we will focus on the issues associated with demographic changes and taxation.

Demographic changes and housing demand Mankiw and Weil (1989) argue that demographic changes, such as baby booms and busts, have an impact on housing demand and therefore on housing prices. This directly follows from the Ando-Modigliani life cycle hypothesis; namely, people buy houses during their working career and sell them in old age. A similar idea has recently been proposed by Nishimura (2011), who argues that the real estate bubble in

[3] See Okina et al. (2001), Okina and Shiratsuka (2002), and Shiratsuka (2005) for more details on each factor.

[4] See Ueda (1998), Hoshi and Kashyap (2000), and Baba et al. (2005) on banks' lending behavior during the bubble and its burst.

the 1980s and its collapse in the 1990s were caused, at least partially, by demographic changes. His argument is based on a simple comparison of the inverse dependency ratio (i.e., the ratio of the working-age population between 15 and 64 years of age to the rest of population, who are dependent on the working-age population) and real land prices. The inverse dependency ratio hit a trough in 1980 and exhibited a gradual increase until reaching a peak in 1990, when the children of the post-World War II baby boomers reached working age. The ratio then started to decline again from 1990 onward and has continued to do so in the more than two decades since then. These ups and downs in the inverse dependency ratio more or less coincide with those in real land prices. Nishimura (2011) interprets this as evidence that demographic changes affect land prices through changes in land demand.

Employing international panel data on 22 countries for 1970-2009, Takáts (2012) finds that an increase in the population by 1 percent is associated with an increase in real house prices by 1 percent, and that an increase in the old age dependency ratio by one percent is associated with a decrease in real house prices by 2/3 of a percent. These results imply that, on average, demographic factors raised house prices in advanced economies by around 30 basis point per annum in the past 40 years and will, on average, decrease house prices in advanced economies by around 80 basis points per annum over the next 40 year.

Following Mankiw and Weil (1989), Shimizu and Watanabe (2010) estimate the weighted average of housing demand of each generation, using the population share of each generation as weights, for each prefecture from 1975 to 2008. They find that the change in housing demand and the change in house price are positively correlated, implying that house prices tend to go up in a particular prefecture when housing demand increases there due to demographic reasons. Focusing on the bubble period (i.e., 1985-1990), they find that the cross-sectional correlation is particularly high, although the very sharp price increases in prefectures with large populations, such as Tokyo and Osaka, cannot be accounted for solely by demographic factor.

Land taxation and its reform in the early 1990s It has been pointed out by many practitioners and researchers that the taxation system in Japan gave land owners an incentive to hold land for speculative purposes, thereby making land supply inelastic to price changes.[5] In this respect, the following two characteristics of the Japanese tax system play a key role. First, tax imposed on the holding of land has been extremely low in Japan when compared with other countries. This is particularly true for agricultural land, and the rate of tax imposed on the holding of agricultural land, even for agricultural land in urban outskirts, has been extremely low. This low cost of holding land, especially agricultural land, results in the inefficient use

[5]In Japan, land is taxed at three stages: inheritance tax, registration and license tax, and/or real estate acquisition tax are applied when land is acquired; and fixed property tax, urban planning tax, and/or special land-holding tax are applied while land is held; corporation tax, income tax, and/or inhabitant tax are applied to capital gains when land is transferred.

Table 1: Chronology of policy actions related to the real estate bubble in the 1980s and its collapse in the 1990s

1986	Tokyo Metropolitan Government introduces regulation on land transactions
1987	National Land Use Planning Act is amended
	Closer monitoring of land transactions starts
1990	Local Property Tax Act is amended
	Tax enforcement for agricultural land in urban areas is strengthened
1990	National Land Use Planning Act is amended
	Surveillance of resale of properties within two years is introduced
1990	Ministry of Finance issues a directive and guidance requesting financial institutions
	(1) to contain the increase in lending to property-related firms to within the increase
	in total lending, and (2) to report lending to the real estate industry,
	construction industry, and non-banks
1991	Land Value Tax Act is enacted
1992	Land Value Tax is newly introduced
1993	Cooperative Credit Purchasing Co. is established
1995	Tokyo Kyodo Bank (Japanese RCC) is established
1998	Act on Emergency Measures for Early Strengthening of Financial Functions is enacted

of land and in land being left unused, and makes holding land more profitable than holding financial and other non-financial assets, promoting speculative land transactions. Second, as for taxes on capital gains, the tax rates were not sufficiently high to prevent speculative transactions; that is, for individuals, an income tax of 20 percent was imposed on capital gains up to 40 million yen and a consolidated tax of 50 percent above this threshold: for companies, apart from ordinary corporation tax, an additional tax of 20 percent was imposed on capital gains on land held for ten years or less.

In the early 1990s, a series of tax reforms was implemented as a part of policy responses to suppress speculative land transactions, following the publication of the report on "Basic Issues Regarding Revisions in the Land Tax System" in May 1990 by the Land Tax System Subcommittee.[6] The chronology of tax reforms is presented in Table 1. The tax reforms implemented in the early 1990s consist of three parts. First, in January 1992, property tax rates were raised and a land value tax was newly introduced to make land less attractive as an investment, thereby suppressing speculative land transactions. Second, capital gains taxation was strengthened. In particular, a punitively high tax rate was imposed on capital gains from transactions in land which has been held only for a short period. For example, in the 1991 reform, a separation tax of 30 percent was additionally introduced for capital gains on transactions in land held by companies for two years or less. Finally, the tax rate associated with the holding of land for agricultural purposes in urban areas was raised to the same level as that for residential purposes in order to eliminate the special treatment of land for agricultural

[6]See Morinobu (2006) for more details on the tax reforms during this period.

purposes. Practices and methods regarding land appraisals for inheritance tax purposes were also amended.

3 Data and Empirical Strategy

3.1 Data

We use a unique dataset that we compiled from individual listings in a widely circulated real estate advertisement magazine, which is published on a weekly basis by Recruit Co., Ltd., one of the largest vendors of residential lettings information in Japan. The dataset covers the Greater Tokyo Area for the period 1986 to 2009, thus including the bubble period in the late 1980s and its collapse in the first half of the 1990s. It contains 724,416 listings for condominiums and 1,602,918 listings for single family houses.[7] In this paper we will use data only for condominiums. The Greater Tokyo Area covered in the dataset includes the 23 special wards of Tokyo, other areas making up the Tokyo Metropolis, as well as adjacent cities and suburbs. According to Shimizu et al. (2004), the dataset covers more than 95 percent of all transactions in the 23 special wards (i.e., central Tokyo), while the coverage for the other areas is somewhat more limited. This dataset has been used in a series of papers, including Shimizu et al. (2010), which compares hedonic and repeat-sales measures in terms of their performance.

3.2 Empirical strategy

A widely used approach to deal with product heterogeneity in terms of quality is hedonic analysis, which has been applied in a number of studies to analyze real estate prices. The core idea of hedonic analysis is that the value of a product is the sum of the values of individual product characteristics. For example, Shimizu et al. (2010) start their analysis by assuming that the value of a house is the sum of the values of attributes such as its floor space, its age, the commuting time to the nearest station, and so on, and run hedonic regressions using these attributes as independent variables.

This idea has important implications regarding the shape of the cross-sectional distribution of house prices. To show this, let us start by assuming that the price of house i at a particular point in time, which is denoted by P_i,[8] is the sum of K components:

$$P_i = F(X_{i1}, X_{i2}, \ldots, X_{ik}, \ldots, X_{iK}) \tag{1}$$

[7]The dataset contains full information about the evolution of the posted price for a housing unit from the week when it was first listed until the final week when it was removed because of a successful transaction. In this paper, we only use the price in the final week, since this can be safely regarded as sufficiently close to the contract price. The number of listings shown in the text does not include those prices listed before the final week.

[8]Note that the subscript for time is dropped here to simplify the exposition.

where P_i and X_{ik} are both random variables and $X_{i1}, \ldots, X_{iK}$ are assumed to be independent from each other. Furthermore, we assume a multiplicative functional form such that

$$P_i = \prod_{k=1}^{K} X_{ik}. \tag{2}$$

Taking the logarithm of both sides of this equation leads to

$$\ln P_i = \sum_{k=1}^{K} x_{ik}, \tag{3}$$

where $x_{ik} \equiv \ln X_{ik}$. This equation appears frequently in hedonic analyses of house prices. It simply states that the price of a house is equal to the sum of K random variables.

Given this setting, the central limit theorem tells us that the sum of these random variables converges to a normal distribution if the number of attributes, K, goes to infinity. Let us denote the variance of x_{ik} by s_k^2 and define the average variance $\bar{s}_K^2$ as

$$\bar{s}_K^2 \equiv \frac{1}{K}\left(s_1^2 + s_2^2 + \cdots + s_K^2\right).$$

Then, according to the Lindberg-Feller central limit theorem, the sum of random variables $\sum_{k=1}^{K} x_{ik}$ converges to a normal distribution as K goes to infinity, if the average variance $\bar{s}_K^2$ converges to a finite constant (namely, $\lim_{K\to\infty} \bar{s}_K^2 = \bar{s}^2$) and the following condition is satisfied:

$$\lim_{K\to\infty} \frac{\max_{k\leq K}\{s_k\}}{K\bar{s}_K} = 0. \tag{4}$$

In other words, the theorem states that the sum of random variables, regardless of their form, will tend to be normally distributed. A notable feature of this result is that it does *not* require that the variables in the sum come from the same underlying distribution. Instead, the theorem requires only that no single term dominates the average variance, as stated in (4). Put differently, condition (4) states that none of the random variables is dominantly large relative to their sum.[9] A famous textbook example of the central limit theorem is the distribution of persons' height. The height distribution of, say, mature men of a certain age can be considered normal, because height can be seen as the sum of many small and independent effects. Similarly, the log price of houses will be normally distributed if house prices are determined as the sum of many small and independent effects.

The above argument suggests that the lognormal distribution can be seen as a benchmark for the cross-sectional distribution of house prices. However, some previous studies on house price distributions find that the actual distributions have fatter tails than a lognormal distribution.

[9] For more on this theorem, see Feller (1968). Greene (2003) provides a compact description of various versions of the central limit theorem including this one.

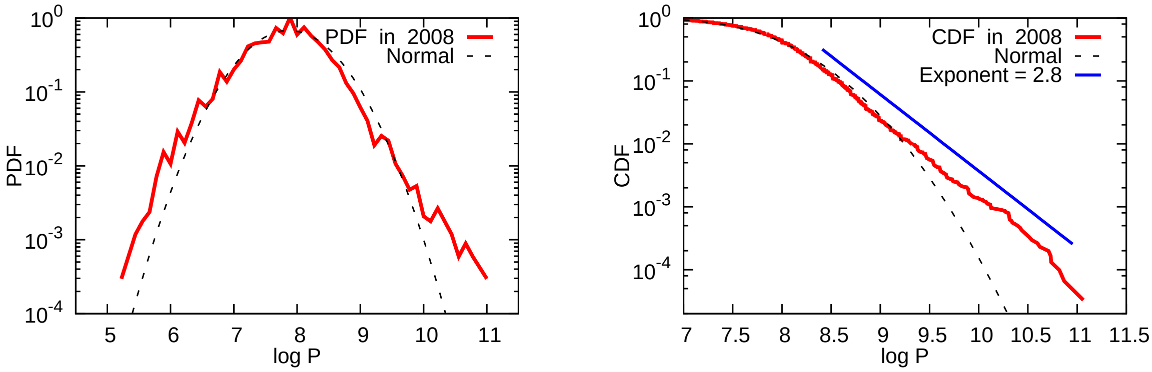


Figure 2: House price distribution in 2008

For example, McMillen (2008), using data on single family houses in Chicago for 1995, shows that the kernel density estimates for the log price are asymmetric, with a much fatter lower tail. Against this background, we examine the extent to which the house price distribution deviates from a lognormal distribution using our observations for 2008. The results are presented in Figure 2, where the left panel shows the probability density function (PDF), with the horizontal axis representing the yen price in logarithm and the vertical axis representing the corresponding density, also in logarithm. The empirical distribution is shown by the red line, while the lognormal distribution with the same mean and standard deviation is shown by the black dotted line. The figure indicates that the tails of the empirical distribution are fatter than those of the lognormal distribution. In particular, the upper tail of the empirical distribution is much fatter than that of the lognormal distribution. To examine the differences in the upper tail more closely, we accumulate the densities from the right (upper) tail to produce the cumulative distribution function (CDF), which is shown in the right panel. In this panel, the value on the vertical axis corresponding to the value of 9.2 on the horizontal axis, for example, is 0.01, meaning that the fraction of houses whose prices are equal to or higher than that price level is 1 percent. We now see more clearly that the upper tail of the empirical distribution is fatter than that of the lognormal distribution. For example, the fraction of housing units whose price deviates from the mean by more than 3σ is about 1.47 percent, while the corresponding number for the lognormal distribution is only 0.26 percent.

What causes the empirical distribution to deviate from the benchmark (i.e., the lognormal distribution)? This is the main topic we address in this paper. Our hypothesis is that some of the factors that determine house prices are dominantly volatile, so that condition (4) is violated. Denoting these dominant factors by vector Z_i, the house price distribution, $\Pr(P_i = p)$, can be decomposed as follows:

$$\Pr(P_i = p) = \sum_z \Pr(P_i = p \mid Z_i = z) \Pr(Z_i = z). \tag{5}$$

Note that the house price distribution conditional on Z_i, namely $\Pr(P_i = p \mid Z_i = z)$, should be a lognormal distribution, since the dominant factors are now fully controlled for. This means that the right-hand side of equation (5) is a weighted sum of lognormals, with the weights being given by $\Pr(Z_i = z)$. We know that the sum of lognormals with different means and variances is no longer a lognormal (see, for example, Feller (1968)), and the hypothesis we examine is that this is why the house price distribution deviates from the benchmark. Given this hypothesis, we proceed as follows in the remainder of the paper: we first specify the dominant factors and then eliminate them, thereby constructing prices that are adjusted for these factors; finally, we examine whether these adjusted prices follow a lognormal distribution.

Diewert et al. (2010) argue that there are three important price determining characteristics: the land area of the property; the livable floor space area of the structure; and the location of the property. Similarly, previous studies on house prices in Japan, including Shimizu et al. (2010), find that the floor space of a housing unit (especially in the case of condominiums) and its location play dominantly important roles in determining its price. This empirical evidence suggests that the size and the location of a property are key candidates for the Z variables. We will identify and eliminate the size effect in the next section, and the location effect in Section 5.

4 Size-adjustment to House Prices

4.1 Distribution of unadjusted house prices

Figure 3 presents the PDFs and CDFs of the cross-sectional price distribution for each year from 1986 to 2009. To make the price distributions in different years comparable, we normalize the log prices in year t by subtracting the mean in year t (i.e., the mean of log prices in year t) and dividing by the standard deviation in year t (i.e., the standard deviation of log prices in year t). The lognormal lines in the figure represent the CDF of a standard lognormal distribution. Note that the CDFs are constructed in the same way as in Figure 2, that is, the value on the vertical axis corresponding to a price level is the sum of the densities *above* that price level.

The first thing we see from the figure is that, as in Figure 2, the PDFs and the CDFs show fatter upper tails than a lognormal distribution. More importantly, we see that the deviation from a lognormal distribution tends to be larger for the late 1980s and the first half of the 1990s. Specifically, the PDFs in these years are substantially skewed to the right, indicating that during the bubble period house prices did not rise by the same percentage for every housing unit; instead, price increases were concentrated in particular housing units, so that relative prices across houses changed significantly.

The CDFs in this figure provide more detailed information regarding the shape of the price distributions. We see that the CDF for each year forms an almost straight line in this log-

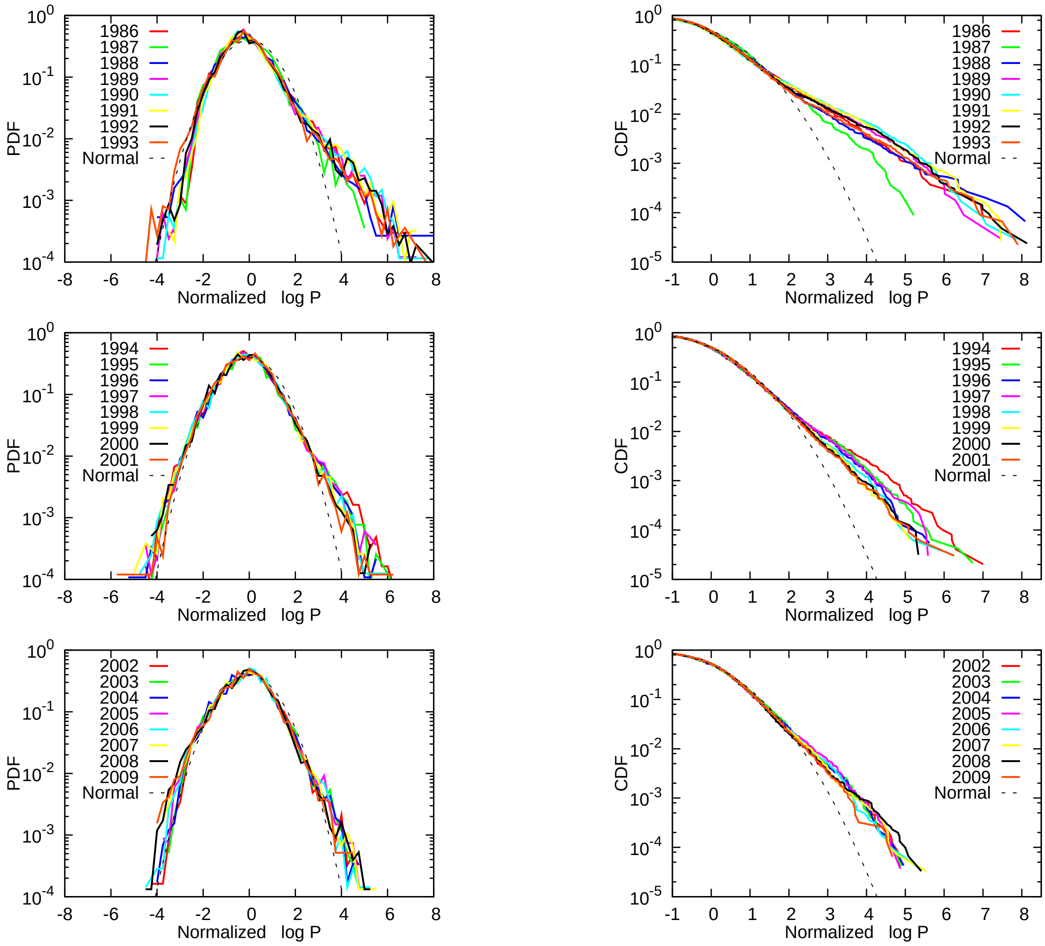


Figure 3: House price distributions by year

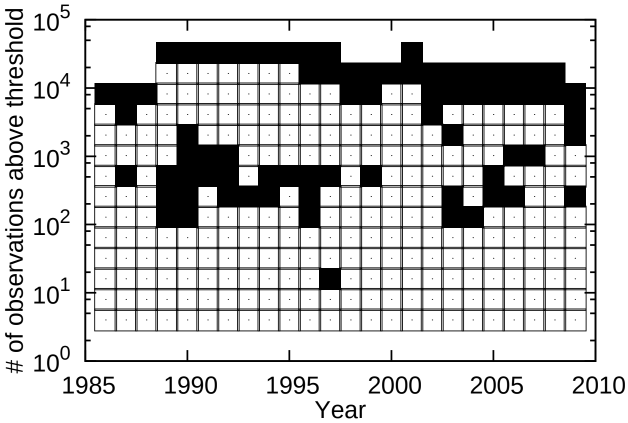
log graph, implying that the house price distribution is well approximated by a power law distribution (or a Pareto distribution) at least in the tail part, the PDF and CDF of which are given by

$$\Pr(P_{it} = p) = \frac{\zeta_t m_t^{\zeta_t}}{p^{\zeta_t+1}}; \quad \Pr(P_{it} \geq p) = \left(\frac{m_t}{p}\right)^{\zeta_t}; \quad p > m_t > 0 \tag{6}$$

where P_{it} denotes the price of house i in period t, and ζ_t and m_t are time-variant positive parameters.[10] The shape of a power law distribution is mainly determined by the parameter ζ_t, which is referred to as the exponent of the power law distribution. Smaller values for ζ_t imply fatter tails. Note that the CDF given in (6) implies that

$$\ln \Pr(P_{it} \geq p) = -\zeta_t \ln p + \zeta_t \ln m_t.$$

[10] See Gabaix (2008) for an extensive survey of empirical and theoretical studies on power laws in various economic contexts such as income and wealth, the size of cities and firms, and stock market returns.



■: The null (power law) is rejected at the 1 percent significance level.
⊡: The null is not rejected at the 1 percent significance level.

Figure 4: Power law distribution versus lognormal distribution

In other words, the log of the cumulative probability should be linearly related to the log price, and the slope of a linear line between the two variables should be equal to $-\zeta_t$. The CDFs in Figure 3 suggest the presence of such a linear relationship between the log price and the log of the cumulative probability. We see from the CDF in Figure 2 that ζ_{2008} is about 2.8. Similarly, we find from the corresponding figures for the other years (which are not shown due to space limitations) ζ also all take values of around 3.[11]

As a goodness-of-fit test, we employ the test proposed by Malevergne et al. (2011). Specifically, we test the null hypothesis that, beyond some threshold u, the upper tail of the house price distribution is characterized by a power law distribution,

$$\Pr(P = p;\ \alpha) = \alpha \cdot \frac{u^{\alpha}}{p^{\alpha+1}} \cdot 1_{p \geq u}$$

, against the alternative that the upper tail follows a lognormal beyond the same threshold, i.e.,

$$\Pr(P = p;\ \alpha, \beta) = \left[\sqrt{\frac{\pi}{\beta}} \exp\left(\frac{\alpha^2}{4\beta}\right)\left(1 - \Phi\left(\frac{\alpha}{\sqrt{2\beta}}\right)\right)\right]^{-1} \frac{1}{p} \exp\left(-\alpha \ln \frac{p}{u} - \beta \ln^2 \frac{p}{u}\right) \cdot 1_{p \geq u},$$

[11]Note that we cannot obtain estimates for ζ_t from Figure 3. The CDFs in Figure 3 are for normalized prices, which are defined by $[P_{it} \exp(-\mu_t)]^{1/\sigma_t}$, where μ_t and σ_t are the mean and the standard deviation in year t. Therefore, the slope of each CDF in Figure 3 is given by $\sigma_t \zeta_t$ (rather than ζ_t), if the original price follows the power law distribution given by (6). Taleb (2007) provides many examples of power law distributions. For example, the net worth of Americans follows a power law distribution with an exponent of 1.1; the frequency of the use of words follows such a distribution with an exponent of 1.2; the population of U.S. cities has an exponent of 1.3; the number of hits on websites has an exponent of 1.4; the magnitude of earthquakes has an exponent of 2.8; and price movements in financial markets have an exponent of 3 (or lower). The exponents for the house price distributions estimated here are greater than most of these figures, implying that the tail parts of the house price distributions are less fat than those in the other examples of power law distributions.

where $\Phi(\cdot)$ represents the CDF of a standard normal distribution. Note that this is equivalent to testing the null that the upper tail of the *log* price follows an exponential distribution against the alternative that it follows a normal distribution. For this transformed test, Del Castillo and Puig (1999) have shown that the clipped empirical coefficient of variation $\hat{c} \equiv \min(1, c)$ provides the uniformly most powerful unbiased test, where c is the empirical coefficient of variation.

The result of our goodness-of-fit test is presented in Figure 4, where the horizontal axis represents the year and the vertical axis represents the number of observations above the threshold u. For example, 10^3 on the vertical axis means that the threshold u is set such that the number of observations above u is 10^3. A black square indicates that the null is rejected at the 1 percent significance level for a particular year-threshold combination, while a white square indicates that the null is not rejected at the same significance level. The figure shows that a power law distribution provides a good approximation for the 500 most expensive houses, while a lognormal distribution provides a better approximation for the set of less expensive houses.

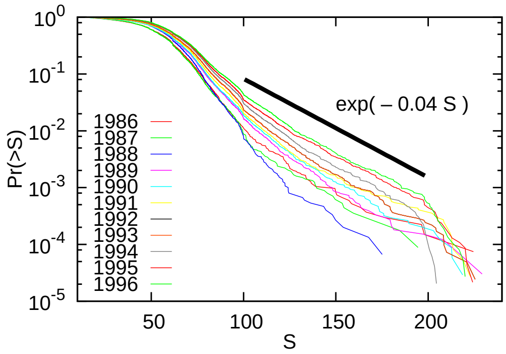
4.2 Distribution of house sizes

Previous studies on wealth (or income) distributions across households have typically found that those distributions are characterized by fat upper tails, and that they follow a power law distribution (see Pareto (1896)). Given that houses form an important part of households' wealth, it may be not that surprising that we detect a similar pattern in the house price distribution. However, the result that house prices follow a power law distribution is not consistent with the argument based on the central limit theorem. Why do house prices follow a power law distribution rather than a lognormal distribution? As a first step to address this question, we decompose the house price distribution as follows:

$$\Pr(P_{it} = p) = \sum_{s} \Pr(P_{it} = p \mid S_i = s) \Pr(S_i = s), \tag{7}$$

where S_i represents the size of housing unit i, which is measured by the floor space of that unit. The term $\Pr(S_i = s)$ represents the distribution of house sizes, while the term $\sum \Pr(P_{it} = p \mid S_i = s)$ represents the distribution of house prices conditional on house size. An important thing to note is that even if each of these conditional distributions is lognormal, the weighted sum of lognormals with different mean and variance is not a lognormal distribution. This is a potential source of the power law tails that we observed in our house price data.

We start by examining the term $\Pr(S_i = s)$ in equation (7). Figure 5 presents the CDFs of house sizes for each year, with the floor space, measured in square meters, on the horizontal axis and the log of the CDF on the vertical axis. We see that the CDF for each year is close to a straight line in this semi-log graph, implying that the size distribution can be approximated



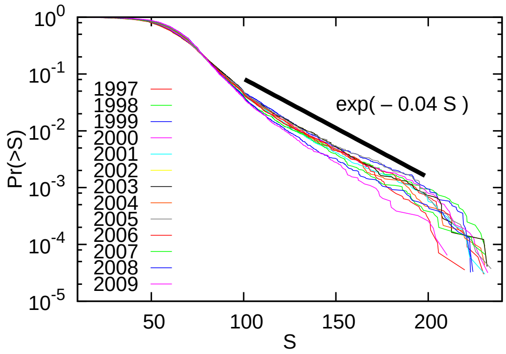


Figure 5: Cumulative house size distributions

by an exponential distribution whose PDF and CDF are given by

$$\Pr(S_i = s) = \lambda_t \exp\left(-\lambda_t s\right); \quad \Pr(S_i \geq s) = \exp\left(-\lambda_t s\right); \quad \lambda_t > 0. \tag{8}$$

Note that the CDF shown above implies that

$$\ln \Pr(S_i \geq s) = -\lambda_t s,$$

so that the log of the CDF depends linearly on house size. This is what we see in Figure 5. The slope of the CDF line, namely the value of λ, is almost identical for the different years and is somewhere around 0.04.

The fact that house sizes follow an exponential distribution implies that the tails of the size distribution are less fat than those of the price distribution. For example, for 2008, the fraction of housing units whose size deviates from the mean by more than 3σ is only 0.94 percent, while the corresponding number for the price distribution is 1.47 percent.[12]

4.3 Size-adjusted prices

We now turn to the relationship between the price of a house and its size, which is represented by the conditional probability $\Pr(P_{it} = p \mid S_i = s)$ in equation (7). We propose a hedonic model which is consistent with the fact that house prices and sizes follow, respectively, a power law distribution with an exponent of ζ_t and an exponential distribution with an exponent of λ_t. That is, the log prices are determined by

$$\ln P_{it} \sim \left(\frac{\lambda_t}{\zeta_t}\right) S_i + \epsilon_{it}, \tag{9}$$

[12]To see why the tails of the house size distribution are less fat than the tails of the price distribution, consider a simple example in which household A has 100 times as much wealth as household B, so that A spends 100 times as much money on a house as B. What does A's house look like? Does it have a bathroom that is 100 times larger than the one in B's house? Alternatively, does it have 100 bathrooms? Needless to say, neither is true, because even a person of A's wealth would have little use for such a gigantic bathroom (or so many bathrooms). Instead, it is more likely that the size of A's house (and therefore the size and/or number of its bathroom) is only, say, 10 times greater and, consequently, the unit area price of A's house is 10 times higher than B's.

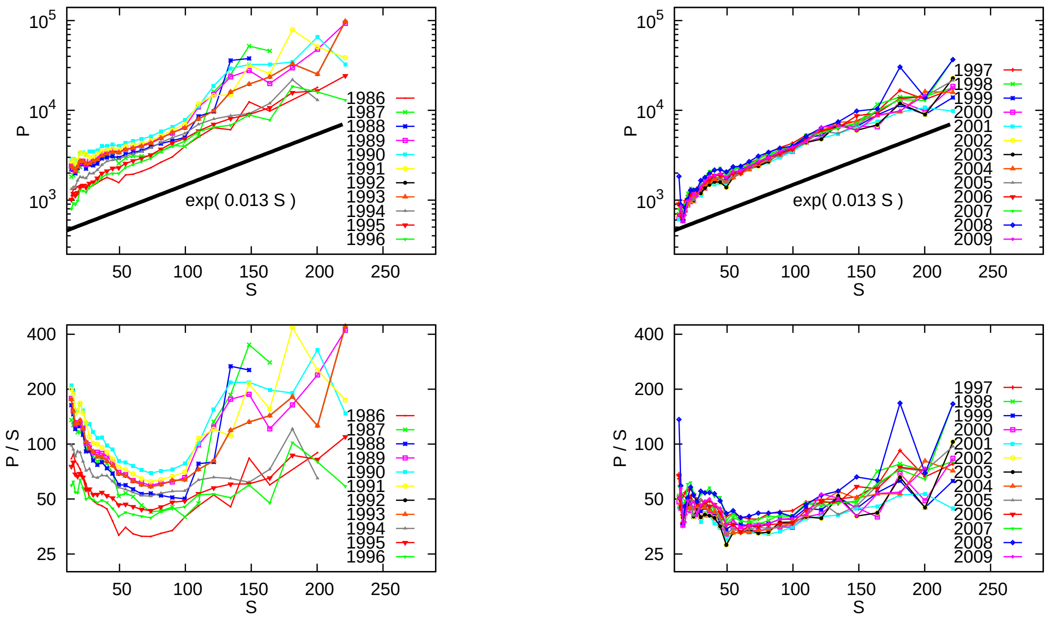


Figure 6: Relationship between house size and price

where ϵ_{it} is a normally distributed random variable, which, as we saw in Section 3.2, can be interpreted as the sum of many small and independent factors. To show equation (9), we first note that the PDF of the exponential distribution given in (8) implies that $(\lambda_t/\zeta_t)S_i$ follows an exponential distribution with an exponent of ζ_t if S_i itself is an exponential distribution with an exponent of λ_t. Next, we can show that the sum of the random variable that follows an exponential distribution and the random variable that follows a normal distribution is well approximated by the exponential distribution when the sum takes large values (because of the much fatter tails of an exponential distribution). Combining the two, the right-hand side of (9) is well approximated by an exponential distribution with an exponent of ζ_t when the sum of the two terms on the right-hand side takes large values. On the other hand, the fact that P_{it} follows a power law distribution with an exponent of ζ_t implies that $\ln P_{it}$ follows an exponential distribution with an exponent of ζ_t. In this way we can confirm that each side of equation (9) follows an identical distribution with an identical exponent.[13]

To empirically test the hedonic model given by (9), we first examine for a linear relationship between the log price of houses and their size. The upper panels of Figure 6 show the floor space on the horizontal axis and the median of the log price corresponding to that size on the vertical axis. These panels indicate that there exists a stable linear relationship between

[13]The price-size relationship described by equation (9) provides an answer to the question regarding the choice of functional form for hedonic price equations, which has been extensively discussed in previous studies such as Cropper et al. (1988), Diewert (2003), and Triplett (2004). The novelty of our approach is that we derive this functional form not from economic theory but from the statistical fact that house prices and sizes follow a power law and an exponential distribution, respectively.

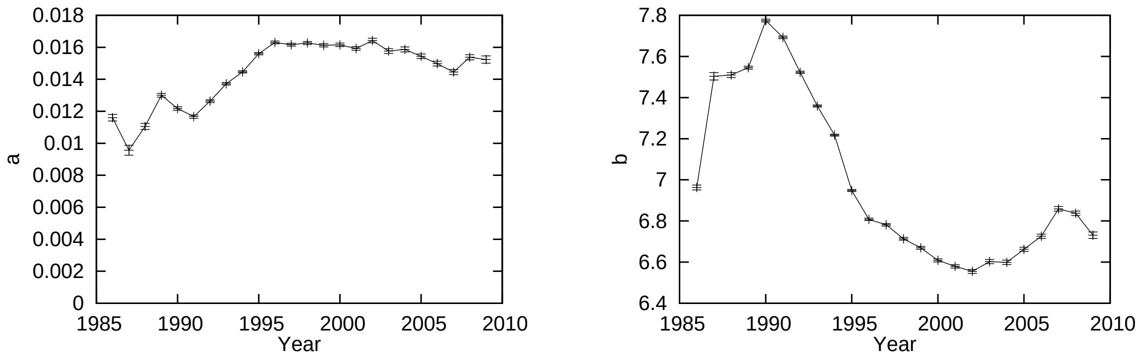


Figure 7: Price-size regressions

the two variables. Furthermore, equation (9) implies that the per unit area price, $P/S = [\exp(\lambda/\zeta)S + \text{positive constant}]/S$, decreases with S when S is small and increases with S when S is sufficiently large, so that there should exist a U-shaped relationship between the per unit area price and the house size. The lower panels of Figure 6, in which the vertical axis now measures P/S, confirms this prediction.

Second, we run an OLS regression of the form

$$\ln P_{it} = a_t S_i + b_t + \eta_{it} \tag{10}$$

to see whether the disturbance term η_{it} is indeed normally distributed as assumed in (9). The regression results are presented in Figures 7 and 8. Figure 7 shows the estimates of a and b for each year. The estimate of a is almost identical across years and is around 0.013, implying that an increase in the house size by a square meter leads to a 1.3 percent increase in the house price. More importantly, the estimate of a is very close to the value predicted by (9). That is, the value of ζ is around 3, as we saw in Section 4.1, and the value of λ is about 0.04, as we saw in Section 4.2, so that the coefficient on S_i, namely λ/ζ, should be something around 0.013 ($= 0.04/3$). This is quite close to the point estimate of a for each year.[14] Turning to the estimate of b, this exhibits substantial fluctuations: it increases by more than 20 percent per year from 1986 to 1990 and then declines by 10 percent per year from 1990 to 2002.

Figure 8 shows the CDFs of size-adjusted prices, which are defined by

$$\tilde{P}_{it} \equiv \left[P_{it} \exp(-\hat{a}_t S_i - \hat{b}_t)\right]^{1/\hat{\sigma}_t}, \tag{11}$$

where $\hat{a}_t$ and $\hat{b}_t$ are the estimates of a_t and b_t, and $\hat{\sigma}_t$ is the estimate for the standard deviation of η_{it}. Note that the hedonic model given by (9) implies that $\tilde{P}_{it}$ should be a lognormal

[14] Note that the per unit area price, $\exp(aS + b)/S$ takes its minimum value when S is equal to $1/a$. Given the estimate of a, this implies that the per unit area price takes its minimum value when $S = 1/0.013 \approx 75$, which is consistent with what we see in the lower two panels of Figure 6.

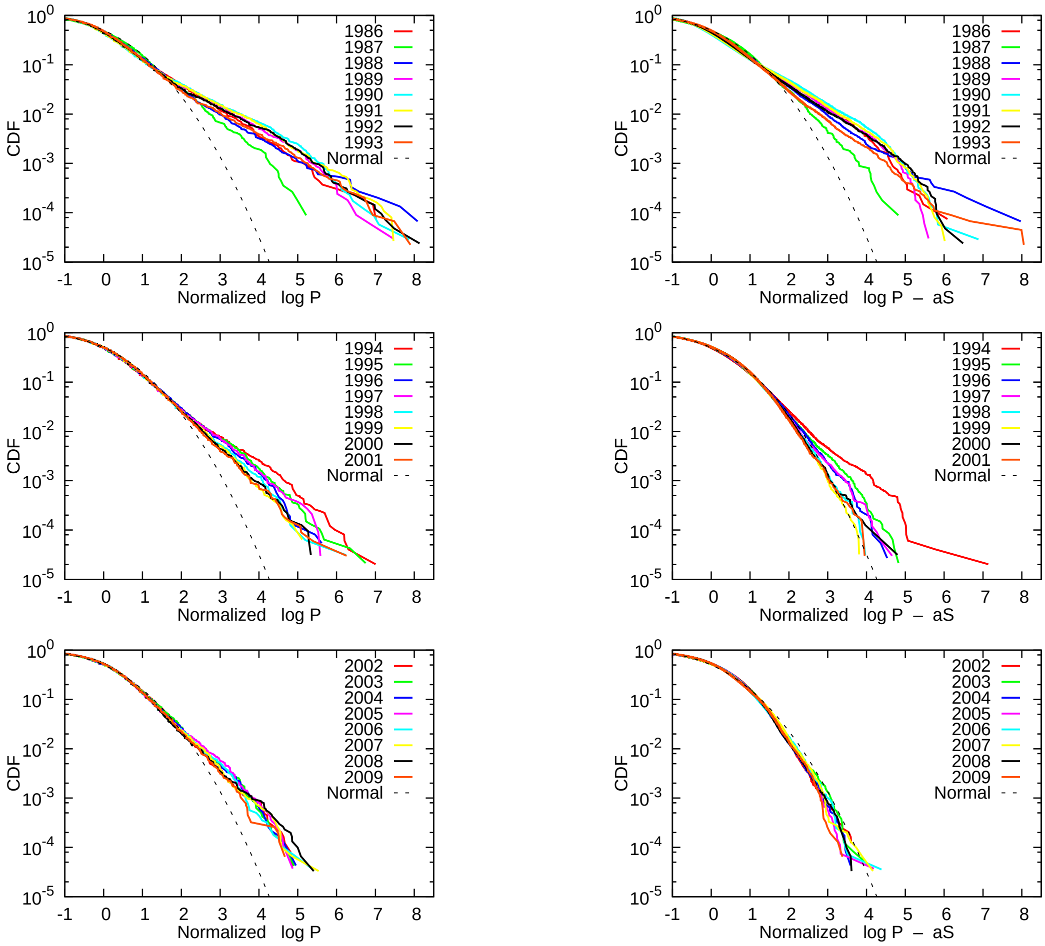


Figure 8: Cumulative distributions of size-adjusted house prices

distribution. The CDFs of the size adjusted prices are shown in the three panels on the right-hand side of Figure 8, while the price distributions *without* size adjustments from Figure 3 are replicated on the left-hand side. Comparing these two sets of CDFs, we see that the CDFs of the size-adjusted prices are much closer to the CDF of a lognormal distribution. More specifically, the CDFs for 2002 to 2009, which are shown in the lower right panel, are almost identical to the CDF of a lognormal distribution. The same applies to the CDFs for 1994 to 2001, which are shown in the middle right panel. However, the CDFs for 1986 to 1993, which are presented in the upper right panel, are still far from the CDF of a lognormal distribution, although they are slightly closer to it than the CDFs of the non-adjusted prices.

5 Location Adjustment to House Prices

The analysis in the previous section suggested that size-adjusted prices followed a lognormal distribution at least for quiet periods without large price fluctuations. This is consistent with

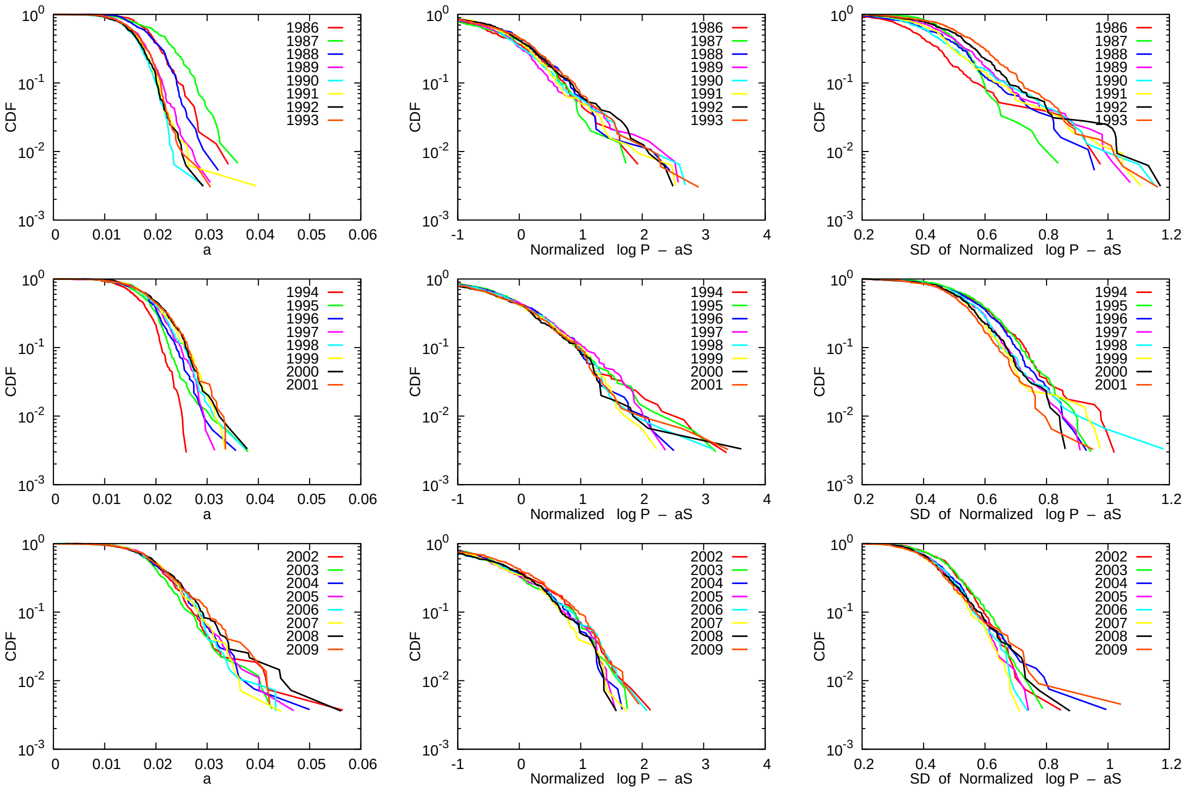


Figure 9: Dispersion of a_r, b_r and σ_r across pixels

the idea that, as stated in (7), the power law tails of the original prices stem from the mixture of lognormal distributions with different mean and variance. At the same time, the analysis in the previous section showed that the fat tails of the price distribution remain largely unchanged for the bubble period (i.e., the late 1980s and the first half of the 1990s) even after controlling for the size effect. This suggests that there still remains some mixture of lognormal distributions.

In this section, we test the hypothesis that the power law tails of the size-adjusted price distribution during the bubble period arise due to the mixture of different lognormal distributions corresponding to different regions. To do so, we start by decomposing the size-adjusted price distribution into the sum of conditional distributions:

$$\Pr\left(\tilde{P}_{i_r t} = p\right) = \sum_{\theta} \Pr\left(\tilde{P}_{i_r t} = p \mid \theta_{rt} = \theta\right) \Pr\Big(\theta_{rt} = \theta\Big), \tag{12}$$

where $\tilde{P}_{i_r t}$ denotes the size-adjusted price for a house located in region r, which is defined by $\tilde{P}_{i_r t} \equiv P_{i_r t} \exp(-a_{rt} S_{i_r} - b_{rt})$. The vector of parameters θ_{rt} is defined by

$$\theta_{rt} \equiv (a_{rt}, b_{rt}, \sigma_{rt}), \tag{13}$$

where the parameters a_{rt}, b_{rt}, and σ_{rt} are the coefficient on the house size variable, the constant term, and the standard deviation of the disturbance term in equation (10), but it is assumed in this section that they could differ depending on the location. The location effect is fully controlled for in the conditional distributions $\Pr(\tilde{P}_{i_r t} = p \mid \theta_{rt} = \theta)$, so that they should be lognormal. According to equation (12), the distribution of $\tilde{P}_{i_r t}$ is a mixture of these lognormal distributions, each of which is for a different region.

We first examine the distribution of θ_{rt} across different regions. Specifically, we divide the Greater Tokyo Area into pixels of 0.033 degrees latitude and 0.033 degrees longitude or roughly 3.3 by 3.3 km.[15] Then, using size-adjusted prices within a pixel, we run a regression of the form

$$\ln P_{i_r t} = a_{rt} S_{i_r} + b_{rt} + \eta_{i_r t} \tag{14}$$

for each combination of r and t and obtain $\hat{\theta}_{rt} \equiv (\hat{a}_{rt}, \hat{b}_{rt}, \hat{\sigma}_{rt})$. The regression results are presented in Figure 9.[16] The three panels on the left show the CDFs of $\hat{a}_{rt}$, while the panels in the middle and those on the right respectively show the CDFs of $\hat{b}_{rt}$ and $\hat{\sigma}_{rt}$. The CDFs of $\hat{a}_{rt}$ indicate that a is less dispersed across pixels for the period of the bubble and its collapse (1987-1995) than in the other years. On the other hand, the CDFs of $\hat{b}_{rt}$ and $\hat{\sigma}_{rt}$ show that these parameters are more highly dispersed during the same period, implying that the sharp price hike during the bubble period was concentrated in particular pixels. Put differently, the housing market was segmented during this period.

Next, we investigate whether the conditional distributions are close to a lognormal distribution. Using the estimates of θ_{rt} obtained from the regression, we calculate the size-adjusted prices for each pixel, which is given by:

$$\tilde{P}_{i_r t} \equiv \left[P_{i_r t} \exp(-\hat{a}_{rt} S_{i_r} - \hat{b}_{rt}) \right]^{1/\hat{\sigma}_{rt}}. \tag{15}$$

The estimated CDFs of $\tilde{P}_{i_r t}$ are presented in Figure 10 for the years 1986, 1990, 1994, 1998, 2002, and 2006. Note that each of the six panels contains four different lines, each of which corresponds to a different pixel size, namely 4.190 by 4.190 degrees, 0.524 by 0.524 degrees, 0.263 by 0.263 degrees, and 0.033 by 0.033 degrees. The results for 1998, 2002, and 2006 indicate that the CDFs are very close to a lognormal distribution, irrespective of pixel size. This is not very surprising given that, as we saw in the previous section, the CDFs in these years were already close to a lognormal distribution before controlling for the location effect. For the period of the bubble and its collapse, we see more interesting results: for 1986, 1990, and 1994, the estimated CDF tends to be closer to a lognormal distribution the smaller the pixel size.[17]

[15] Note that one degree is approximately 100 km.

[16] In conducting these regressions, we use only those pixels with more than twenty transactions in a year. The number of pixels used in the regressions is about 300 for each year.

[17] It should be noted that the estimated CDF does not fully converge to a lognormal even in the case of the

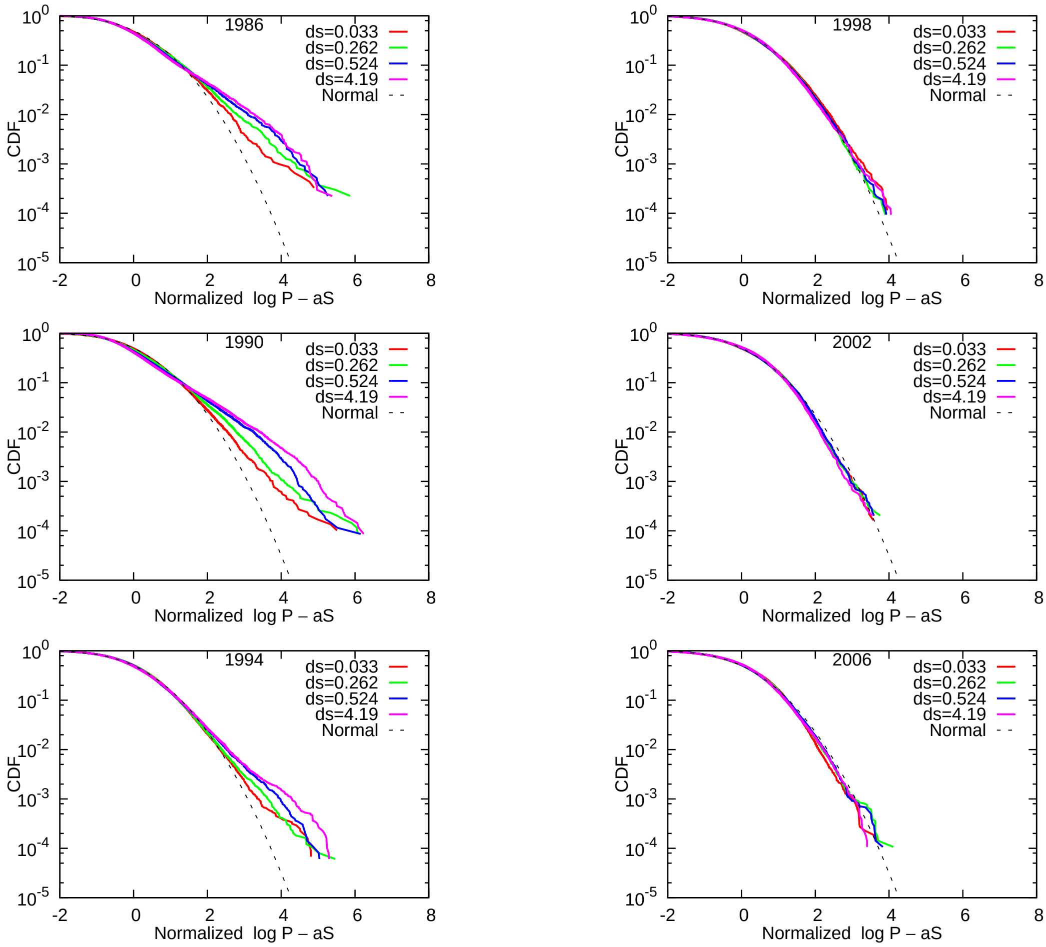


Figure 10: Cumulative distributions of size-adjusted house prices for different pixel sizes

In sum, the analysis in this section shows that the distribution of size-adjusted prices *within* a pixel is fairly close to a lognormal distribution even for the period of the bubble and its collapse, but its mean and standard deviation are highly dispersed *across* different pixels. As a result, the sum of these lognormals turns out to be far from a lognormal distribution during this period. In other words, heterogeneity across pixels in terms of the mean and standard deviation is the source of the fat upper tail of the size-adjusted price distribution during the period of the bubble and its collapse.

smallest pixels. It may be the case that the CDF becomes much closer still to a lognormal distribution if we were able to reduce the pixel size even further. Unfortunately, we cannot do so because of the limited number of observations.

6 Summary and Some Policy Implications

In this paper, we found that the cross-sectional distribution of house prices in the Greater Tokyo Area has a fat upper tail and that the tail part is close to a power law distribution. On the other hand, the cross-sectional distribution of house sizes measured in terms of floor space has less fat tails than the price distribution and is close to an exponential distribution. We proposed a hedonic model consistent with these findings and, using data for Greater Tokyo, confirmed that size-adjusted prices follow a lognormal distribution except for the period of the asset bubble and its collapse, for which the price distribution remains asymmetric and skewed to the right even after controlling for the size effect. As for the period of the bubble and its collapse, we found some evidence that the sharp price movements were concentrated in particular areas, and that this spatial heterogeneity is the source of the fat upper tail.

The analysis in this paper shows that the cross-sectional distribution of size-adjusted prices is very close to a lognormal distribution during regular times but deviates substantially from a lognormal for the bubble period. This suggests that the shape of the size-adjusted price distribution, especially the shape of the tail part, may contain information useful for the detection of housing bubbles. That is, the presence of a bubble can be safely ruled out if recent price observations are found to follow a lognormal distribution. On the other hand, if there are many outliers, especially near the upper tail, this may indicate the presence of a bubble, since such price observations are very unlikely to occur if prices follow a lognormal distribution. This method of identifying bubbles is quite different from conventional ones based on aggregate measures of housing prices, which are estimated either by hedonic or repeat-sales regressions, and therefore should be a useful tool to supplement existing methods.

References

[1] Baba, N., S. Nishioka, N. Oda, M. Shirakawa, K. Ueda, K., and H. Ugai (2005), "Japan's Deflation, Problems in the Financial System, and Monetary Policy," *Monetary and Economic Studies*, 23(1), 47-111.

[2] Cochrane, J. H. (2002), "Stocks as Money: Convenience Yield and the Tech-Stock Bubble," *NBER Working Paper* 8987.

[3] Cropper, M. L., L. B. Deck, and K. E. McConnell (1988), "On the Choice of Functional Form for Hedonic Price Functions," *Review of Economics and Statistics* 70(4), 668-675.

[4] Del Castillo, J., and P. Puig (1999), "The Best Test of Exponentiality Against Singly Truncated Normal Alternatives," *Journal of the American Statistical Association* 94, 529-532.

[5] Diewert, W. E. (2003), "Hedonic Regressions: A Consumer Theory Approach," in R. C. Feenstra and M. D. Shapiro (eds.), *Scanner Data and Price Indexes*, National Bureau of Economic Research Studies in Income and Wealth, Vol. 64, Chicago, IL: University of Chicago Press, 317-48.

[6] Diewert, W. E., J. de Haan, and R. Hendriks (2010), "The Decomposition of a House Price Index into Land and Structures Components: A Hedonic Regression Approach," *Discussion Paper* 10-01, University of British Columbia.

[7] Feller, W. (1968), *An Introduction to Probability Theory and Its Applications.* Vol. 1, Third Edition, New York: Wiley.

[8] Gabaix, X. (2008), "Power Laws in Economics and Finance," *NBER Working Paper* 14299.

[9] Greene, W. (2003), *Econometric Analysis*, Fifth Edition, Upper Saddle River, NJ: Prentice Hall.

[10] Gyourko, J., C. Mayer, and T. Sinai (2006), "Superstar Cities," *NBER Working Paper* 12355.

[11] Himmelberg, C., C. Mayer, and T. Sinai (2005), "Assessing High House Prices: Bubbles, Fundamentals, and Misperceptions," *Journal of Economic Perspectives* 19, 67-92.

[12] Hoshi, T., and A. Kashyap (2000), "The Japanese Banking Crisis: Where Did It Come from and How Will It End?" *NBER Macroeconomics Annual 1999*, Volume 14. MIT, 2000. 129-212.

[13] Kiyotaki, N., and J. Moore (1997), "Credit Cycles," *Journal of Political Economy* 105(2), 211-248.

[14] Maattanen, N., and M. Tervio (2010), "Income Distribution and Housing Prices: An Assignment Model Approach," *CEPR Discussion Paper* 7945.

[15] Malevergne, Y., V. Pisarenko, and D. Sornette (2011), "Testing the Pareto Against the Lognormal Distributions with the Uniformly Most Powerful Unbiased Test Applied to the Distribution of Cities," *Physical Review E* 83(3), 036111.

[16] Mankiw, N. G., and D. N. Weil (1989), "The Baby Boom, the Baby Bust, and the Housing Market," *Regional Science and Urban Economics* 19(2), 235-258.

[17] McMillen, D. P. (2008), "Changes in the Distribution of House Prices Over Time: Structural Characteristics, Neighborhood, or Coefficients?" *Journal of Urban Economics* 64(3), 573-589.

[18] Morinobu, S. (2006), "The Rise and Fall of the Land Myth in Japan: Some Implication to the Chinese Land Taxation," PRI Discussion Paper Series No.06A-08, Policy Research Institute, Ministry of Finance.

[19] Nishimura, K. (2011), "Population Ageing, Macroeconomic Crisis and Policy Challenges," Prepared for the Panel "The General Theory and the Policy Reponses to Macroeconomic Crisis" at the 75th Anniversary Conference of Keynes' General Theory, University of Cambridge, June 19-21, 2011.

[20] Okina, K., M. Shirakawa, and S. Shiratsuka (2001), "The Asset Price Bubble and Monetary Policy: Japan's Experience in the Late 1980s and the Lessons," *Monetary and Economic Studies* (Special Edition) 19(2), 395-450.

[21] Okina, K., and S. Shiratsuka (2002), "Asset Price Bubbles, Price Stability, and Monetary Policy: Japan's Experience," *Monetary and Economic Studies* 20(3), 35-76.

[22] Pareto, Vilfredo (1896), *Cours d'Economie Politique*. Lausanne and Paris: Rouge.

[23] Shimizu, C., K. G. Nishimura, and Y. Asami (2004), "Search and Vacancy Costs in the Tokyo Housing Market: An Attempt to Measure Social Costs of Imperfect Information," *Review of Urban and Regional Development Studies* 16(3), 210-230.

[24] Shimizu, C., K. G. Nishimura, and T. Watanabe (2010), "Housing Prices in Tokyo: A Comparison of Hedonic and Repeat-Sales Measures," *Journal of Economics and Statistics*, Volume 230, Issue 6, 792-813.

[25] Shimizu, C., and T. Watanabe (2010), "Housing Bubbles in Japan and the United States," *Public Policy Review* 6(3), 431-472.

[26] Shimizu, C., and T. Watanabe (2013), "Sticky Housing Rents and Monetary Policy: The Japanese Experience," Paper presented at the KDI Workshop on "Real Estate Driven Systemic Risk: Country Cases and Their Policy Implications," December 13-14, 2012, Seoul.

[27] Shiratsuka, S. (2005), "The Asset Price Bubble in Japan in the 1980s: Lessons for Financial and Macroeconomic Stability," in Real Estate Indicators and Financial Stability, BIS Papers No. 21, 42-62.

[28] Taleb, N. N. (2007), *The Black Swan: The Impact of the Highly Improbable*, New York: Random House, Inc.

[29] Takáts, E. (2012), "Aging and House Prices," *Journal of Housing Economics* 21(2), 131-141.

[30] Triplett, J. (2004), "Handbook on Hedonic Indexes and Quality Adjustments in Price Indexes: Special Application to Information Technology Products," *OECD Science, Technology and Industry Working Papers* 2004/9.

[31] Ueda, K. (1998), "The Japanese Banking Crisis in the 1990s," Bank for International Settlements.

[32] Van Nieuwerburgh, S., and P. O. Weill (2010), "Why Has House Price Dispersion Gone Up?" *Review of Economic Studies* 77(4), 1567-1606.