House Price Index Construction in the Nascent Housing Market: The Case of China

Jing Wu
Yongheng Deng
Hongyu Liu

September 2012

House Price Index Construction in the Nascent Housing Market: The Case of China

Jing WU

Department of Construction
Management, Tsinghua University
and Institute of Real Estate Studies,
National University of Singapore
ireswujing@mail.tsinghua.edu.cn

Yongheng DENG

Institute of Real Estate Studies,
National University of Singapore
ydeng@nus.edu.sg

Hongyu LIU

Department of Construction
Management, Tsinghua University
liuhy@tsinghua.edu.cn

Abstract

Most existing house price index construction methods are developed mainly based on transaction data from the secondary housing market, and are not necessarily suitable for the nascent housing markets where a predominant portion of housing transactions are new units. Using the booming market in China as an example, we evaluate and compare the performances of three most common house price measurement methods in the newly-built housing sector, including the simple average method without quality adjustment, the matching approach with the repeat sales modeling framework, and the hedonic modeling approach. Our analyses suggest that the simple average method fails to account for the substantial complex-level quality changes over time of sales during our sample period, and the matching model fails to control for the effect of developers' pricing behaviors when adopted in the newly-built sector, hence both are downward biased. Based on this finding, we apply a hedonic method, which allows us to control for both quality changes over time of sales and developers' pricing behaviors, to 35 major newly-built housing markets and provide the first multi-city constant-quality house price index in China. The new index reveals that the current Chinese housing market is facing a greater risk of mispricing than reported by the existing official metrics.

Key words: house price index; hedonic method; quality change; nascent housing market

JEL Classification: C43, C81, R31

1.   Introduction

The dramatic rise of house prices in major Chinese cities has generated global interest among investors, policy makers, scholars and journalists. Due to China's rising economic importance, there has been growing concern that a potential house price bubble in China and its aftermath would be a catastrophe not only to China but also to the world economy. In February 2011, IMF (2011) explicitly listed "a potentially steep price correction in Chinese property markets" as one major risk in global recovery from the financial crisis. Accordingly recent researches have sought to provide a more rigorous test of the sustainability of Chinese house prices by detecting and measuring the potential mispricing,[1] which should undoubtedly depend on an accurate measurement of the level and movement of house prices.

A high-quality house price indicator is also of enormous importance in the policy perspective. Policy makers in most major economics including China are increasingly sensitive to changes in housing market conditions, while the house price is always adopted as one major yardstick (Crowe *et al*, 2011). The appropriateness of government interventions directly anchoring on house price is debatable and well beyond the scope of this research, but it highlights the importance of a reliable house price index – the incorrect signal from systematic errors in measurement may lead to improper intervention policies and great damage to the market.

Currently two official house price indices are regularly updated in China, namely, the "Average Selling Price of Newly-Built Residential Buildings" (abbreviated as the

---

[1] Among others, see Hui and Shen (2006), Yang and Shen (2008), Chen, Hao and Stephens (2010), Ahuja *et al* (2010), Dreger and Zhang (2010), and Wu, Gyourko and Deng (2012) for example. There are also a large number of related studies published in Chinese.

"Average Price Index" henceforth),[2] and the "Price Indices for Real Estate in 70 Large- and Medium-sized Cities" ("70 Cities Index"),[3] both calculated and reported by the National Bureau of Statistics of China (NBSC). They provide almost the only consistent sources for description and analysis of Chinese house prices. Unfortunately, both these indices are mistrusted and widely criticized at present. Figure 1 provides the real quarterly series for newly-built houses in four "first-tier" cities as the example. First, these two indicators published by the same official agency seriously conflict, which inevitably puzzles users (Ahuja *et al*, 2010). Secondly and even more importantly, the magnitudes of price growths suggested by these indices are widely suspected to be underestimated, especially for the "70 Cities Index". It provides an almost flat real house price path in Shanghai, Guangzhou, and Shenzhen between 2005 and 2010, and only a very modest price growth in Beijing, which deviates significantly from common experience in the market.[4] These suspicions and criticisms finally forced the NBSC to suspend reporting any house price indicator at the national level since February 2011.[5]

* Insert Figure 1 about here*

---

[2] The "Average Price Index" covers all Chinese cities since mid-1990s. It is the obligation by law for all developers in China to regularly report stated business indicators to the government statistics agency, including the total volume (in floor area) of newly-built housing units sold within this period and the total price of these units. By aggregating these reported figures at various levels, and dividing the total price by total floor area of the transacted units, the average house prices are calculated and reported at the city, province, and national level, respectively.

[3] The "70 Cities Index" originally covered 35 major cities since 1997 and then expanded to 70 cities in 2005. In each month technicians from local statistics authorities are sent to sample housing complexes to collect raw information on house transaction prices. The matching approach is then adopted to calculate the index, which will be discussed in detail later in Section 3.2.

[4] As a more well-known example, while the government and the general public in China were highly concerned with the apparent surge in house prices in 2009, the "70 Cities Index" suggested that nominal house price at the national level only increased by 1.5% in 2009 compared with the previous year, which immediately generated great suspicions and criticisms. See the reports from *Financial Times* ("Fears of China Property Bubble Grow", Mar 10, 2010) or *China Daily* ("Doubts over Increase in Property Price", Feb 27, 2010) for more details.

[5] See the reports from *Wall Street Journal* ("China Scraps High-profile Property Data", Feb 17, 2011) for more details.

In this paper we focus on the appropriate measurement of house prices in China, especially in the newly-built housing markets. There are several reasons to set measurement of newly-built house prices as our major concern. First, in current Chinese housing markets a major portion of housing transactions concentrate in the newly-built sector, which should also be an indispensable stage in other nascent housing markets' initial developments. Second, considering its potential influences on macroeconomics, for policy makers the newly-built house price index could be even more informative than the resale price indicator: besides the effects on household sector, changes in newly-built house prices directly determine the performance of real estate and related industries, as well as the quality of construction and land loans. Third, one major challenge for measuring (resale) house price is to isolate the age effect from the pure price changes (Chau et al, 2005). By focusing on the newly-built sector we can perfectly avoid this problem. Finally and most importantly, most existing researches on house price index constructions are (explicitly or implicitly) based on the resale sector. However, the newly-built sector is distinguished from the resale sector in several aspects: the quality changes over time of sales are always larger in the newly-built sector since it is the market for flow supplies, and there also exist some unique factors affecting new units' prices such as developers' pricing strategies. Consequently the methods developed in the resale sector are not necessarily suitable for newly-built house price measurement; instead careful methodological re-evaluations and comparisons based on thorough understandings of newly-built residential markets are called for.[6]

---

[6] As another unique factor, currently the reported transaction prices of resale units are not considered of high quality in China, at least partially because an unknown number of buyers/sellers are reporting lower values to avoid transaction taxes and capital gain taxes. By contrast, house price data in the first hand sales are much more reliable, since developers are facing stricter audits.

Three general price estimation methodologies are included in the comparisons: the simple average methods without any adjustment for quality change or variability among housing units, the hedonic method, and the repeat sales method. The procedures for the first two are generally consistent with those employed in existing researches and practices, but the conventional repeat sales method developed by Cash and Shiller (1987, 1989) could not be directly applied in the newly-built sector due to the lack of a sample of paired repeat-sales transaction data. Instead the matching approach developed by McMillen (2008), Deng, McMillen and Sing (2012) is introduced: each observation on the sale of a new unit is matched with the most similar unit within the same complex sold previously, and the classical weighted repeat sales procedures are then applied to these matched pairs to estimate the price index.

The performances of these three methodologies are compared via both theoretical analysis and empirical test based on a unique, large transaction dataset from a typical Chinese newly-built housing market. The results suggest that the simple average pricing methods fail to account for housing quality adjustment and hence are biased during the sampling period due to the ongoing trends of housing suburbanization and building density increasing in most Chinese cities. On the other hand, although the matching approach perfectly controls for the complex-level quality changes and has been proved to work well in the resale sector, it does not account for developers' pricing behaviors in the newly-built sector, hence can also lead to a downward bias during our sampling period. By contrast, we adopt a hedonic method that allows us to control for both factors that are common in many nascent housing market, which is suggested as the preferable choice for measuring newly-built house prices.

These findings also suggest that, since the "Average Price Index" is calculated by the weighted average formula without quality adjustment, and the "70 Cities Index" with a simplified form of the matching approach, these two existing official indices are both problematic in the methodological aspect. As an initial attempt to correct such biases, we apply the hedonic modeling method to the newly-built housing markets in 35 major Chinese cities. Recent house price appreciation rates are significantly higher according to our results than reported by the two official indexes, which implies the current Chinese housing markets are facing an even greater risk of mispricing. We suggest the intervention policies, as well as other market participants' investment decisions, need to be re-evaluated based on this new finding.

This paper contributes to the literature in two folds. From the methodological aspect, while most existing literatures focus on the resale housing sector, the analyses on behaviors of the newly-built sector and their effects on price measurement could help better design the construction method of newly-built house price index, which is important to China as well as nascent markets in other emerging and developing economies. From the practical aspect, this paper provides the first multi-city hedonic house price index in China. It will not only help market participants, researchers and policy makers better understand market conditions, but also facilitate more rigorous empirical researches on Chinese housing issues.

The paper proceeds as follows. The next section describes the unique features and behaviors of the newly-built housing markets in China, especially factors with potential effects on house price index estimation. Section 3 briefly reviews the three major house price estimation approaches and theoretically discusses whether they are able to control

for these potential problems, followed by the empirical test based on a unique transaction dataset in Section 4. The preferable method suggested by these analyses, the hedonic method, is then applied to 35 major Chinese cities, and the new index is discussed in Section 5. Section 6 concludes.

2. Features of Chinese Newly-Built Housing Markets

2.1 Housing Reform and Booming of Newly-Built Housing Markets

In most Chinese cities there has only been a truly private market for housing units since the 1990s. Before that the Chinese urban housing sector was dominated by the housing provision system, in which almost all urban housing units were built and owned by the employers (work units), and allocated to individual households at low rent. The private housing markets did not exist during that period.

The Chinese government embarked on the housing reform in the early 1980s, which expanded and advanced step by step during the following two decades. Finally in 1998, the State Council issued the 23rd Decree, according to which work units were no longer allowed to develop new housing units for their employees in any form. Instead, they had to integrate all implicit housing benefits into employees' salaries, and the households had to buy or rent their residential housing units in the private housing markets (or public housing system for low-income households).[7]

This led to the dramatic booming of the private housing markets in most Chinese cities since late 1990s. During the housing provision era the Chinese urban households in general consumed insufficient housing service, and this misallocation resulted in a

---

[7] Among others, see Wang (2001, 2003), Wu, Xu and Yeh (2007) for more details about Chinese housing reform.

substantial shift out in existing homeowners' housing demand after the reform (Wang, 2011). Meanwhile, the continuous economic growth and urbanization also greatly contribute to the rapid expansion in urban housing demand (Chen, Guo and Wu, 2011; Wu, Gyourko and Deng, 2012). Driven by these huge demands, the annual volume of private housing units completed increased from 140 million square meters in 1998 to over 610 million in 2010, while its share in all housing units completed more than doubled from 30% to over 70% during the same interval (Figure 2).

* Insert Figure 2 about here*

As a direct result of the large volume of new supply, currently housing markets in most Chinese cities are dominated by the newly-built sector. According to the statistics provided by the Ministry of Housing and Urban-Rural Development of China (MOHURD), in 2010 the new units accounted for about 64% in floor area of all private housing units sold at the national level. Hence the construction of newly-built house price index is still the key task in house price measurement in current China.

The continuous prosperity of newly-built housing markets also makes the real estate industry a very important sector in Chinese economy. In 2009 the real estate industry was directly responsible for about 11% of total GDP growth at the national level, which greatly contributed to the quick recovery of Chinese economy from the global financial crisis (Deng *et al*, 2011). This historically highest contribution rate may still be significantly underestimated because the indirect influences via sectors like construction material industries were not accounted. Besides, as reported by the People's Bank of China, at the end of 2010 the outstanding balance of developer loans reached 3.12 trillion yuan RMB, or about 6.1% in all loans (as a benchmark, the balance of residential

mortgage loans was about 6.16 trillion yuan at the same time). The importance of the real estate related sectors provides another reason for us to focus on the appropriate measurement of newly-built house prices.

2.2 Complex-Level Quality Changes and Their Potential Effects on Price Measurement

As emphasized in almost every research in this field, the key challenge in house price measurement is to control for the effect of quality changes among housing units sold in different periods, which is always confounded with the pure changes of house prices. Accordingly before the methodological comparisons it is necessary to understand the major quality changes in current Chinese newly-built housing markets.

So far the urban housing sector in China has been dominated by non-landed condominium units.[8] A typical condominium complex usually contains hundreds (or even thousands) of units located in dozens of high-rise residential buildings on a contiguous land parcel. Therefore the housing attributes can be grouped into two levels: the complex-level attributes cover all locational characteristics and neighborhood amenities, as well as some physical characteristics like building type and construction quality, while the unit-level attributes include most of the physical characteristics such as unit size, floor level, and specific environmental attributes (*e.g.*, noise, view, accessibility to sunshine, *etc.*). Housing units within one complex share exactly the same complex-level attributes, but may vary in unit-level attributes.

---

[8] According to the statistics published by National Bureau of Statistics of China in the "*Statistics Yearbook of China*", the percentage of condominium units in the newly-built housing market kept around 94%-96% during the past decade in the national level. For example, in 2010 condominium units accounted for 95.5% in floor area of all newly-built housing units sold.

During the period when the reported China's house price index covers, there are substantial complex-level quality changes in current Chinese newly-built housing markets, which are mainly driven by the rapid expansion in new housing supply. In most Chinese cities, especially the developed mega cities, the scarcity in available land resources has become a major constraint in new housing supply. Two shifts in residential land usage patterns are resulted as the solutions, both of which significantly affect the newly-built housing markets in turn.

First, the urban area in most Chinese cities keeps expanding, and increasing land parcels in urban fringes are developed for residential usage. During the recent decade, the total size of developed urban area at the national level increased from 22,439 square kilometers in 2000 to 40,058 square kilometers in 2010, or an average annual growth rate of about 5.96%, while the total area of developed residential land parcels also grew at a similar speed, from 7,122 square kilometers in 2000 to 12,404 square kilometers in 2010 (5.71% annually).[9] This implies a continuous trend of suburbanization in most Chinese newly-built housing markets.

The other way to provide more housing units is to utilize available land parcels more "efficiently" by setting higher permitted floor area ratio (FAR). As an approximate estimation, the annual ratio between the total floor area of private housing completions and the area of land parcels they occupy has increased from 2.15 in 2000 to 3.57 in 2010.[10] In other words, the average FAR of residential units sold in the newly-built housing markets increased by over 60% during the recent decade.

---

[9] Source: MHURD, "*China Urban Construction Statistical Yearbook*". Besides, see Wu, Xu and Yeh (2007), Deng *et al* (2009) for example for a more detailed review on housing suburbanization and urban expansion in China.
[10] Source: Authors' calculation based on the "*Statistics Yearbook of China*" published by the NBSC.

Controlling for other factors, in general residential units with longer distances to city center or higher FAR could be expected to achieve in lower transaction prices; that is, these two trends could be regarded as trended "degrading" in new residential units' complex-level qualities. Accordingly they will make the price measurement results misleading if not adjusted properly in the index construction.

2.3 *Developers' Pricing Behaviors and Their Potential Effects on Price Measurement*

Besides quality changes over time of units sold, one unique factor affecting house price index construction in the newly-built sector is developers' pricing behaviors.

Such effects mainly result from new residential units' unique selling process, which is rather different with that of the resale units discussed in literature. In China the seller (developer) of a newly-built housing complex is required to list all the units included when it receives the selling permit from local government.[11] The developer sets the listing price for each unit, and can adjust any unsold unit's listing price during the following selling process. The households come to investigate the complex and choose their target units. As a most noteworthy fact, in most cases the potential buyers cannot bargain with the developer; in other words, the household that is interested in certain unit can only choose to accept the given listing price, or reject it and turn to other units/complexes. Thus a newly-built unit's transaction price should always equal its

---

[11] Note that presale of newly-built housing units is permitted and very popular in China, which means that developers can get the permit and sell housing units to households before the units are completed and ready to deliver. Both the developer and the buyer of a presale unit are contractually bound to complete the transaction upon completion of the unit.

listing price set by the developer. Typically the whole selling process for one newly-built housing complex would last for several months.[12]

Due to the inherent equivalency between transaction price and listing price, the new units' transaction prices observed may deviate from their "fair values" purely determined by housing attributes, but largely affected by developers' marketing and pricing behaviors. The developers may intentionally overprice or underprice, or just unintentionally misprice some units. The effect of such pricing behaviors on specific complex varies with time and developer, and thus their common patterns, if exist, should be most important here when considering house price measurement.

For example, it has been known that developers in China's housing market are commonly adopt a sales strategy that to reduce the listing prices of unsold units (or more precisely, during the period when the market price is trend up, developers tends to raise the listing prices of unsold units with a margin less than the prevailing market prices) as the selling process proceeds. The key reason comes from developers' trade-off between transaction price and expected time to sale in determining units' listing prices. In general a unit with higher listing price would result in a higher transaction price, but also a longer expected time-on-market (TOM), *vice versa*, and hence the developer has to choose a balance between these two sides to maximize its selling revenue.[13] As the selling of a newly-built complex proceeds, the scale effect in developer's marketing/selling efforts gradually mitigates with the decreasing in number of units left, which means an

---

[12] For instance, according to the statistics provided by the MOHURD, in 2010 the average time-on-market (TOM) of all newly-built housing transactions in 35 major cities was about 9.2 months.

[13] The determinants of listing prices and their effects on expected TOM in the resale housing markets have been explored in many existing researches (Haurin, 1988; Forgey, Rutherford and Springer, 1996; Anglin, Rutherford and Springer, 2003), although so far there are still no similar researches in the newly-built sector.

increasing duration cost of marketing for each unsold unit (*i.e.*, the cost per unit per day).[14] This change will force the developer to shift the balance to the TOM side by lowering listing prices for unsold units in order to reduce their expected TOMs. As a result, a unit's listing price could be expected to be negatively correlated with its TOM (or transaction sequence) in the complex.

Anchoring effect may play another role in influencing the pricing behavior in the housing market. People tend to irrationally anchor on properties' earlier transaction prices, initial asking prices or so on.[15] This effect can be expected to be especially significant in the newly-built housing market, since transactions within the same new complex are always concentrated within a few months, and hence potential buyers can easily observe the complex's past price path. Thus if a developer substantially increases the unsold units' listing price, potential buyers are very likely to reject the new prices according to their judgments based on past price level, even if such adjustment does not exceed the magnitude of market price appreciation during the same interval. This makes developers tend to anchor on earlier price level and reluctant to greatly raise the listing price of unsold units, which will also result in a negative correlation between units' listing prices and their TOM during the market booming period.[16]

---

[14] In the Chinese newly-built housing markets, typically the developers' marketing cost include rental and maintenance cost of the display space, wages and bonuses for the sales staff, advertising cost, *etc*. Most of these costs are duration costs, which means that they would happen at almost the same rate every day no matter whether any transactions are achieved. Besides, a large portion of the costs are fixed and could not be easily adjusted according to the number of units left (*e.g.*, the cost of the display space, which contributes for a major portion of the marketing cost).

[15] Among others, see Northcraft and Neale (1987), Genesove and Mayer (2001), Leung and Tsang (2011) for example.

[16] Similar phenomenon also exists in the recession period. It is difficult for developers to lower the listing price of unsold units' even if the market conditions turn down, because the households that already purchased units in the same complex always strongly oppose, or even require to refund.

As emphasized earlier, the effects of pricing behaviors discussed above only work for units within the same complex. Therefore, their effects on price measurement could be offset and hence mitigated when observations from multiple complexes are pooled together, but amplified if the price index is constructed mainly relying on intra-complex price changes. We will discuss this in detail in the following theoretical analyses.

3.   Theoretical Analysis on Method Comparison

The theoretical analyses are based on the framework of hedonic model. Based on the features discussed in last section, the hedonic model in Chinese newly-built housing markets can be expressed as:

$$P_{ijt} = \alpha \cdot OU_{it} + \lambda \cdot UU_{it} + \beta \cdot OC_{jt} + \varphi \cdot UC_{jt} + \theta \cdot PB_{ijt} + \delta_t \cdot D_{ijt} + \mu_{ijt}, \tag{1}$$

where, P_{ijt} is the observed transaction price of unit i in complex j sold at time t; U_{it} and C_{jt} are sets of unit-level and complex-level housing characteristics, respectively; considering that in reality it is almost impossible to capture all housing attributes in the available dataset, these housing characteristic variables are divided into observed (OU_{it} and OC_{jt}) and unobserved (UU_{it} and UC_{jt}) groups; PB_{ijt} reflects the effect of developers' pricing behaviors; D_{ijt} is a set of time dummies (equals 1 in period t, and 0 in other periods); μ_{ijt} is an i.i.d. error term.

While other parameters might also be of interest, the key task of house price index construction is the accurate estimation of time dummies' coefficients δ_t. Accordingly the key concern in the following theoretical analyses is whether the candidate methodologies could achieve in an unbiased estimation of δ_t or not when adopted in Chinese newly-

built housing markets, especially with the existence of the challenges discussed in the previous section.

3.1 The Simple Average Method

The simplest house price index construction method is to directly calculate the unweighted or weighted average of units' observed transactions prices. In the framework of eq.(1), this equals only to include the time dummies as explanatory variables of house prices, leaving all other factors in a new error term ε_{ijt} (eq.(2)).

$$P_{ijt} = \delta'_t \cdot D_{ijt} + \varepsilon_{ijt},$$

where, $\varepsilon_{ijt} = \alpha \cdot OU_{it} + \lambda \cdot UU_{it} + \beta \cdot OC_{jt} + \varphi \cdot UC_{jt} + \theta \cdot PB_{ijt} + \mu_{ijt}$

$$(2)$$

In eq.(2), if none of the characteristics affecting housing units' prices experience a trended change over time, and the sample volume is large enough to offset the units' quality variance in each period, the new error term ε_{ijt} can still meets i.i.d. and hence the coefficients of δ'_t is the unbiased estimate of δ_t, although the variance of ε_{ijt} would be substantially larger than that of μ_{ijt}. In this case the simple average methods could achieve in an unbiased house price index without any requirements on housing characteristics information. This advantage in feasibility makes the simple average method well fit the data condition in emerging housing markets like China, and this is one most important reason why the "Average Price Index", the first house price index in China, chose the weighted average formula.

However, the simple average methods can be biased by not controlling for quality changes. Any trended change in housing characteristics over time of sales will lead to a correlation between time dummies and the error term ε_{ijt} in eq.(2), and hence a bias in the estimated δ'_t. Therefore the complex-level quality changes discussed in Section 2.2

15

should be an important issue here. Since both the ongoing trends of housing suburbanization and density increasing could be regarded as trended "degrading" of transacted units' qualities over time, they would lead to a downward bias in the results of simple average methods, although the magnitude of bias may vary with time and city.

3.2   *The Matching Approach with the Repeat Sales Modeling Framework*

Considering the importance of quality control in house price measurement, two quality-adjusted house price index construction methods have been developed in literature. The repeat sales method aims to take advantage of the inherent homogeneousness of the same unit (Bailey, Muth and Nourse, 1963; Case and Shiller, 1987, 1989). By restricting the sample to housing units sold at least twice, and with the assumption of homogeneousness of the same unit between two sale dates, the repeat sales method can achieve in constant quality house price index without detailed information on housing characteristics. Although the limitation of the repeat sales method has been well-documented,[17] it is still the "gold standard" of house price index construction at present, especially in the U.S. and other matured housing markets in the developed economies.

But the repeat sales method cannot be directly applied in the newly-built housing market, where obviously should not exist real repeat sale units. However, as suggested by McMillen (2008), Deng, McMillen and Sing (2012), instead of strictly focusing on repeat sales only, we can choose to match each transacted unit with the most similar unit sold previously, and then apply the repeat sales approach to these matched pairs. Suppose unit i in complex j sold in period t is matched with the unit k in complex l sold in period

---

[17] Among others, see Case, Pallakowski and Watcher (1991), Haurin and Hendershott (1991), Case and Szymanoski (1995), Deng and Quigley (2008), Nagaraja, Brown and Wachter (2010) for a literature review on the application and potential problems of the repeat sales method.

$\tau\left(t>\tau\right)$ based on certain matching procedures (*e.g.*, via the propensity score method), their difference in transaction price is:

$$P_{ijt} - P_{kl\tau} = \alpha \cdot \left(OU_{it} - OU_{k\tau}\right) + \beta \cdot \left(OC_{jt} - OC_{l\tau}\right) + \left(\delta_t' D_{ijt} - \delta_\tau' D_{kl\tau}\right) +$$
$$\gamma \cdot \left(UU_{it} - UU_{k\tau}\right) + \varphi \cdot \left(UC_{jt} - UC_{l\tau}\right) + \theta \cdot \left(PB_{ijt} - PB_{kl\tau}\right) + \left(\mu_{ijt} - \mu_{kl\tau}\right) \quad (3)$$

Particularly, if we only allow for matching between units within the same complex, eq.(3) becomes:

$$P_{ijt} - P_{kj\tau} = \alpha \cdot \left(OU_{it} - OU_{k\tau}\right) + \left(\delta_t' D_{it} - \delta_\tau' D_{k\tau}\right) +$$
$$\gamma \cdot \left(UU_{it} - UU_{k\tau}\right) + \theta \cdot \left(PB_{ijt} - PB_{kl\tau}\right) + \left(\mu_{ijt} - \mu_{ij\tau}\right) \quad (4)$$

In eq.(4) all complex-level attribute terms (either observed or not) are dropped due to the inherent homogeneousness between units within the same complex, which implies that it can perfectly control for the complex-level quality changes described in Section 2.2 even without any information on complex-level characteristics. The observed unit-level attributes are controlled to be homogeneous to the largest extent via the matching procedures, while the remained differences, as well as the unobserved unit-level characteristics, can also be offset if the sample volume of matched pairs is large enough. Then if the change in effect of developers' pricing behaviors, ($PB_{it} - PB_{k\tau}$), is uncorrelated with the term of ($\delta_t D_{ijt} - \delta_\tau D_{ij\tau}$), we can apply the standard weighted repeat sales procedures to these matched pairs and achieve in an unbiased price index.[18] This is why it has been proved to work well in price index construction in the secondary markets (Deng, McMillen and Sing, 2012).

---

[18] Theoretically, besides the repeat sales method, we could also apply the hybrid method developed by Case and Quigley (1991), Quigley (1995), Hill, Knight and Sirmans (1997) to the matched pairs, where information from both repeat sales and single sales are both utilized. However, since only few (if any) transactions are unmatched in the matching approach, which means that only little information is dropped in the estimation of repeat sales method for these matched pairs, the hybrid method's improvement compared with the repeat sales method could be expected to be very limited. Accordingly we do not discuss the hybrid method detailedly in this paper.

A simplified form of this approach has been adopted in the "70 Cities Index" in China. For each housing complex, the average transaction price is calculated in each month and compared with that of the same complex in the previous month. The monthly house price growth rate at city level is then calculated as the average (weighted by transaction volume) of all complexes' growth rates in the corresponding month. This equals to apply the matching procedures at the complex level by matching each complex in each month with itself in the previous month.

As a major advantage, the complex-level quality changes are perfectly controlled in the matching approach. However, if the pattern in developers' pricing strategy discussed in Section 2.2.2 do exist and units' listing prices are negatively correlated with their selling durations, the premise of no trended change in developers' pricing behaviors' effect would be violated. Instead the $\left(PB_{it} - PB_{k\tau}\right)$ term will be negatively correlated with the $\left(\delta_t D_{ijt} - \delta_\tau D_{ij\tau}\right)$ term, and also result in a downward bias in the price index.

3.3   The Hedonic Method

The other quality-adjusted house price index construction method is the hedonic method, which seeks to incorporate the quality adjustment directly based on the estimation results of the hedonic model. In its most frequently adopted form (time dummy form), housing transactions from multiple periods are pooled into a single hedonic model to estimate the vector of time dummy coefficients δ_t, and then the house price index is calculated based on δ_t (Kain and Quigley, 1970; Thibodeau, 1989; Kiel and Zabel, 1997; Gourieroux and Laferrere, 2010).

A major challenge for the hedonic method is the high data requirement. Besides transaction price, detailed housing attribute information is required for the proper

implementation of the hedonic method. As pointed out by Clapham *et al* (2006), the lack of standard and extensive datasets of housing attributes in U.S. makes it difficult to apply the hedonic method in official statistics. Similar reasons also explain why it has not been applied in China yet.

But things started to change recently. In China all housing transaction contracts should be registered in local housing authorities, and since 2003 municipal housing authorities in major cities gradually started to electronically record each transaction's key information via the so-called "Real Estate Market Information System (REMIS)". Finally in 2006 this became a compulsory task for local housing authorities according to the requirement of MOHURD. In the official technical code released in April 2007, the MOHURD explicitly prescribes the list of variables to be recorded, as well as the definition and data format of each variable included. Since this official variable list covers transacted units' major housing attributes, it provides a basic data foundation for the implementation of the hedonic method in China.

Nevertheless, this does not imply that all factors affecting house prices could be captured in the dataset. Some complex- or unit-level attributes may still be missing in REMIS, besides it is also very difficult to explicitly measure the effect of developers' pricing behaviors, which suggests that the omitted variable issue still matters. In this case the hedonic model in eq.(1) should be re-written as eq.(5), where all effects of omitted variables are grouped into the error term.

$$P_{ijt} = \alpha \cdot OU_{it} + \beta \cdot OC_{jt} + \delta_t^{''} \cdot D_{ijt} + s_{ijt} ,$$

where, $s_{ijt} = \gamma \cdot UU_{it} + \varphi \cdot UC_{jt} + \theta \cdot PB_{ijt} + \mu_{ijt}$

$$(5)$$

Recall the earlier discussion on eq.(2), the estimated $\delta_t^{''}$ could be an unbiased estimation of δ_t only if the new error term s_{ijt} still meets i.i.d., which means that neither the unobserved housing characteristics nor the effect of developers' pricing behaviors experiences a trended change over time. Therefore the most important task in raw data preparation is to capture the complex-level attributes undergoing continuous quality changes, especially the two trends discussed in Section 2.2. On the other hand, the effects of omitted unit-level attributes or developers' pricing behaviors could be expected to be offset when observations from multiple complexes are pooled, which will lead to a larger variance of the error term, but not a substantial bias in the long term trend of price growth.

As a summary of the theoretical analyses, of the three candidates, the simple average methods without quality adjustment are vulnerable to bias resulted from ongoing complex-level quality changes in most Chinese cities, while the matching approach with the repeat sales modeling framework may also be biased when adopted in the newly-built sector due to developers' pricing behaviors. Only the hedonic method could be expected to well handle both these effects and achieve in an unbiased estimation, if information on complex-level quality changes is included in the underlying dataset and hence reasonably reflected in the hedonic model. These suggest the hedonic model to be a preferable choice for price measurement in Chinese newly-built housing markets, and we will further test this in the following empirical analyses.

4. Empirical Analysis

4.1 Data

One typical large Chinese city is selected for the empirical test.[19] As most other Chinese cities, the newly-built housing market in this sample city has experienced a rapid expansion during recent years. In our sample period from 2004 to 2009, the annual transaction volume of newly-built private residential units kept increasing from 4.21 million square meters in 2004 to 14.56 million square meters in 2009, despite a short and small downturn in the market recession in 2008.

Supported by the local housing authority, we are able to export full-sample transaction data of new units from the REMIS in this city. During the period of 2004-2009, 539,067 newly-built non-landed condominium units in 2,534 complexes were sold in the 6 districts in the city. The variables available in this dataset are listed in Table 1. Besides the transaction date and price, the complex-level attributes include distance to city center, distance to nearest subway station, floor area ratio, complex size, green space ratio, and expected completion date, while the unit-level attributes include room area, floor level, and total floor level. This is also the typical information available in REMIS in most Chinese cities under current conditions.

* Insert Table 1 about here*

4.2 Evidences on Factors Affecting Price Measurement

Although our main focus is the performance of the house price construction methods, it is difficult to directly achieve in any definitive conclusions simply by comparing the results from different methods, because it is impossible to judge what the

---

[19] We are required by the data provider not to report the name of the city.

"true" price path should be. Our strategy here is to test the existences of the factors affecting price index construction discussed in Section 2, in order to indirectly prove the existence of potential problems in corresponding index construction methodologies as suggested in the theoretical analyses.

4.2.1 Results of the Hedonic Model

The baseline hedonic model is estimated as the first step. The dependent variable is the transaction price of each unit in yuan RMB per square meter of floor area (in natural logarithmic term), while the explanatory variables include the distance to city center (with a square in the central region as the widely-accepted landmark of the city center; *D_CENTER*; in log term), the distance to nearest subway station (*D_SUBWAY*; in log term), complex size in land area (*SIZE*; in log term), floor area ratio (*FAR*; in log term), green space ratio (*GREEN*), interval between listing date and expected completion date (*PRESALE*; in log term), unit size in floor area (*UNITSIZE*; in log term), total floor level (*TOTALFLOOR*; in log term), ratio between floor level and total floor level (*FLOOR*), whether the unit is on the first floor (*FIRSTFLOOR*) or top floor (*TOPFLOOR*), and the monthly time dummies. The model is estimated via OLS, and the results are listed in Table 2 (column 1).

* Insert Table 2 about here*

Generally the results are consistent with expectations. As for the complex-level attributes, units in complexes nearer to city center or subway station, with larger size, lower density or more green space could achieve in higher transaction prices, controlling for other factors. Besides, the presale complexes listed too early before expected completion date will get lower prices, which may result from the additional fund cost and

higher risk for buyers. As for the unit-level attributes, holding other factors constant, in general units in higher buildings or on higher floors could get higher prices, with units on the first and top floors as the exception. The larger units are more expensive even in terms of unit price.

4.2.2 Complex-Level Quality Changes

The two trends of complex-level quality changes discussed in Section 2.2 both significantly exist in this sample city. As depicted by the dark line in Figure 3, the average distance to city center of transacted new units increased from 4.20 km in 2004 to 6.06 km in 2007, which provides a clear evidence of housing suburbanization in the newly-built sector during this period. The spatial pattern generally remained stable after that in 2008 and 2009. Meanwhile, the average FAR of the new units sold gradually increased from 3.01 in 2004 to 4.37 in 2009 (the light line in Figure 3).

* Insert Figure 3 about here*

As suggested in Section 3.1, these complex-level quality changes should significantly affect the results by simply average methods without quality adjustment. The effect of the housing suburbanization is especially large because location is a most important fact in determining house prices. According to the estimated coefficient of *D_CENTER* in column 1 in Table 2, if the shift in average distance to city center from 4.20 km to 6.06 km within four years is not properly controlled, the accumulated newly-built house price growth rate during these 4 years would be underestimated by about 14.73%. The effect of FAR change is comparably limited because density is much less important than the locational factors in determining house prices. Based on the estimated

coefficients, this quality change could only lead to an underestimation in accumulated price growth by about 0.68% during the period of 2004-2009.

4.2.3 Developers' Pricing Behaviors

In order to test the existence of developers' pricing behaviors discussed in Section 2.3, firstly we choose to introduce the units' selling sequences as the proxy, which is defined as the ratio between number of units sold before the unit and the total number of units in the complex (*PERCENT*). As listed in column 2 in Table 2, this variable is significantly negative in the hedonic model, which means that, controlling for other factors, units sold later are listed (and thus achieve in) a lower price than units sold earlier in the same complex. More precisely, its coefficient suggests that the transaction price of the last unit sold in a complex could be about 10.92% lower than the first unit sold in the same complex, controlling for other factors including observed housing attributes and market conditions. The result is robust if we introduce units' time-on-markets (*TOM*; in log term) as the proxy instead (column 3 in Table 2), or even introduce *PERCENT* and *TOM* to the model simultaneously (column 4 in Table 2). Although we cannot provide a definitive conclusion whether this phenomenon results from the reasons discussed in Section 2.3, the significantly negative correlation between units' prices and transaction sequences/duration is enough to prove the existence of bias in the matching approach when adopted in the newly-built sector.

4.3 *Comparison of Performance between Various Methods*

Since both the complex- and unit-level quality changes are proved to exist in the newly-built housing market, according to the earlier theoretical analyses, the simple average methods and the matching approach should both suffer from downward bias. In

order to test this we finally turn to the direct comparison between price indices calculated via various methods.

4.3.1 Method Specification

The monthly house price index between 2004 and 2009 in this sample city is calculated by the three methodologies respectively, with the specifications as follows.

(1) Simple Average Method

Both the unweighted and weighted (with unit size of each transaction as the weight) average transaction prices are calculated relative to a fixed base. The weighted average formula is exactly how the "Average Price Index" is calculated.

(2) Matching Approach

Following the strategy suggested by McMillen (2008), Deng, McMillen and Sing (2012), we replicate the "70 cities index" in a stricter way by allowing each unit to be matched with the most similar unit within the same complex sold during the previous months via the propensity score approach. More precisely, the following procedures are applied to each complex included in the dataset one by one. For each month after the initial transaction in this complex (*i.e.*, starting from month t+1 if the first transaction in this complex appears in month t), all housing units within this complex sold in and before that month are pooled together in a probit model (the dependent variable equals 1 for units sold in this month, and 0 for units sold in previous months; the explanatory variables are the unit-level attributes listed in Table 1). Then the propensity score is predicted for all these units based on the coefficients estimated, and each unit sold in this month is matched with its "nearest neighbor" in propensity score sold in previous

months.[20] 385,179 pairs of transactions are matched via these procedures. Finally the standard procedures of weighted repeat sales method developed by Case and Shiller (1987, 1989) are applied to these matched pairs to estimate the price index.

(3) Hedonic Method

The hedonic house price index is calculated based on coefficients of the time dummies listed in column 1 in Table 2.

4.3.2 Results and Analyses

The house price indexes calculated by these three methodologies are depicted in Figure 4, and Table 3 provides four indicators on the performance of the methods which are applied in most existing methodological comparison researches (Case, Pallakowski and Wachter, 1991; Case and Szymanoski, 1995; Dorsey *et al*, 2010; Nagaraja, Brown and Wachter, 2010): (1) the average monthly growth; (2) the standard deviation of the monthly growth; (3) the average of width of 95% confidence interval; and (4) the standard deviation of error term in the model. The first two indicators reflect the divergence in long-term trend or short-term variance of the price index series, although can only serve as qualitative analysis because there are no absolute standards for these two indicators. The latter two are more quantitative – methods achieving in narrower confidence interval or smaller variance in error term are always expected to be better.

<center>* Insert Figure 4 about here*</center>

<center>* Insert Table 3 about here*</center>

On the one hand, the magnitude of house price appreciation rates suggested by the simple average methods are lower than that resulting from the hedonic method as

---

[20] If more than one unit sold in previous periods shares the same distance in propensity score to the object unit, the unit sold last would be selected. See Deng, McMillen and Sing (2011) for more details about the matching procedures.

expected. The average annual growth rate by the unweighted and weighted average formulas are 0.25 and 0.19 percentage points (or 18.4% and 14.0%) lower than the hedonic method, respectively. Moreover, consistent with the analysis on Figure 3, the difference mainly exists during the period of rapid housing suburbanization in 2004-2007 (average annual price growth rate of 1.55% by the weighted average formula and 1.82% by the hedonic method), but comparably smaller (0.41% and 0.46%) during the period of 2008-2009 when the spatial pattern kept generally stable. This provides an evidence for the existence of the non-constant-quality bias in the simple average methods. It also suggests that the level of such bias depends on the magnitude of complex-level quality changes of the transacted units, and hence may vary with time and city.

On the other hand, the indices estimated by the matching approach greatly diverge from results by other methods in the magnitude of price growth – the average monthly growth rate by the matching approach is only about one third of that by the hedonic model (0.42% by the matching approach and 1.36% by the hedonic method), and is also far lower than the average indices. Such difference reflects the downward bias resulting from developers' pricing strategy.

4.3.3 Summary

The study shows hedonic method allows researchers to control both the complex-level quality changes and the effect of developers' pricing behaviors; hence in the newly-built housing markets in China, it offers better measurement for housing price movement. The other two candidates suffer from server downward bias by not appropriately controlling for drift in quality change (in the case of simple average methods), and not

27

appropriately controlling for developers' pricing strategy (in the case of matching approach).[21]

The results could also help us understand the potential problem of the two official house price indices currently published in China. Despite any possible flaw in micro-level data underlying the calculation, in the newly-built housing markets both the "Average Price Index" and the "70 Cities Index" could be expected to suffer from substantial bias due to their inappropriate index construction methods. This well explains the divergent paths of these two indices as depicted in Figure 1, especially the extremely low price growth rate of the "70 Cities Index".

5. Hedonic Price Indices for Major Chinese Newly-Built Housing Markets

The above theoretical and empirical analyses suggest the hedonic method as a preferable choice for newly-built house price measurement. Therefore we try to apply this methodology to the newly-built housing markets in 35 major Chinese cities, as a first step in a series of efforts to provide a high-quality house price indicator for market analyses, researches and policy designs.

As mentioned earlier, since 2006 the local housing authorities in all major cities have been collecting full-sample newly-built housing transaction data via REMIS. But currently the raw data are still not open to public due to privacy considerations. As an

---

[21] Because both the trended changes in unobserved unit-level characteristics and effects of developers' pricing strategy only exist in the newly-built housing markets, these results do not violate the conclusions in earlier researches (McMillen, 2008; Deng, McMillen and Sing, 2011) about the appropriateness of the matching approach in the resale sector. Actually our empirical tests based on resale housing transactions in the same sample city point out that, the unit-level matching approach could achieve in consistent results with the hedonic method, while it does not require any information on complex-level attributes. It also overbids the other methods judging from indicators of width of 95% confidence interval and standard deviation of error term. Therefore we believe the matching approach with the repeat sales modeling framework is still a preferable choice in resale house price measurement in nascent markets like China. The results are available on request.

initial attempt, we developed a program which could automatically read raw data from REMIS, construct and estimate the hedonic model,[22] and calculate the index based on the estimated coefficients. We sent this program to local housing authorities in the 35 major cities and finally succeeded in persuading them all to run this program, which enabled us to get the hedonic price indices in the 35 major newly-built housing markets from 2006 to 2010.[23] These city level series are then aggregated into a national level indicator using the weighted average formula, with the total transaction volume in these 5 years in each city as the weight.

This new hedonic index reflects a dramatic house price surge during the sample period, with an average appreciation rate substantially higher than the two existing indicators (Table 4). At the national level, the real quarterly price appreciation rate of the hedonic index reaches 3.94% (equaling an average real appreciation growth rate of 16.71%), which is significantly higher than that suggested by the "Average Price Index" (1.87% quarterly, or 7.69% annually) and the "70 Cities Index" (1.02% quarterly, or 4.14% annually). Similarly, as for the city level, the real quarterly price appreciation rates of the hedonic indices are much larger than those by the "70 Cities Index" in all the four first-tier cities, and are also higher than the "Average Price Index" in three of the four cities, with Shenzhen as the only exception.[24]

---

[22] We had to define a standard specification for hedonic models in all these cities (which was almost the same as the specification discussed in Section 4.2.1; the only difference is the variables of *TOPFLOOR* and *FIRSTFLOOR* were not included), since we were now allowed to choice a most suitable specification for each city respectively in that condition. Besides, it was also difficult for us to adopt more complicated estimation methods than OLS in that simple program. We leave these for future studies.

[23] The accumulated sample volume included in the index calculation in these 35 cities from 2006 to 2010 is about 8.3 million units.

[24] One possible explanation is, at the end of 2005 the ratio between urban area and total area in Shenzhen already reached 0.32 (calculated based on the statistics published by MHURD), the highest in all these 35 major cities. This implies there is comparably less space left for further urban expansion in Shenzhen. Therefore the trend of housing suburbanization is weaker in Shenzhen during the sample period.

Such difference in house price measurement results could lead to totally different judgment on market conditions – the current Chinese housing market may be even more overheated and risky than is revealed in existing researches and analyses based on the official price indicators. Given the fact that most key fundamental factors in China kept growing at real speeds around 10% annually during recent years, [25] a real annual appreciation rate of 7.69% or 4.14% for newly-built house price could be expected to be reasonable. However, a house price appreciation rate of 16.71% annually obviously could not be fully supported by the fundamental factors; instead some mispricing should exist. So far most bubble gauge researches (*e.g.*, Ahuja *et al*, 2010) believe there is no evidence of national-level house price bubble in China, and the Chinese government also holds similar opinions. Such judgments, as well as the intervention policies based on these judgments, need to be re-evaluated carefully when the potential bias in house price index is considered.

Figure 5 investigates such potential risks more detailedly. The horizontal axis provides the real annual growth rate of per capita disposable income in each city, while the vertical axis provides the real average annual appreciation rate of our hedonic house price indices. During the period between 2006 and 2010, the appreciation rate of the hedonic index in 27 of these 35 major cities exceeded the growth rate of household income, most of which concentrated in the east region. In cities like Beijing, the average appreciation rate of house price is over three times as much as the growth rate of income. This provides a much more risky picture compared with similar results including Yang

---

[25] According to the statistics published by the NBSC, between 2006 and 2010 the real annual growth rate for GDP, per capita GDP and per capita disposable income in urban area in national level was 10.8%, 10.5% and 9.5%, respectively.

and Shen (2008), Chen, Hao and Stephens (2010), Ahuja *et al* (2010) and Wu, Gyourko and Deng (2012), which all adopt the official house price index in their price-to-income analyses.

<center>* Insert Figure 5 about here*</center>

6. Conclusions

As issues related to Chinese house prices gradually become an international concern, the accurate measurement of Chinese house prices will also become an important issue, especially in the newly-built sector which still dominates China's nascent housing market. However, considering the remarkable distinctions between the newly-built and resale residential sectors, the existing house price index construction methods developed in the resale sector require careful re-evaluation before introduced to the newly-built sector.

In this paper, we apply three major house price indices methodologies to Chinese newly-built housing markets, and compare their performance based on the unique features in the newly-built sector. It is clear that the simple average method without quality adjustment is biased if transacted units' complex-level attributes are experiencing trended changes over time. Particularly, they are more likely to be downward biased in most Chinese cities as a result of the recent trend of housing suburbanization. The matching approach using the repeat sales modeling framework is also biased (downward biased according to the empirical results) when adopted in the newly-built house price measurement due to the unique pricing behavior of newly-built housing units. This makes

<center>31</center>

the hedonic method a better choice of house price index construction in the Chinese newly-built housing markets.

Based on this finding, we apply the hedonic method to the newly-built housing markets in 35 major cities, and result in a substantially higher house price appreciation rate than the two official indicators. In other words, the current Chinese housing markets are even more risky than what have been reported based on these two official indices.

Nevertheless, the hedonic method adopted in this paper is surely far from a perfect solution. Several important issues remain on the agenda for further researches. First, the specification of the hedonic model should be investigated more carefully. Inclusion of more housing attributes in the hedonic model could help further improve the accuracy of the index. Besides, currently the coefficients of housing attributes are set to be consistent during the whole sample period and for all complexes, which may be improper and thus result in bias. Introducing some less restrictive form of specification may help further improve the accuracy of the index. Similarly, the current log-log form of specification may be not a best choice for all cities. Second, besides those discussed in this paper, developers' marketing and pricing behaviors may still affect newly-built house price measurement via other channels, which is an interesting topic for future researches. Finally, the comparison in this paper mainly concentrates on the long-term trend of the index resulted. The methods' performances in other aspects, like the sensitivity to potential turning points, may also be very important and require further research.

References

Ahuja, A., Cheung, L., Han, G., Porter, N., & Zhang, W. (2010). Are house prices rising too fast in China? IMF Working Paper No. WP/10/274.

Anglin, P., Rutherford, R., & Springer, T. (2003). The trade-off between the selling price of residential properties and time-on-the-market: the impact of price setting. *Journal of Real Estate Finance and Economics*, 26(1), 95-111.

Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963). A regression method for real estate price index construction. *Journal of the American Statistical Association*, 58(304), 933-942.

Case, B., Pollakowski, H., & Wachter, S. (1991). On choosing among house price index methodologies. *AREUEA Journal*, 19(3), 286-307.

Case, B., & Quigley, J. (1991). The dynamics of real estate prices. *Review of Economics & Statistics*, 73(1), 50-58.

Case, B., & Szymanoski, E. (1995). Precision in house price indices: findings of a comparative study of house price index methods. *Journal of Housing Research*, 6(3), 483-496.

Case, K. E., & Shiller, R. (1987). Prices of single family homes since 1970: new indexes for four cities. *New England Economic Review*, 1, 45-56.

Case, K. E., & Shiller, R. (1989). The efficiency of the market for single-family homes. *American Economic Review*, 79(1), 125-137.

Chau, K. W., Wong, S. K., & Yiu, C. Y. (2005). Adjusting for non-linear age effects in the repeat sales index. *Journal of Real Estate Finance and Economics*, 31(2): 137-153.

Chen, J., Hao, Q., & Stephens, M. (2010). Assessing housing affordability in post-reform China: a case study of Shanghai. *Housing Studies*, 25(6), 877-901.

Chen, J., Guo, F., & Wu, Y. (2011). One decade of urban housing reform in China: urban housing price dynamics and the role of migration and urbanization, 1995–2005. *Habitat International*, 35(1), 1-8.

Clapman, E., Englund, P., Quigley, J., & Redfearn, C. (2006). Revisiting the past and settling the score: index revision for house price derivatives. *Real Estate Economics*, 34(2), 275-302.

Crowe, C., Dell'Ariccia, G., Igan, D., & Rabanal, P. (2011). Policies for macrofinancial stability: options to deal with real estate booms. IMF Staff Discussion Note No. SDN/11/02.

Deng, X., Huang, J., Rozelle, S., & Uchida, E. (2010). Economic growth and the expansion of urban land in China. *Urban Studies*, 47(4), 813-843.

Deng, Y., McMillen, D., & Sing, T. F. (2012). Private residential price indices in Singapore: a matching approach. *Regional Science and Urban Economics*, 42(3), 485-494.

Deng, Y., & Quigley, J. (2008). Index revision, house price risk, and the market for house price derivatives. *Journal of Real Estate Finance and Economics*, 37(3), 191-209.

Deng, Y., Morck, R., Wu, J., & Yeung, B. (2011). Monetary and fiscal stimuli, ownership structure, and China's housing market. NBER Working Paper No. 16871.

Dorsey, R., Hu, H., Mayer, W., & Wang, H. (2010). Hedonic versus repeat-sales housing price indexes for measuring the recent boom-bust cycle. *Journal of Housing Economics*, 19(2), 75-93.

Dreger, C., & Zhang, Y. (2010). Is there a bubble in the Chinese housing market? European University Viadrina Frankfurt (Oder), Department of Business Administration and Economics Discussion Paper No. 290.

Forgey, F., Rutherford, R., & Springer, T. (1996). Search and liquidity in single-family housing. *Real Estate Economics*, 24(3), 273-292.

Genesove, D. & Mayer, C. (2001). Loss aversion and seller behavior: evidence from the housing market. *Quarterly Journal of Economics*, 116(1), 1233-1260.

Gourieroux, C, & Laferrere, A. (2009). Managing hedonic housing price indexes: the French experience. *Journal of Housing Economics*, 18(3), 206-213.

Haurin, D. (1988). The duration of marketing time of residential housing. *Real Estate Economics*, 16(4), 396-410.

Haurin, D., & Hendershott, P. (1991). House price indexes: issues and results. *AREUEA Journal*, 19(3), 259-269.

Hill, R. C., Knight, J. R., & Sirmans, C. F. (1997). Estimating capital asset price indexes. *Review of Economics and Statistics*, 79(2), 226-233.

Himmelberg, C., Mayer, C., & Sinai, T. (2005). Assessing high house prices: bubbles, fundamentals and misperceptions. *Journal of Economic Perspectives*, 19(4), 67-92.

Hui, E., & Shen, Y. (2006). Housing price bubbles in Hong Kong, Beijing and Shanghai: a comparative study. *Journal of Real Estate Finance and Economics*, 33(4), 299-327.

IMF. (2011). Global economic prospects and policy challenges. http://www.imf.org/external/np/g20/pdf/021811.pdf.

Kain, J., & Quigley, J. (1970). Measuring the value of housing quality. *Journal of the American Statistical Association*, 65(330), 532-548.

Kiel, K., & Zabel, J. (1997). Evaluating the usefulness of the American housing survey for creating house price indices. *Journal of Real Estate Finance and Economics*, 14(1/2), 189-202.

Leung, T. C. & Tsang, K. P. (2011). Anchoring and loss aversion in the housing market: implications on price dynamics. HKIMR Working Paper #2011-28.

McMillen, D. (2008). Changes in the distribution of house prices over time: structural characteristics, neighborhood or coefficients? *Journal of Urban Economics*, 64(3), 573-589.

Nagaraja, C., Brown, L., & Wachter, S. (2010). House price index methodology. working paper.

Northcraft, G., & Neale, M. (1987). Experts, amateurs, and real estate: an anchoring-and-adjustment perspective on property pricing decisions. *Organizational Behavior and Human Decision Processes*, 39(1), 84-97.

Quigley, J. (1995). A simple hybrid model for estimating price indexes. *Journal of Housing Economics*, 4(1), 1-12

Thibodeau, T. (1989). Housing price indexes from the 1974-1983 SIMSA annual housing surveys. *AREUEA Journal*, 17(1), 100-117.

Wang, S. (2011). State misallocation and housing prices: theory and evidence from China. *American Economic Review*, 101(5), 2081-2107.

Wang, Y. (2001). Urban housing reform and finance in China: a case study of Beijing. *Urban Affairs Review*, 36(5), 620-645.

Wang, Y. (2003). Social and spatial implications of housing reform in China. *International Journal of Urban and Regional Research*, 24(2), 397-417.

Wu, F., Xu, J., & Yeh, A. (2007). *Urban Development in Post-Reform China: State, Market, and Space*. London: Routledge.

Wu, J., Gyourko, J., & Deng, Y. (2012). Evaluating conditions in major Chinese housing markets. *Regional Science and Urban Economics*, 42(3), 531-543.

Yang, Z., & Shen, Y. (2008). The affordability of owner occupied housing in Beijing. *Journal of Housing and the Built Environment*, 23(4), 317-335.

Figure 1: Real Quarterly Series of Two Official House Price Indices in Four Major Newly-Built Housing Markets (2005Q3-2010Q4; 2005Q3=100)

(A) Beijing

(B) Shanghai

(C) Guangzhou

(D) Shenzhen

Source: Calculated based on statistics published by local statistics authorities in corresponding cities.

Figure 2: Floor Area of Private Housing Units Completed and its Share in Total Annual Flow Supply, 1991-2010

Million SQ.M.

Source: Calculated based on statistics published by National Bureau of Statistics of China in "Statistics Yearbook of China" and "Real Estate Statistics Yearbook of China".

Figure 3: Complex-Level Quality Changes in the Sample City

Source: Authors' calculations.

Figure 4: Newly-Built House Price Indices in the Sample City by Various Methods

Figure 5: Comparison between House Price Appreciation and Income Growth (35 Major Cities, 2006-2010)

Table 1: Definition and Descriptive Statistics of Variables

Variable	Description	Mean	Std. Dev.
A. Complex-Level Attributes			
D_CENTER	Distance to the central point of the city (with a well-known square in the central region as the landmark); in kilometers.	5.82	2.45
D_SUBWAY	Distance to the nearest subway station; in kilometers.	2.32	1.81
SIZE	Land area of the complex; in ten thousand square meters.	4.16	3.78
FAR	Floor area ratio of the complex.	3.84	1.62
GREEN	Green space ratio of the complex.	0.34	0.10
PRESALE	For presale complexes, the interval between listing date and expected completion date of the complex; in months. For spot complexes, equaling 1.	20.08	5.42
B. Unit-Level Attributes			
UAREA	Room area of the unit; in square meters.	98.10	37.55
TFLOOR	Total floor of the building.	21.39	8.89
FLOOR	Ratio between floor level of the unit and total floor of the building.	0.54	0.27
FIRSTFLOOR	Whether the unit is on the ground floor; 1=yes, 0=o/w.	0.03	0.18
TOPFLOOR	Whether the unit is on the top floor; 1=yes, 0=o/w.	0.05	0.22
PERCENT	Ratio between the number of units sold before this unit and the total number of units in the complex.	0.44	0.27
TOM	Interval between the listing date and the transaction date; in months.	6.80	8.13

Table 2: Results of the Hedonic Model

Dependant: log(*PRICE*)

	(1)	(2)	(3)	(4)
log(*D_CENTER*)	-0.251 (-325.78)***	-0.254 (-332.11)***	-0.256 (-333.17)***	-0.256 (-333.72)***
log(*D_SUBWAY*)	-0.027 (-61.88)***	-0.027 (-62.65)***	-0.029 (-65.90)***	-0.028 (-64.26)***
log(*SIZE*)	0.026 (64.83)***	0.026 (65.30)***	0.024 (59.98)***	0.025 (62.90)***
log(*FAR*)	-0.010 (-11.46)***	-0.010 (-12.13)***	-0.011 (-12.78)***	-0.010 (-12.59)***
GREEN	0.151 (44.38)***	0.162 (47.87)***	0.149 (44.06)***	0.159 (47.01)***
log(*PRESALE*)	-0.036 (-40.34)***	-0.034 (-38.94)***	-0.032 (-36.50)***	-0.033 (-37.44)***
log(*UAREA*)	0.013 (13.55)***	0.022 (23.88)***	0.020 (21.84)***	0.024 (25.53)***
log(*TFLOOR*)	0.013 (14.42)***	0.010 (11.52)***	0.007 (7.33)***	0.008 (8.89)***
FLOOR	0.050 (39.20)***	0.054 (43.17)***	0.052 (41.55)***	0.055 (43.44)***
FIRSTFLOOR	0.059 (32.61)***	0.070 (39.05)***	0.064 (35.67)***	0.070 (39.15)***
TOPFLOOR	-0.017 (-11.25)***	-0.008 (-5.20)***	-0.013 (-8.59)***	-0.008 (-5.20)***
PERCENT	-	-0.116 (-101.39)***	-	-0.093 (-67.59)***
log(*TOM*)	-	-	-0.036 (-81.26)***	-0.016 (-30.49)***
Constant	8.210 (1199.61)***	8.232 (1214.11)***	8.244 (1209.91)***	8.243 (1215.11)***
Time Dummies	Yes	Yes	Yes	Yes
Observations	515,523	515,523	515,479	515,479
Adjusted R2	0.54	0.55	0.55	0.55

Notes: (1) t-statistics are reported in parentheses.

(2) ***: significant at the 1% level; **: significant at the 5% level;*: significant at the 10% level.

(3) See Table 1 for variable definitions.

Table 3: Summary Statistics of the House Price Construction Methods

	Average Monthly Increase	Std. Dev. of Monthly Increase	Width of 95% Confidential Interval	Std. Dev. of Error Term
Unweighted Average	1.11%	3.93%	0.023	0.256
Weighted Average	1.17%	4.31%	0.025	0.256
Matching Approach	0.42%	1.41%	0.016	0.125
Hedonic Method	1.36%	3.91%	0.020	0.216

Table 4: Comparison between the Hedonic Price Index and Two Official House Price Indices (2006Q1-2010Q4)

	"Average Price Index"	"70 Cities Index"	Authors' Calculation Based on Hedonic Method
(A) Average of Real Quarterly Increase			
Aggregated Index	1.87%	1.02%	3.94%
Beijing	4.60%	1.86%	6.94%
Shanghai	3.33%	0.26%	4.72%
Guangzhou	4.03%	0.58%	4.26%
Shenzhen	6.45%	0.08%	3.84%
(B) Standard Deviation of Real Quarterly Increase			
Aggregated Index	4.57%	4.14%	3.57%
Beijing	11.21%	2.17%	6.15%
Shanghai	12.51%	1.79%	4.86%
Guangzhou	10.40%	3.14%	7.34%
Shenzhen	24.20%	3.56%	9.76%

Note: The aggregated index for "Average Price Index" and authors' calculation based on hedonic method cover 35 major cities, while the "70 cities index" covers 70 cities. Thus the gap between "70 cities index" and the other two indicators in aggregated index may partly due to the difference in market coverage.