

The User Cost of Housing and the Price-Rent Ratio in Shanghai

Jie Chen¹, Yu Chen², Robert J. Hill³, and Pei Hu⁴

¹ Professor, School of International and Public Affairs, Shanghai Jiao Tong University, China (chenjie100@sjtu.edu.cn)

² Deceased, formerly at Department of Economics, University of Graz, Austria

³ Professor, Department of Economics, University of Graz, Austria (robert.hill@uni-graz.at)

⁴ Postdoc at Shanghai Pudong Development Bank and Fudan University (hululu 1990@126.com)

September 25, 2020

Abstract: We show that simple median price-rent ratios in Shanghai are distorted by quality differences between sold and rented properties. Correcting for these quality differences using hedonic methods reduces the price-rent ratio by 14%. Even so, the price-rent ratio in Shanghai (at about 67) is still extremely high by international standards. From a user cost perspective, such a large price-rent ratio is driven mainly by the very high rate of expected capital gains on housing. If households form their expectations by simply extrapolating past price trends, we find that the user cost of owner-occupying in Shanghai is negative (implying that everyone except short-term residents wants to owner occupy rather than rent). While for many years the user cost was probably negative, such a situation is not sustainable going forward. By international standards, house prices in Shanghai are already high, which limits the potential for further growth. Expected capital gains, therefore, need to start falling soon to prevent the emergence of a housing bubble. (JEL Codes: C43; E01; E31; R31)

Keywords: Shanghai housing market; Pricerent ratio; User cost; Hedonic quality adjustment; Capital gains

Acknowledgements:

We dedicate this paper to the memory of our coauthor and friend Yu (Larry) Chen.

Jie Chen acknowledges financial support from the NSFC-ESRC Joint Funding (NSF71661137004) and National Natural Science Foundation of China (NSFC71974125; NSF71573166).

Robert Hill acknowledges financial support from the Austrian Research Promotion Agency (FFG), grant #10991131.

Declarations of interest: none

1 Introduction

Price-rent ratios for housing in Chinese cities are much higher than elsewhere in the world (Tsai and Chiang, 2019). Here we focus specifically on Shanghai. We have two main objectives. First, we use hedonic methods and micro-level data to accurately measure the price-rent ratio in Shanghai. Second, in a user cost setting we consider why the price-rent ratio is so high.

The price-rent ratio is an important indicator of the state of a housing market. In particular, a high and rising price-rent ratio can signal the emergence of a housing bubble (see Shiller, 2010, and Bourassa, Hesli and Oikarinen, 2019). However, a simple ratio of median house price to median rent can be a misleading indicator since the median sold property is typically of better quality than the median rented property. Adjusting for this quality difference reduces the observed price-rent ratio. Even with this adjustment, we find that the price-rent ratio is still very high in Shanghai.

A useful benchmark against which to evaluate the observed price-rent ratio is its equilibrium counterpart obtained by equating the user cost of owner occupying with the cost of renting. A crucial component of the user cost formula is the expected capital gain, which is not directly observable. If households form their expectations by extrapolating from the past performance of the housing market, we find that the user cost in Shanghai is negative. In this case the equilibrium price-rent ratio is not defined. By implication, however high the actual price-rent ratio is in Shanghai, we cannot say it is above its equilibrium level.

Alternatively, the expected capital gain can be derived as a residual from the user cost equilibrium condition if it is assumed that households are indifferent between owning and renting. In this case expected capital gains are considerably lower than if households simply extrapolate from past performance, although still very high by international standards.

One important sign that households prefer to owner-occupy than rent is the high and rising vacancy rate in Shanghai (see Glaeser et al., 2017). Hence we argue that households are indeed forming their expectations by extrapolating from the past performance of the housing market. The resulting negative user cost implies a preference for owner-occupying.

Hence the question of whether Shanghai is experiencing a housing bubble hangs on whether or not such high expected capital gains are plausible. The Chinese economy has been growing rapidly for four decades, and there is clearly potential for further growth. For example, according to the International Comparisons Program (ICP) (see World Bank, 2020), China's per capita gross domestic product in US dollars in 2017 (converted at the purchasing power parity exchange rate) is \$14 150. This is only 23.6% of the level in the US. The problem is that some of this future growth of per capita income has already been factored into Shanghai's house prices, which are already high by international standards (see section 4.7). This is one reason why price-income ratios are so high in Shanghai (see Figure 1). Hence house prices can no longer be expected to rise faster than per capita income, and it remains to be seen whether expectations will adjust quickly enough to prevent a bubble emerging.

The remainder of this paper is structured as follows. Section 2 provides a short literature review of recent work on the Chinese housing market and on the application of the user cost equilibrium condition to housing markets. Section 3 focuses on our methodology for constructing quality-adjusted price-rent ratios, and explains the user cost equilibrium condition and how it can be used to construct the equilibrium price-rent ratio. Our empirical results for Shanghai are presented and interpreted in Section 4. Our main conclusion are then summarized in section 5.

2 Literature Review

A number of recent papers have debated whether the Chinese real estate market is in the midst of a bubble. The general consensus is that it is difficult to say. The very rapid rise in house prices and extremely high price-rent and price-income ratios are certainly indicative of a bubble. However, in an economy that has been growing as rapidly as China, much rests on whether and how long the Chinese growth miracle continues. Glaeser et al. (2017) argue that house prices can only stay so high if new construction of housing is sufficiently restricted. House prices have risen faster than income in most Chinese cities (as is shown for the case of Shanghai in Figure 1). Fang et al. (2016) and Chen and Wen (2017) argue

that one reason for this is the lack of alternative assets that households and firms can invest their wealth in. A second reason according to Fang et al. (2016) is that the low level of housing affordability does not look as bad if one assumes that incomes will continue rising fast. In other words, the expectation of future income growth is already factored into current house prices. However, they then note that:

regression to the mean is the single-most robust and empirically relevant fact about cross-country growth rates. (Fang et al., page 108)

China has been a large positive outlier for the last 40 years. They argue that sooner or later it must succumb to regression to the mean. When the Chinese growth rate starts to significantly slow, expectations will gradually adjust and then both the price-income and price-rent ratios should fall. This is because the high price-income is driven by expectations of rising incomes, while the high price-rent ratio is driven by expectations of capital gains. This downward adjustment in the price-income and price-rent ratio is likely to come more from falling prices than rising incomes and rents. Hence sometime in the future the transition to a slower growth rate in China may lead to a major correction in the housing market.

Traditionally, most of the analysis related to the Shanghai housing market is based on aggregate statistics (Guo and Qu, 2019; Zhou, 2016; Du and Zhang, 2015; Tsai and Chiang, 2019), and quality adjustment on the housing price data is largely ignored. Only recently there has been a growing number of papers applying hedonic analysis to micro-level data in Shanghai (Li, Wei, and Wu, 2019; Li et al., 2019; Lu, 2018; Zhang et al., 2018; Chen, Hao, and Yoon, 2018; Zhang and Chen, 2018). In addition to the high rate of expected capital gains, another factor that may be contributing to the high price-rent ratio in Shanghai is the unequal policy of school enrollment for children between renter families and homeowner families (Zhang and Chen, 2018).

The price-rent ratios we observe in Shanghai are extremely high by international standards (see for example Bourassa, Hoesli, and Oikarinen, 2019). This paper focuses specifically on the measurement issues that arise in the construction of these price-rent ratios, and the extent to which they can be explained by the user cost of housing. This user

cost approach to interpreting price-rent ratios dates back to Poterba (1984). Until recently, most applications have focused on measuring and interpreting changes in the price-rent ratio (e.g., Himmelberg, Mayer and Sinai, 2005). The construction of quality-adjusted price-rent ratios in levels is more complicated requiring the application of hedonic methods to micro-level data. Hill and Syed (2016) apply such an approach to Sydney data, and find that quality adjustment reduces the observed price-rent ratio in Sydney. We find the same thing is true for Shanghai.

Wu, Gyourko and Deng (2012, 2016) assess price-rent ratios in China specifically through the lens of a user cost approach. An important element of the user cost formula is the expected capital gain on housing, which is not directly observable. Wu, Gyourko and Deng compute the implied expected capital gain required for households to be indifferent between owner occupying and renting. We undertake a similar exercise here, except that we then argue that the Shanghai housing market is in a corner solution where everyone prefers to owner occupy. As far as we are aware, the implications of a sustained negative user cost (or at least the perception that it is negative) for a housing market have not previously been explored in the literature.

3 Theory and Analytic Framework

3.1 Imputation of price-rent ratios from hedonic models

A hedonic model regresses the price of a product on a vector of characteristics, whose prices are not independently observed (see Hill, 2013 and Silver, 2016). The hedonic equation is a reduced form that is determined by the interaction of supply and demand (see Rosen, 1974).

Here we estimate separate hedonic models for sold and rented apartments in Shanghai for each quarter using a semilog functional form.^{1,2} The hedonic model for apartments sold

¹This section draws on methods developed in Hill and Syed (2016).

²Alternative functional forms, such as linear or Box-Cox transformations, are sometimes also considered. See Malpezzi (2003) for a discussion of some of the advantages of semilog in a hedonic context.

in quarter t can be written as follows:

$$y_{Pt} = X_{Pt}\beta_{Pt} + u_{Pt}, \quad (1)$$

where y_{Pt} is an $H_{Pt} \times 1$ vector with elements $y_{Ph} = \ln p_{Ph}$ (where H_{Pt} denotes the number of apartments sold in quarter t), X_{Pt} is an $H_{Pt} \times C$ matrix of characteristics, β_{Pt} is a $C \times 1$ vector of characteristic shadow prices, and u_{Pt} is an $H_{Pt} \times 1$ vector of random errors. Examples of characteristics in our housing context include the number of bedrooms, number of bathrooms, land area, and district.

Similarly, the hedonic rent equation for each quarter t is written as follows:

$$y_{Rt} = X_{Rt}\beta_{Rt} + u_{Rt}, \quad (2)$$

where y_{Rt} is the vector of log rents of the apartments rented in quarter t , and X_{Rt} is the corresponding matrix of rented apartment characteristics.

Our objective is to compute a matched price-rent ratio for each individual apartment. A price for each rented apartment can be imputed from the hedonic price model, and a rent for each sold apartment imputed from the hedonic rent model. In this way a price-rent ratio can be calculated for each rented apartment and each sold apartment. A feature of this approach is that the hedonic price and rent models need to be defined on the same set of characteristics. In sections 3.2 and 3.3 we develop extensions of our basic method to account for missing characteristics (i.e., characteristics that are missing for only some apartments in our data set) and omitted variables (i.e., characteristics that are missing for all apartments in our data set).

A rent for each apartment h sold in quarter t is imputed from (2) as follows:

$$\hat{r}_{th}(x_{Pth}) = \exp \left(\sum_{c=1}^C \hat{\beta}_{Rtc} x_{Pthc} \right), \quad (3)$$

where $c = 1, \dots, C$ indexes the list of characteristics over which the price and rent hedonic models are defined, and $\hat{\beta}_{Rtc}$ denotes the estimated characteristic shadow rental price of characteristic c obtained from (2). Similarly, a price for each apartment j rented in quarter

t is imputed from (1) as follows:

$$\hat{p}_{tj}(x_{Rtj}) = \exp \left(\sum_{c=1}^C \hat{\beta}_{Ptc} x_{Rtjc} \right), \quad (4)$$

where $\hat{\beta}_{Ptc}$ denotes the estimated characteristic shadow price of characteristic c obtained from (1). We can also use the hedonic rent equation to impute a rent for an apartment j actually rented in quarter t :

$$\hat{r}_{tj}(x_{Rtj}) = \exp \left(\sum_{c=1}^C \hat{\beta}_{Rtc} x_{Rtjc} \right), \quad (5)$$

and the hedonic price equation to impute a price for an apartment h actually sold in quarter t :

$$\hat{p}_{tj}(x_{Pth}) = \exp \left(\sum_{c=1}^C \hat{\beta}_{Ptc} x_{Pthc} \right). \quad (6)$$

Strictly speaking, \hat{r} and \hat{p} are biased estimates of r and p since by exponentiating we are taking a nonlinear transformation of a random variable. Possible corrections have been proposed by Kennedy (1981) and others. From our experience, however, these corrections are small enough that they can be ignored. They are also partially offsetting when we divide \hat{p} by \hat{r} .

A price-rent ratio $P/R(sold)_{th}$ for an apartment h sold in quarter t is obtained by dividing the imputed price for apartment h obtained from (6) by its imputed rent obtained from (3):

$$P/R(sold)_{th} = \frac{\hat{p}_{th}(x_{Pth})}{\hat{r}_{th}(x_{Pth})} = \frac{\exp \left(\sum_{c=1}^C \hat{\beta}_{Ptc} x_{Pthc} \right)}{\exp \left(\sum_{c=1}^C \hat{\beta}_{Rtc} x_{Pthc} \right)}. \quad (7)$$

We can likewise generate a price-rent ratio $P/R(rented)_{tj}$ for each apartment j rented in quarter t . This is obtained by dividing the imputed price for apartment j obtained from (4) by its imputed rent obtained from (5):

$$P/R(rented)_{tj} = \frac{\hat{p}_{tj}(x_{Rtj})}{\hat{r}_{tj}(x_{Rtj})} = \frac{\exp \left(\sum_{c=1}^C \hat{\beta}_{Ptc} x_{Rtjc} \right)}{\exp \left(\sum_{c=1}^C \hat{\beta}_{Rtc} x_{Rtjc} \right)}. \quad (8)$$

This approach of imputing both the numerator and denominator is referred to in the hedonic literature as double imputation. An alternative would be to impute only the denominator in (7) and only the numerator on (8). This is what is referred to as single imputation. We prefer double imputation here since it is more robust to omitted variables (see Hill and Melsner, 2008).

Now let $Med[P/R(sold)]$ denote the median price-rent ratio derived from the price-rent distribution of sold apartments in (7), while $Med[P/R(rented)]$ denotes the corresponding median price-rent ratio defined on the apartments actually rented in (8). An overall median is obtained by averaging these two population specific medians as follows:

$$Med[P/R] = \sqrt{Med[P/R(sold)] \times Med[P/R(rented)]}. \quad (9)$$

3.2 Correcting for missing characteristics

A detailed description of the dataset is provided in section 4.2. However, at this point it is useful to flag one problem with the dataset. For many of the apartments one or more of the characteristics included in the hedonic model are missing. In the price data the characteristics that are sometimes missing are: floor level of apartment, total number of floors, building orientation, and building age. In the rent data only building orientation and building age are sometimes missing.

Instead of deleting price and rent observations for which some characteristics are missing, we estimate a number of different versions of our basic hedonic price and rent model. In the price data, floor level and total number of floors are either both present or both missing. This means that we have eight possible permutations of characteristics in the price data.

HM1: The complete hedonic model

HM2: Floor level and total number of floors are missing

HM3: Building orientation is missing

HM4: Building age is missing

HM5: Floor level, total number of floors, and building orientation are missing

HM6: Floor level, total number of floors, and age are missing

HM7: Building orientation and age are missing

HM8: Floor level, total number of floors, building orientation, and age are all missing

We estimate each of these models using all apartments that have at least that mix of characteristics available. For example, HM8 is estimated using the full dataset. HM1 by contrast can only be estimated using the subset of apartments for which none of the characteristics are missing. Once these eight models have been estimated, the price and rent for each apartment is imputed from the hedonic model that is tailored to its particular mix of available characteristics. For example, if the age of an apartment is missing, then its price is imputed from HM4.

For the rental data, only two characteristics (orientation and building age) are sometimes missing. So to impute rents for the rental dataset we only need to estimate HM1, HM3, HM4 and HM7. However, given that we also need to impute rents for apartments in the price dataset, it is still necessary to estimate all eight models (H1-H8) for the rent data.

3.3 Correcting for omitted variables

Omitted variables are a problem in all our hedonic models, even in HM1. The omitted variables may be physical (e.g., the quality of the structure, its energy efficiency, the general ambience, floor space, sunlight, the availability of parking, and the convenience of the floor plan), or locational (e.g., street noise, air quality and the availability of public transport links).³ Omitted variables bias may also result from nonequivalence between the bedroom and bathroom characteristics in the price and rent data sets. For example, a bathroom in a sold apartment may on average be of better quality than a bathroom in a rented apartment.

These two sources of omitted variables bias should reinforce each other since both the included and omitted characteristics are likely to be of better quality in the price data set than in the rent data set. This implies that our hedonic price-rent ratios, by failing to fully adjust for quality differences, will be biased upward.

Our first step in correcting for omitted variables bias is to obtain reference quality-

³The impact of some locational characteristics are captured by locational dummies.

adjusted price-rent ratios that are free of bias. This can be done by selecting apartments that are in both the price and rent data sets. We use a house price index and rent index to extrapolate forwards and backwards prices and rents on the same apartment in different quarters. In our dataset we only have two quarters of data. So the temporal extrapolation using price or rent indexes is only forward or back one quarter. For example, suppose apartment h sells in quarter s at the price p_{sh} and is rented in quarter t at the rate r_{th} . An address-matched (AM) price-rent ratio for this apartment in quarter s can be calculated by extrapolating the rental rate back to quarter s using a rental index $R_{s,t}$ as follows:

$$P/R_{sh}^{AM} = \frac{p_{sh} \times R_{s,t}}{r_{th}}, \quad (10)$$

or by extrapolating the selling price forward to quarter t using a price index $P_{s,t}$ as follows:⁴

$$P/R_{th}^{AM} = \frac{p_{sh} \times P_{s,t}}{r_{th}}. \quad (11)$$

The rent and price indexes $R_{s,t}$ and $P_{s,t}$ in (10) and (11) are calculated using the Törnqvist-type hedonic imputation formula in (14), which is itself derived from the Paasche and Laspeyres-type formulas in (12) and (13).

$$\text{Paasche - Type Imputation : } P_{s,t}^{PI} = \prod_{h=1}^{H_t} \left[\left(\frac{\hat{p}_{t,h}(x_{Pth})}{\hat{p}_{s,h}(x_{Pth})} \right)^{1/H_t} \right] \quad (12)$$

$$\text{Laspeyres - Type Imputation : } P_{s,t}^{LI} = \prod_{h=1}^{H_s} \left[\left(\frac{\hat{p}_{t,h}(x_{Psh})}{\hat{p}_{s,h}(x_{Psh})} \right)^{1/H_s} \right] \quad (13)$$

$$\text{Törnqvist Imputation : } P_{s,t}^{TI} = \sqrt{P_{s,t}^{PI} \times P_{s,t}^{LI}} \quad (14)$$

In (12), (13), and (14) we focus on the price data. Equivalent formulas exist for the rent data. The term x_{Pth} denotes the vector of characteristics of an apartment h sold in quarter t , and $\hat{p}_{t,h}(x_{Pth})$ refers to the imputed price in quarter t of this apartment. The Laspeyres-type formula takes all the apartments sold in the earlier quarter s , and imputes prices for

⁴For apartments with multiple prices and rents in our sample, we select the chronologically closest price and rent observations to construct our address-matched price-rent ratio. For apartments that sell and rent in the same quarter, we count these price-rent ratios twice. Hence we have exactly two address-matched price-rent ratios for each apartment that both sold and rented.

them in both quarters s and t . The Paasche-type formula takes all the apartments sold in the later quarter t , and imputes prices for them in both quarters s and t . Törnqvist, by taking the geometric mean of Laspeyres and Paasche, gives equal weight to apartments sold in both quarters s and t .

After pooling the address-matched properties, we then take the median for each quarter t :

$$\text{Med}(\text{AMS}_t) = \text{Med}_{h=1, \dots, H_t}[P/R_{th}^{AM}], \quad (15)$$

where $h = 1, \dots, H_t$ indexes all the address-matched price-rent ratios in quarter t in our data set. The notation AMS_t in (15) stands for “address-matched sample”. Since it is constructed from actual prices and rents on the same apartments, $\text{Med}(\text{AMS}_t)$ should by construction be free of omitted variables bias.

With our methodology in place for constructing quality-adjusted price-rent ratios that are free of omitted variables bias, we can now compute bias correction factors for models $\text{HM1}, \dots, \text{HM8}$. We consider first the bias of our HM8 model, since this is the only hedonic model that allows us to use the full sample of price-rent ratios. We calculate the bias as follows:

$$\lambda_{t, \text{HM8}} = \frac{\text{Med}[\text{HM8}(\text{AMS}_t)]}{\text{Med}(\text{AMS}_t)}, \quad (16)$$

where $\text{Med}[\text{HM8}(\text{AMS}_t)]$ denotes the median price-rent ratio obtained from (9) using the hedonic model HM8 applied to the address-matched sample (AMS) in quarter t . More precisely, we estimate the HM8 model over the full data set and then pick out the imputed price-rent ratios for apartments in the address-matched sample (AMS). The median is then calculated only over the imputed price-rent ratios in the address-matched sample. The median in the denominator of (16) [i.e., $\text{Med}(\text{AMS}_t)$] is obtained from (15). Both medians $\text{Med}[\text{HM8}(\text{AMS}_t)]$ and $\text{Med}(\text{AMS}_t)$ are therefore calculated over the same address-matched sample.

Any systematic deviation of $\lambda_{t, \text{HM8}}$ from 1 can hence be attributed to omitted variables bias in the $\text{HM8}(\text{AMS}_t)$ median price-rent ratio. In our empirical results we find in both quarters that $\lambda_{t, \text{HM8}} > 1$, indicating that omitted variables bias is causing the price-rent ratios obtained from the HM8 model to be systematically too high.

The omitted variables bias for each of our other models HMj (where $j = 1, \dots, 7$) is calculated as follows:

$$\lambda_{t, HMj} = \frac{Med[HMj(AMS_t \cap HMjS_t)]}{Med(AMS_t \cap HMjS_t)}, \quad (17)$$

where $AMS_t \cap HMjS_t$ means that the median price-rent ratio is calculated only over observations which are in the address matched price-rent dataset and have all the characteristics required to estimate HMj. For example, $AMS_t \cap HM8S_t = AMS_t$, since all apartments are in the HM8 sample. There are less observations in $AMS_t \cap HM4S_t$, since to qualify in this case an apartment has to have an actual price and rent observation and the age characteristic must be available. Still more restrictive is the $AMS_t \cap HM1S_t$ sample, which contains apartments that have an actual price and rent observation, and where floor level, total number of floors, orientation, and age are all available. Returning to (17), the numerator is the imputed median price-rent ratio for the $AMS_t \cap HMjS_t$ sample of apartments obtained from hedonic model HMj, while the denominator is the corresponding price-rent ratio from the actually observed prices and rents in this same sample.

Given our hypothesis that sold apartments perform better than rental apartments on the omitted variables, it follows that we expect that all our λ estimates will be greater than one. This is indeed what we find (see Table 6). In both quarters, the observed omitted variables bias corrections are very similar for each of our eight hedonic models. This indicates that when missing, floor level, total number of floors, building orientation, and age do not contribute much to the observed omitted variables bias in the price-rent ratios.

To correct for omitted variables bias we adjust the price-rent ratio of an apartment h sold in quarter t with the HMj mix of characteristics by dividing it by $\lambda_{t, HMj}$ as follows:

$$P/R(sold)_{th, HMj}^{adj} = \frac{P/R(sold)_{th, HMj}}{\lambda_{t, HMj}}. \quad (18)$$

Similarly, an apartment j with the HMj mix of characteristics rented in quarter t is adjusted

for omitted variables bias as follows:

$$P/R(\text{rented})_{th, HMj}^{adj} = \frac{P/R(\text{rented})_{tj, HMj}}{\lambda_{t, HMj}} \quad (19)$$

We refer to the price-rent ratios derived from (18) and (19) as *adjusted* price-rent ratios.

3.4 The user cost equilibrium condition

The user cost of a durable good is the present value of buying it, using it for one period and then selling it (see Hicks, 1946). In equilibrium this should equal the cost of renting the good for one period. Following Poterba (1984), Himmelberg et al. (2005), and Girouard, Kennedy, Noord and André (2006), the equilibrium condition can be written as follows:

$$R_t = u_t P_t, \quad (20)$$

where R_t is the period t rental price, P_t the purchase price, $u_t P_t$ is user cost, and u_t the per yuan user cost. In a housing context, per yuan user cost can be calculated as follows:

$$u_t = (r_t - \pi^e) + \omega_t + \delta_t + \gamma_t - g_t, \quad (21)$$

where r denotes an appropriate nominal interest rate, π^e the expected rate of inflation, ω is running and average transaction costs, δ the depreciation rate for housing, γ the risk premium of owning as opposed to renting, and g the expected real capital gain on housing. That is, an owner occupier foregoes interest/makes mortgage interest payments on the market value of the apartment, incurs running/transaction costs, depreciation, and risk (mainly due to the inherent uncertainty of future price and rent movements in the housing market) and benefits from any capital gains on the apartment.⁵ If $R_t > u_t P_t$, owner-occupying becomes more attractive and hence this should exert upward pressure on P and downward pressure on R until equilibrium is restored. The converse argument applies when $R_t < u_t P_t$.

⁵In some countries owner-occupiers can tax deduct mortgage interest payments (see Girouard et al. 2006 for a list of OECD countries providing such benefits). For these countries, r_t should be adjusted to include the offsetting tax benefit. However, no such benefit is provided to the owner occupiers in China.

Rearranging (20), we obtain that in equilibrium the price-rent ratio should equal the reciprocal of per dollar user cost (i.e., $P_t/R_t = 1/u_t$). The equilibrium price-rent ratio can then be compared with the actual prevailing price-rent ratio to detect departures from equilibrium.

The equilibrium condition (20), however, implicitly assumes that P_t and R_t are calculated for properties of equivalent quality. If instead P_t refers to a apartment that is of superior quality to the apartment referred to by R_t , then P_t/R_t in (20) is overestimated and, as a result, the user cost equilibrium condition will be biased towards finding that the price-rent ratio is above its equilibrium level. It is important therefore that the price-rent ratio in (20) is computed using hedonic methods.

4 An Application to Shanghai

4.1 Institutional background of the Shanghai housing market

Shanghai, located at the mouth of the Yangtze River on Chinas central eastern coast, is the largest city in China in terms of both population and GDP. With a resident population of more than 24 million and land space of 6 341 square km in 2019, Shanghai is one of only two cities in mainland China classified as Alpha + cities on the Globalization and World Cities (GaWC) ranking for 2020.⁶

In the 20th century, Shanghai was Chinas largest industry center. Since the beginning of the 21st century, Shanghai has rapidly transformed itself into a global financial and trade center, and its service sector now contributes nearly 70% of gross output. According to official statistics, in 2019 Shanghais GDP was 3 815.5 billion yuan and per capita GDP was 157.3 thousand yuan.⁷

Like other cities in mainland China, the private housing market in Shanghai was annulled during the 1950-80s under the central-planned economic system. While it revitalized somewhat in the early-1990s, it only started properly developing after the 1998 reform of the urban housing system which completely removed the welfare allocation of

⁶<https://www.lboro.ac.uk/gawc/world2020t.html>

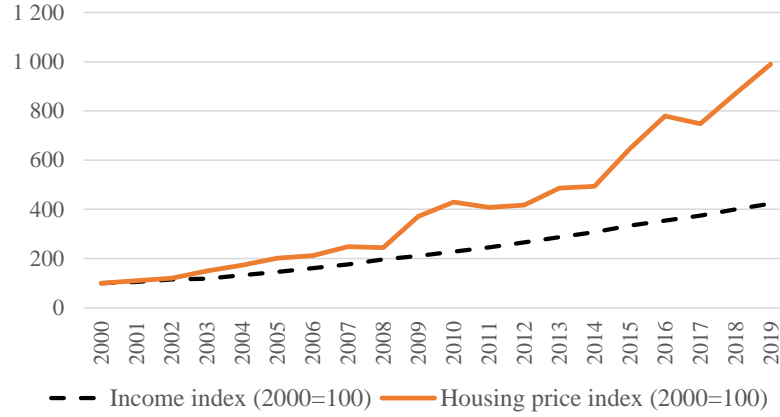
⁷The mean yuan-US dollar exchange rate was roughly 7:1 in 2019.

housing as a form of in-kind compensation to employees of public sector and stated-owned enterprises (Chen, Hao, and Stephens, 2010). Since the early 2000s, the housing markets in all major Chinese cities have witnessed continuous booms. Shanghai has had one of the strongest upward trends, and it has attracted a lot of international investment qualifying it as a superstar city (Glaeser et al., 2017). To curb the rapid price spiraling of housing that has triggered an affordability crisis and even threatened social stability, the Chinese central government has since the mid-2000s implemented several rounds of cooling policies and regulations (Zhou, 2016). In April 2010, on the order of the State Council, the home purchase restriction (HPR) policy was implemented in Shanghai alongside other major cities (Li, Cheng, and Cheong, 2017). Under the HPR policy, non-local residents (people without hukou) need to accumulate at least two years of local tax-paying record to buy homes in the city and local households cannot purchase and hold more than two properties. Since January 28, 2011, Shanghai became one of the first two cities (and also so far the only two cities) in China to levy a property tax; buyers of second properties in Shanghai are subject to a tax at the annual rate of 0.4-0.6% of the assessed value of the property, although 60 square meters per family member can be exempted from the tax (Du and Zhang, 2015). However, despite these regulations, house prices in Shanghai continued rising fast in the 2010s. In the first three quarters of 2016, the Shanghai housing market experienced another round of sharp price growth, approximately 40-50% year-to-year, and the municipality government responded by implementing several new curbing policies including two-round increases of the minimum down-payment ratio for home mortgages, forbidding some innovative mortgage products, and extending the minimum year requirement of tax-paying record history for non-local buyers to buy homes in the city. The housing market in Shanghai cooled in the fourth quarter of 2016 and slightly dropped in 2017 but rebounded sharply in 2018-2019 (see Figure 1). Currently, the Shanghai housing market is not only the most expensive in mainland China but also ranks very high even by any international standard.⁸

In 2019, even according to official statistics that arguably may underestimate the true price, the average price of newly-built residential housing in Shanghai is 32 926 yuan/sqm.

⁸<https://www.statista.com/statistics/1040698/most-expensive-property-markets-worldwide/>

Figure 1: The Income and Housing Price Trend in Shanghai (2000-2019)



Data source: Statistics Bureau of Shanghai Note: Income in the figure refers to the per capita disposable income of each year, Housing Price refers to the mean sale price of newly-built commodity residential housing per sqm.

In the central downtown area (the “inner ring” covering 120 square km) the average price is a stunning 111 307 yuan/sqm. The average price in the “mid-ring” covering 400 sq km is still very high at 75 796 yuan/sqm.⁹ By contrast, the mean annual disposable income per capita of urban individuals in Shanghai is only 73 615 yuan in 2019. Hence the average individual can only afford to buy at most 2.3 sqm of housing per year at its mean price even when all other consumption is ignored.

4.2 The data set and descriptive facts

The data sets used in this paper are from Hongwei Wang studio, a research hub of the Shanghai government development research center. The proxies of sale prices of dwellings used here are the listed prices from the real estate agencies’ websites. The data for prices consist of 409 277 observations from the fourth quarter of 2016 and the first quarter of 2017. The characteristics included in the data sets are the listed price, exact date of listing, floor area, number of bedrooms, number of living rooms, the total number of floors, floor level, location (block number), building orientation and building age. The proxies for rentals in

⁹<http://tjj.sh.gov.cn/tjfx/20200205/0014-1004435.html>

this paper are the listed rental prices. Exactly the same list of characteristics is available for rentals. The data for rentals consist of 113 828 observations. Since the data sets are expected to provide a comprehensive picture of the purchase and rental markets in Shanghai, on top of the difficulty of acquiring registered transaction prices and rentals, using listed data may be more relevant than using the actual transaction data in our analysis. In China, the registered transaction prices are normally lower than the actual market prices due to tax avoidance by transacting parties. Similarly, rentals are often not registered for tax avoidance reasons and due to the complex procedures of registration. In this respect, listed prices from the real estate agencies could be the closest available estimates of the actual transaction prices. The real estate agencies are an important force in the real estate market and the most sensitive and fast reacting players to new developments. Before the estimation of our hedonic models, we collated and cleaned the data. The same apartment may be listed for sale or rent on different agency websites. Therefore, we first delete the duplicate parts. Besides, before an apartment is sold or rented, the owner may keep listing it online and even revise the price according to market conditions. In this case, we only keep the latest prices. We also remove some extreme observations. Apartments with listed prices lower than 10000 yuan/m², or floor area bigger than 450 m², or the number of bedrooms greater than 10 are deleted. In addition, villas are removed from our datasets, since they are significantly different in price and area from the common dwellings in China. We also undertake some further deletions in order to implement our standard hedonic approach. In particular, if the hedonic price model includes apartments in a particular area, then the rental model must include apartments rented in the same area. In the end, 121 116 observations in the price data (see Table 1 for detailed descriptions of the price data set) and 112 199 observations in the rental data (see Table 2 for detailed descriptions of the rental data set) remain after deletion of outliers and other problematic observations. Basic descriptions of the price and rent data sets are provided in Tables 1 and 2 below.

A problem with the datasets is that there are many observations for which one or more characteristics are missing. In particular, all the characteristics are available for 43.7% of the price data and for 39.1% of the rental data (see Table 3). For the price data,

Table 1: Description for Price Data Set

Statistics	Obs	Mean	Std. Dev.	Min	Max
listed price (10 thousand yuan)	121 116	574.4093	520.537	27	16000
floor area (square meters)	121 116	97.27187	58.54025	20.01	450
No. of bedrooms	121 116	2.28211	0.997511	1	9
No. of living rooms	121 116	1.485493	0.687955	0	9
Floor level	117 010	1.9137	0.829212	1	3
Total number of floors	117 010	11.35869	7.869224	1	63
building orientation	102 714	6.345153	1.206106	1	10
location	121 116	103.3188	56.5942	1	201
exact date of listed	121 116	118.2055	45.43409	53	182
Building age	61 773	19.73801	9.568321	4	113

Table 2: Description for Rental Data Set

Statistics	Obs	Mean	Std. Dev.	Min	Max
listed rental (yuan)	112 199	5843.258	5632.265	543	280 000
floor area (square meters)	112 199	77.98334	39.42213	10	450
No. of bedrooms	112 199	1.93581	0.829358	1	9
No. of living rooms	112 199	1.286678	0.641236	0	9
Floor level	112 199	2.002977	0.843463	1	3
Total number of floors	112 199	11.60052	7.856862	1	35
building orientation	77 003	6.262379	1.136195	1	10
location	112 199	100.9971	56.60181	1	201
exact date of listed	112 199	85.82604	49.44529	1	168
Building age	60 186	20.43113	9.814125	4	113

four characteristics (floor level, total number of floors, building orientation, building age) are sometimes missing. While for the rental data, only building orientation and building age are sometimes missing.

In Section 3.2, we explain how we deal with this problem. It is plausible to assume that the missing data are randomly missing, in the sense that the probability of an observation being missing does not depend on the value of the observation. The original data sources are the real estate agencies. All the physical characteristics information are important for these agencies to demonstrate the quality and practicality of the apartment. Therefore, these data could be missing at the outset of the data entry process. Tables 1 and 2 show that the mean number of bedrooms and living rooms and mean floor area of sold dwellings are all higher than those of rental dwellings. Table 4 compares the bedrooms, living rooms, floor areas, and locational distributions of the price and rental data, from which the rented dwellings are clearly concentrated relatively more in the cheaper locations. From

Table 3: Percentage of Observations Having the Following Characteristics

Statistics	Price data	Rental data
floor level, total number of floors, building orientation, building age	43.70%	100%
floor level, total number of floors, building orientation	82.30%	100%
floor level, total number of floors, building age	50.10%	100%
total number of floors, building orientation, building age	43.70%	100%
floor level, building orientation, building age	43.70%	100%
floor level, total number of floors	95.82%	100%
floor level, building orientation	82.30%	100%
total number of floors, building orientation	82.30%	100%
floor level, building age	50.10%	100%
total number of floors, building age	50.10%	100%
building orientation, building age	44.40%	39.10%
floor level	95.82%	100%
total number of floors	95.82%	100%
building orientation	84.80%	68.60%
building age	51.00%	53.60%

the aforementioned results, it is not difficult to conclude that sold dwellings are of better quality than rented dwellings on average. This is also consistent with the actual situation in practice. It is because owner-occupiers tend to be richer and stay in the same dwelling for longer than renters, and hence place a higher value on quality. Also, owner-occupied dwellings are better maintained than rented dwellings.

Table 4: Distribution of Characteristics in the Price and Rental Data (in %)

characteristics	data	counts										total
		0	1	2	3	4	5	6	7	8	9	
bedroom	price	n.a.	20.60	44.35	25.20	6.88	2.25	0.55	0.09	0.05	0.02	100.00
	rental	n.a.	32.2	46.33	18.10	2.68	0.53	0.12	0.02	0.01	0.01	100.00
living room	price	7.33	40.25	49.29	2.85	0.26	0.02	0.01	0.00	0.00	0.00	100.00
	rental	9.66	52.62	37.17	0.52	0.03	0.00	0.00	n.a.	n.a.	0.00	100.00
characteristics	data	deciles										total
		1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th	
floor area ^a	price	7.53	8.31	9.27	9.03	9.53	9.65	10.44	10.28	11.72	14.24	100.00
	rental	13.96	11.15	11.29	9.94	10.52	11.24	9.10	9.25	8.15	5.42	100.00
location(by price) ^b	price	10.48	9.65	10.16	9.76	10.45	10.98	8.91	9.80	9.87	9.95	100.00
	rental	12.19	7.98	10.71	11.63	11.08	8.72	8.29	9.70	10.08	9.62	100.00

^aThe price and rental data are pooled before dividing them into deciles in terms of floor area. Therefore, each decile corresponds to the same floor area in both data sets.

^bHouses are ordered from the cheapest to the most expensive in terms of price.

^cHouses are ordered from the cheapest to the most expensive in terms of rent.

4.3 Econometric models

Our base hedonic model HM1 is semilog with $f(\cdot)$ denoting a linear function. We estimate this model separately for prices and rents for each quarter in our dataset.

$$\ln(\text{price}) = f(\text{floor area, floor area squared, no. bedrooms dummy, no. living rooms dummy, floor level dummy, no. floors dummy, orientation, block dummy, building age, building age squared, interaction between floor level and no. floors, interaction between floor area and no. bedrooms, interaction between floor area and no. living rooms, constant}). \quad (22)$$

Given there are two quarters in the dataset, this means we estimate two price and two rent versions of HM1. In addition we estimate the variants on our basic HM1 model denoted by HM2-HM8 as described in section 3.2. Estimating these additional models allows us to impute prices and rents for apartments where one or more of the characteristics is missing. Again two price and two rent versions of each of HM1-HM8 are estimated, one for each quarter. Hence in total we estimate 32 hedonic models. The number of observations and adjusted R-squared for each model are shown in Table 5.

Table 5: HM1-HM8 Regression Results

Stats	Price models				Rent models			
	2016Q4		2017Q1		2016Q4		2017Q1	
	No. obs.	Adj R^2	No. obs.	Adj R^2	No. obs.	Adj R^2	No. obs.	Adj R^2
HM1	25422	0.8968	27301	0.9225	37981	0.7142	5829	0.6288
HM2	25928	0.8954	27301	0.9216	37981	0.7098	5829	0.6241
HM3	30994	0.8832	29360	0.9147	52443	0.688	7655	0.6153
HM4	47902	0.8725	51148	0.9044	66847	0.6801	9970	0.5981
HM5	31713	0.8831	29360	0.9142	52443	0.6835	7655	0.6106
HM6	49907	0.8741	51148	0.9029	66847	0.6732	9970	0.5909
HM7	60185	0.8582	56093	0.8952	98070	0.6478	13877	0.5917
HM8	62994	0.8629	56093	0.8943	98070	0.6409	13877	0.5835

Note: The HM1-HM8 hedonic models are defined in section 3.2.

The full set of estimated shadow prices and t-statistics for the 32 estimated hedonic

models are provided in the Appendix.¹⁰ There are 16 districts in Shanghai and 201 blocks. Each district contains on average about 12-13 blocks. In the hedonic models we include block dummies instead of district dummies, since the blocks provide a more accurate measure of the impact of location on price and rent. In a few cases in some of the HM1-HM8 models we are not able to estimate block dummy (or some other) shadow price due to collinearity issues. When predicting prices and rents for properties in such blocks, we substitute one of the shadow prices for this block from one of the other hedonic models, or from a neighboring block.

The estimated omitted characteristic/variable adjustment factors are shown in Table 6. As is discussed in section 3, all these adjustment factors are greater than one, implying that sold properties perform better on the omitted variables than rented properties. The impact on the overall price-rent ratio for Shanghai of correcting for omitted characteristics and omitted variables is shown in Table 7.

Table 6: Omitted characteristic/variable adjustment factors: λ_{HMj}

	2016Q4	2017Q1
HM1	1.077	1.104
HM2	1.084	1.101
HM3	1.087	1.105
HM4	1.058	1.069
HM5	1.096	1.078
HM6	1.063	1.066
HM7	1.061	1.064
HM8	1.070	1.065

Note: The HM1-HM8 hedonic models are defined in section 3.2. The calculation of the adjustment factors, λ_{HMj} , is explained in section 3.3.

4.4 Price-rent ratios in Shanghai

Our estimated raw median and hedonic quality-adjusted price-rent ratios for Shanghai and its 17 main regions are shown in Table 8 and Figure 2. The first striking aspect of the results is how high the price-rent ratios are. The raw median price-rent ratio (calculated as price divided by annual rent) is 77 for Shanghai as a whole. By comparison with other cities around the world this is very high. For example, Bourassa, Hoesli, and Oikarinen

¹⁰The order of the variables in the Appendix is the same as in (22).

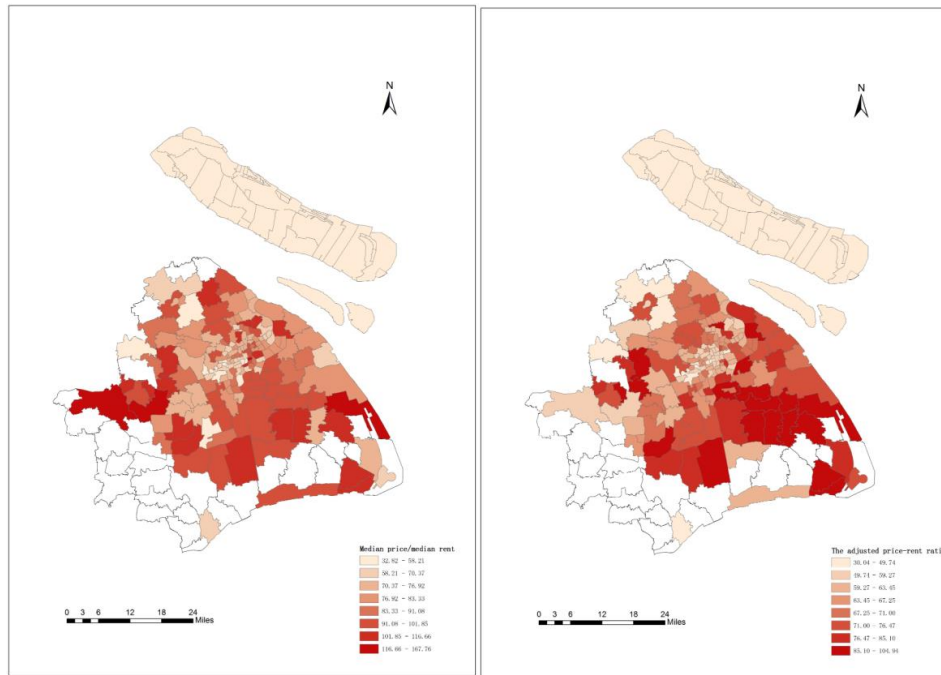
Table 7: Actual and quality-adjusted median price-rent ratios and quality bias

Quarter	2016Q4	2017Q1	Average
Actual unadjusted	75.93	77.78	76.85
Correcting for missing characteristics	70.77	74.45	72.61
Correcting also for omitted variables	66.13	68.35	67.2
Observable bias(%)	7.29	4.48	5.85
Total bias (%)	14.81	13.8	14.29

Note: The actual unadjusted ratios are calculated as the median price divided by the median rent in each quarter. The quality-adjusted price rent ratios are computed from (9). The correction for missing characteristics is explained in section 3.2. The additional correction for omitted variables is explained in section 3.3.

(2019) find that the price-rent ratios for Geneva, Helsinki, Zurich, Chicago, Miami and San Francisco range between 16 and 37 over the period 1980-2011. Even the highest of these (Geneva) is less than half the result we observe for Shanghai. Similarly, Pancak (2017) computes price-rent ratios for 48 US states and District of Columbia for the period 2010-11. She finds that the price-rent ratios range between 6.7 and 29.2.

Figure 2: Raw and quality adjusted median price-rent ratios



Note: The figure on the left shows subdistrict-level raw price-to-rent ratios (2017Q1), while the figure on the right show subdistrict-level quality adjusted price-to-rent ratios (2017Q1).

Table 8: Price-rent ratios for regions in Shanghai

2016Q4 Region	Median price (1000 Yuan)	Median annual rent (1000 Yuan)	Median price-rent ratio	Hedonic adjusted price-rent ratio
Shanghai (whole)	4100	54	75.93	66.13
Jinshan	1900	36	52.78	36.8
Chongming	2100	-	-	-
Fengxian	2560	26.4	96.97	82.47
Jiading	3300	42	78.57	62.36
Yangpu	3680	54	68.15	59.39
Baoshan	3800	45.6	83.33	70.91
Qingpu	4000	38.4	104.17	76.91
Songjiang	4000	50.4	79.37	65.06
Hongkou	4200	57.6	72.92	59.74
Pudong	4200	54	77.78	72.42
Putuo	4200	57.6	72.92	60.73
Zhabei	4250	54	78.7	60.57
Minhang	4300	54	79.63	70.73
Xuhui	5000	66	75.76	62.68
Changning	5000	70.8	70.62	53.53
Jingan	6000	105.6	56.82	52.62
Huangpu	6400	86.4	74.07	58.55
2017Q1 Region	Median price (1000 Yuan)	Median annual rent (1000 Yuan)	Median price-rent ratio	Hedonic adjusted price-rent ratio
Shanghai (whole)	4200	54	77.78	68.35
Jinshan	2120	33.6	63.1	33.92
Chongming	1850	39.6	46.72	35.39
Fengxian	2600	26.4	98.48	87.75
Jiading	3375	42	80.36	62.86
Yangpu	4300	54	79.63	63.44
Baoshan	3900	45.6	85.53	70.28
Qingpu	3500	39.6	88.38	75.81
Songjiang	3850	49.2	78.25	62.55
Hongkou	4600	54	85.19	64.96
Pudong	4455	51.6	86.34	74.66
Putuo	4750	60	79.17	63.28
Zhabei	4750	54	87.96	62.32
Minhang	4350	53.4	81.46	72.44
Xuhui	5000	60.6	82.51	67.82
Changning	5525	68.1	81.13	60.62
Jingan	7500	102	73.53	55.87
Huangpu	7985	84	95.06	64.19

Note: The actual unadjusted ratios are calculated as the median price divided by the median rent in each quarter. The quality-adjusted price rent ratios are computed from (9), with additional corrections for both missing characteristics and omitted variables as explained in sections 3.2 and 3.3.

Second, in all cases the hedonic price-rent ratio is less than its raw median counterpart. For Shanghai as a whole, the raw median is about 14 percent larger. For some regions is rather bigger. Notably in 2017Q1 the raw median for Jinshan (one of the cheapest regions) is 86 percent larger.

A raw median price-rent ratio that is higher than its hedonic counterpart implies that the average property being sold is of better quality than the average property being

rented. This is certainly what we find in Shanghai. Our results indicate that this quality difference is generally similar across more expensive and cheaper regions. Hence when computing price-rent ratios it is important to adjust for differences in the average quality of properties being sold and those being rented. Failure to quality adjust, at least in Shanghai, will result in upward biased estimates of the price-rent ratio.

On average there does not seem to be any systematic pattern between the price-rent ratio and how expensive a region is. For Shanghai as a whole the price-rent ratio is about 67. Across the regions it ranges between 34 and 88 over our sample period.

One further issue of interest is the cross-section distribution of price-rent ratios. Our hedonic approach generates a price-rent ratio for every property in each of the price and rent datasets. The cross-section distributions for the price dataset in 2016Q4 and 2017Q1 are shown in Figure 3, and the corresponding distributions for the rent dataset are shown in Figure 4. Again, there is no clear relationship between how expensive a property is and how high the price-rent ratio is.

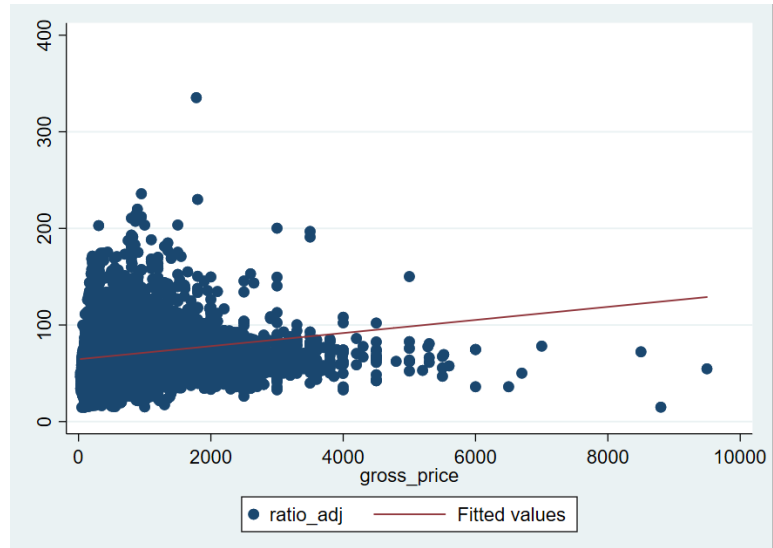
The most interesting result that emerges from Figures 3 and 4 is how much more variability there is in the price-rent ratios at the cheaper end of the market (for both the price and rent data). It is not clear how much of this variation is genuine and how much is attributable to the impact of omitted variables in our hedonic models. It is not immediately clear why omitted variables should be a bigger problem at the cheaper end of the market. This is an issue that warrants further investigation.

4.5 The user cost of owner occupying in Shanghai

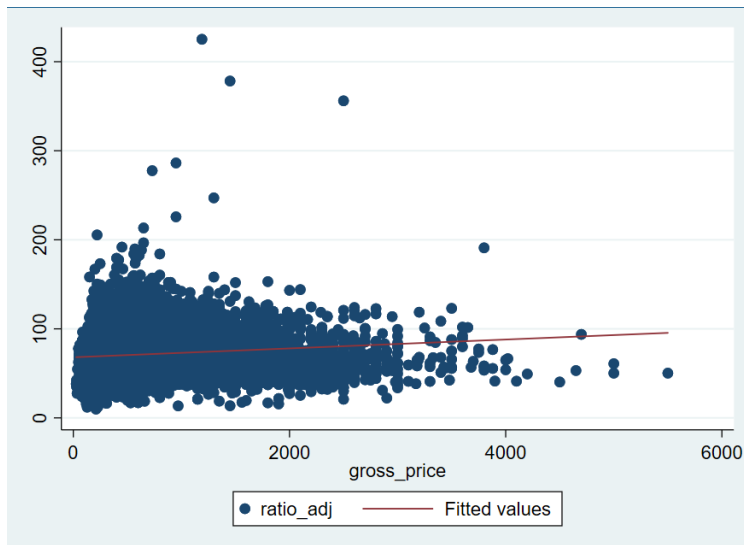
To understand why the price-rent ratio is so high in Shanghai it is informative to compare the costs incurred by owner occupiers and renters. The formula for computing the per yuan user cost, u , of owner occupying is given in (21). We discuss now how we compute each of the elements of u .

The nominal interest rate: r is constructed from two components. For the equity component (i.e., deposit) we use the 10-year government bond rate, which was 3.01% in 2016Q4 and 3.28% in 2017Q1. For the debt component (i.e., the mortgage) we use the benchmark interest rate on bank loans of more than 5 years, which was 4.9% in both

Figure 3: Quality adjusted price-rent ratios for apartments being sold



Results for 2016Q4



Results for 2017Q1

Figure 4: Quality adjusted price-rent ratios for apartments being rented

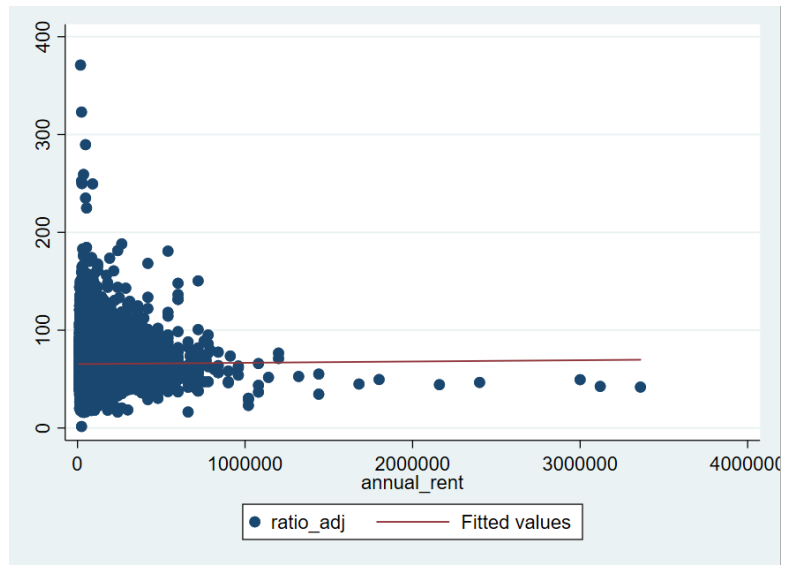
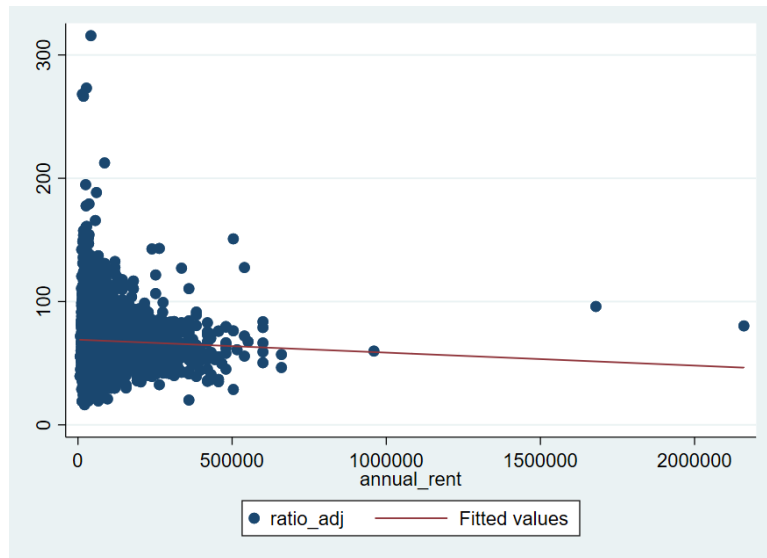


Fig. 3 adjusted price-rent ratio of apartments for rent in 2017Q1

1

Results for 2016Q4



Results for 2017Q1

2016Q4 and 2017Q1. We assume an initial loan-to-value ratio of 70% (see Figure 9 in Li, Wu and Xu, 2020).¹¹ Combining these terms, we obtain the following: $0.7 \times 0.031 + 0.3 \times 0.049 = 0.0436$, or 4.36%.¹²

The expected rate of inflation (π^e): We set this to 2.5%, which is about the average rate of inflation in China from 2007-2017 (see Day, 2017).

The transaction and running costs term (ω_t): Transaction costs consist primarily of real estate agent fees (2%), deed tax (2%) and capital gains tax (1%).¹³ Adding the components of transaction costs we obtain a value of 5%.¹⁴ Assuming buyers hold properties for 10 years before selling, this translates to an annual transaction cost of 0.5% per year.

Running costs consist of repairs, rates and insurance. We are unable to find relevant data to calculate these costs for China. So we assume running costs are 1% (similar to the value used for example by Fox and Tulip, 2014, for Australia). Summing annualized transaction costs and running costs we obtain an estimate of 1.5%.

Depreciation (δ): We use here Herd's (2020) estimate of 2% for depreciation of residential housing in China.

Risk premium (γ): We are not aware of any estimates of the risk premium of owner occupying as opposed to renting in China. Hence we use the standard estimate from the literature of 2% (see Himmelberg, Mayer and Sinai, 2005).

Expected capital gain (g): Gyourko provides a Shanghai house price index on his website.¹⁵ This is calculated using the method developed in Wu, Gyourko, and Deng (2012). The average annual increase from 2006-2017 was 22.81%. Given average inflation

¹¹Banks in China require deposits of at least 30% on mortgage loans (see Fang et al., 2016).

¹²Here we averaged the mortgage interest rates of 2016Q4 and 2017Q1.

¹³For the real estate agent fee we use Home-Link House Agent as an example, as it is the largest housing agency in Shanghai and generally charged 2% of the sale price during 2016-2017 (which was shared equally between the buyer and seller). The deed tax in China depends on the size of the property and how expensive it is relative to other properties in the same region. Common properties (with price less than 1.2 times the average for the region) with 90 or less square meters of living space pay a deed tax of 1%. Common properties with more than 90 square meters pay 1.5%. Other properties pay 3%. Here we will take an intermediate value of 2% for the deed tax. The capital gain tax can be calculated in two ways. If the original value of the dwelling cannot be provided, then the seller pays 1% of the sale price as the capital gains tax. If complete information can be provided, it is calculated at 20% of the tax payable, which is equal to the sale price minus the original value of the dwelling, tax and reasonable expenses incurred during the transaction. To simplify matters we use the first formula (i.e., 1%).

¹⁴By comparison, Cruz (2008) obtains a higher estimate for transaction costs of China for 10.77% for the total process of buying and selling.

¹⁵<http://real-faculty.wharton.upenn.edu/gyourko/chinese-residential-land-price-indexes/>.

of about 2.5% this implies a real capital gain of 20.3% per year. The expected capital gain is directly observed. We will assume here that households form their expectations by extrapolating past performance into the future (see Shiller, 2000, and Glaeser and Nathanson, 2017). In a housing context, Diewert (2009) argues for extrapolating from a past window of at least 10 years (see also Girouard et al., 2006, and Bracke, 2013). We assume that households extrapolate based on average performance since 2006 and hence expect real capital gains of 20.3% per year in 2016Q4 and 2017Q1.

Combining all these terms in (21) we obtain a per yuan user cost (u) of -0.129 (or -12.9%).¹⁶ While there can be some discussion over the particular values chosen for each element of the user cost formula, the calculations strongly suggest that households perceive the user cost to be negative. The implications of this are discussed in section 4.7.

4.6 The break-even expected capital gain

The finding of a negative user cost above, however, is not the only way of interpreting the data. Given expected capital gains are not directly observed, an alternative approach is to assume that households are indifferent between owner occupying and renting and then derive the implied expected capital gain (see Wu, Gyourko and Deng, 2012 and 2016, and Hill and Syed, 2016). Rearranging the user cost formula in (21) and imposing the equilibrium condition in (20) yields the following:

$$g_t = r_t - \pi^e + \omega_t + \delta_t + \gamma_t - \frac{\hat{R}_t}{\hat{P}_t}, \quad (23)$$

where \hat{R}_t/\hat{P}_t is the reciprocal of the median quality-adjusted price-rent ratio in period t obtained from our hedonic model in (9). Inserting in addition estimates of r_t , π^e , ω_t , δ_t and γ_t we obtain an estimate of g_t .

Applying this approach to the price-rent ratios in Table 8 we obtain the results shown in Table 9. The break-even expected real capital gain ranges between 4.4% and 6.2%. While considerably lower than the expected capital gains obtained by extrapolating from the past performance of the Shanghai housing market, by international standards the

¹⁶ $u = 0.436 - 0.025 + 0.020 + 0.015 + 0.020 - 0.203 = -0.129$.

estimates in Table 9 are still very high.¹⁷ By comparison, Gyourko, Mayer and Sinai (2013) compute the average annual real capital gain for 50 US cities over the period 1950 to 2000. They find that the average is 1.7%, with the highest result of 3.5% being observed for San Francisco.

Table 9: Break-even expected annual real capital gains

Region	2016Q4	2017Q1	Region	2016Q4	2017Q1
Shanghai (whole)	5.85%	5.90%	Hongkou	5.69%	5.82%
Jinshan	4.64%	4.41%	Pudong	5.98%	6.02%
Chongming	–	4.53%	Putuo	5.71%	5.78%
Fengxian	6.15%	6.22%	Zhabei	5.71%	5.76%
Jiading	5.76%	5.77%	Minhang	5.95%	5.98%
Yangpu	5.68%	5.78%	Xuhui	5.76%	5.89%
Baoshan	5.95%	5.94%	Changning	5.49%	5.71%
Qingpu	6.06%	6.04%	Jingan	5.46%	5.57%
Songjiang	5.82%	5.76%	Huangpu	5.65%	5.80%

4.7 Insights for the housing market in Shanghai

Compared with other cities around the world outside China, the price-rent ratio is very high in Shanghai. Does this imply that Shanghai is in the midst of a housing bubble? The answer is not necessarily.

A user cost perspective helps shed light on the situation in Shanghai. Suppose first that the user cost equilibrium condition holds in Shanghai. The implied expected annual real capital gain ranges between 4.4% and 6.2% across regions in Shanghai. While this is unusually high in comparisons with US cities (Gyourko, Mayer and Sinai, 2013) it is much lower than the average performance of the Shanghai housing market in recent years. When looked at this way, the observed price-rent ratios in Shanghai do not look so unreasonable.

Our results in section 4.5 strongly suggest that the user cost perceived by households is in fact negative in Shanghai. A negative user cost of owner occupying implies that everyone wants to owner occupy and no-one wants to rent. Under this scenario some households

¹⁷It should also be noted that many households that want to buy are unable to since they have not yet satisfied the residency requirement and/or are credit constrained. As a result, the estimates in Table 9 probably underestimate the required expected capital gain needed make households indifferent between owner occupying and renting.

will still rent for one of three reasons. First, some are sufficiently credit constrained that they are unable to borrow enough funds to buy (particularly given that a deposit of at least 30% is required). Second, as was noted in section 4.1, non-local residents need to accumulate at least two years of local tax-paying records (so that they qualify for the hukou) before they are eligible to buy. Third, short-term residents such as students and foreign workers may not wish to incur the transaction costs and time involved in buying and then soon after selling.

According to the user cost equilibrium condition in (20), a negative user cost (i.e., $u < 0$) implies that the equilibrium expected capital gain is not defined. This leads to the perhaps surprising conclusion that while the price-rent ratio is extremely high in Shanghai (by international standards) we cannot say that it is above its equilibrium level. The price-rent ratio would be even higher if there were not credit constrained households, the hukou residency rules, and short-term residents. In short, expected capital gains are so high that the owner-occupying versus renting equilibrium is a corner solution where everyone that can owner-occupies (except short-term residents).

Furthermore, with such high expected capital gains, investors will still want to buy even if the property is left vacant. The vacancy rate has indeed risen dramatically in first tier Chinese cities (i.e., Beijing, Shanghai, Guangzhou, and Shenzhen) in recent years (see Glaeser et al., 2017). In any OECD country a scenario like this would be interpreted as clear evidence of an emerging bubble. Stiglitz's (1990) definition of a bubble is relevant here.

“[I]f the reason the price is high today is *only* because investors believe the selling price will be high tomorrow – when ‘fundamental’ factors do not seem to justify such a price – then a bubble exists.” (Stiglitz, 1990, p. 13)

At first glance, the Stiglitz quote seems to apply directly to Shanghai. However, as has already been noted, China is currently experiencing an unprecedented economic transformation. During such a transformation, real capital gains on housing of 20% a year may not be unreasonable for a while. Furthermore, within China, Shanghai is a superstar city (in the sense of Gyourko, Mayer and Sinai, 2013). Hence even if a property is vacant now,

the owner may still expect it to yield high rents in the future. To the extent this belief is plausible then there may not be a bubble. Other factors also are relevant here. As was noted in section 2, the lack of alternative investable assets in China is acting to push more wealth into real estate than would otherwise be the case (see Fang et al., 2016, and Chen and Wen, 2017).

Projections of future rents also depend on expectations of new residential construction. The Chinese government has some control over how quickly the housing supply rises. More residential construction increases GDP but exerts downward pressure on house prices. Conversely, reducing residential construction could cause a recession, which might indirectly also adversely affect house prices. Hence the government will need to tread carefully in this regard. A bursting housing bubble would be economically destabilizing and could in turn trigger a financial crisis.

Real house prices cannot keep rising in Shanghai by 20% per year indefinitely. One problem is that current house prices in Shanghai already factor in continued rapid growth in per capita income (this is partly why the price-income ration in Figure 1 is so high). Per capita income has room to continue growing, since China's per capita income in 2017 was only 23.6% of US per capita income according to the World Bank (2020). However, by international standards, house prices are already high in Shanghai. While there is no agreed on ranking of house prices across cities, by some measures Shanghai is already one of the most expensive cities in the world.¹⁸ Hence house prices in Shanghai in the future cannot be expected to continue growing as fast as they have in the past or as fast as per capita income. To put it another way, the user cost of housing in Shanghai was almost certainly negative for many years. However, it is not clear that this should still be the case going forward if households have realistic expectations regarding future capital gains. If expected capital gains do not adjust quickly enough, a bubble could easily emerge. In this context, the quote from Stiglitz is directly applicable. Furthermore, once the expected capital gain falls enough so that the user cost is no longer negative, further falls in expected

¹⁸See for example the following websites: CBRE (<https://www.cbre.com/singapore/about/media-centre/singapore-remains-the-2nd-most-expensive-housing-market-in-the-world-after-hong-kong>), Statista (<https://www.statista.com/statistics/1040698/most-expensive-property-markets-worldwide/>) and the Global Property Guide (<https://www.globalpropertyguide.com/most-expensive-cities>).

capital gains could cause large reductions in the equilibrium price-rent ratio. Under such a scenario, the burden of most of the adjustment will likely fall on prices rather than rents. Hence the eventual transition from the current situation in Shanghai to that of other superstar cities outside of China (e.g., San Francisco or London) will be difficult to smoothly navigate.

5 Conclusion

We have considered here the measurement and interpretation of price-rent ratios in Shanghai. Focusing first on measurement, simple median price-rent ratios are distorted by quality differences between the median properties being sold and rented. These distortions can be corrected using hedonic methods. For Shanghai we find that the simple median price-rent ratio is 14% bigger than its quality adjusted counterpart.

Even with this correction, the price-rent ratio in Shanghai (at about 67) is extremely high by international standards. However, international standards may be misleading when evaluating developments in Chinese cities, especially Shanghai. For four decades China has been growing so rapidly that it was not unreasonable to expect high rates of capital gains on housing. So much so that the user cost of owner occupying was probably negative for many years.

However, looking forward, house prices in Shanghai have already factored in rapid increases in future per capita income. Given also that house prices in Shanghai are already high by international standards, there is limited scope for further increases in house prices. It is important, therefore, that households adjust their expectations to this new reality. Otherwise Shanghai is at risk of a housing bubble.

References

- Bourassa, S. C., M. Hoesli, and E. Oikarinen (2019), “Measuring House Price Bubbles,” *Real Estate Economics* 47(2), 534-563.
- Bracke, P. (2013), “How Long Do Housing Cycles Last? A Duration Analysis for 19

- OECD Countries,” *Journal of Housing Economics* 22, 213-230.
- Chen, J., Q. Hao, and M. Stephens (2010), Assessing Housing Affordability in Post-Reform China: A Case Study of Shanghai, *Housing Studies* 25(6), 877-901.
- Chen, J., Q. Hao, and C. Yoon (2018), Measuring the Welfare Cost of Air Pollution in Shanghai: Evidence from the Housing Market, *Journal of Environmental Planning and Management* 61(10), 1744-57.
- Chen, K. and Y. Wen (2017), “The Great Housing Boom of China,” *American Economic Journal: Macroeconomics* 9(2), 73-114.
- Cruz, P. C. R. (2008), “Transaction Costs and Housing Affordability in Asia,” *International Real Estate Review* 11(1), 128-150.
- Day, I. (2017), “Underlying Consumer Price Inflation in China,” RBA Bulletin, December, Reserve Bank of Australia, 29-36.
- Diewert, W. E. (2009), “Durable and Owner-Occupied Housing in a Consumer Price Index,” in Diewert W. E., J. S. Greenless and C. R. Hulten (eds.), *Price Index Concepts and Measurements*. NBER: University of Chicago Press, Chapter 12, 445-500.
- Du, Z., and L. Zhang (2015), Home-Purchase Restriction, Property Tax and Housing Price in China: A Counterfactual Analysis, *Journal of Econometrics* 188(2), 558-68.
- Fang H., Q. Gu, W. Xiong and L.-A. Zhou (2016), “Demystifying the Chinese Housing Boom,” *NBER Macroeconomics Annual*, University of Chicago Press, vol. 30(1), pages 105-166.
- Fox, R. and P. Tulip (2014), “Is Housing Overvalued?,” Research Discussion Paper RDP 2014-06, Reserve Bank of Australia, Sydney.
- Girouard, N., M. Kennedy, P. van den Noord, and C. André (2006), “Recent House Price Developments: The Role of Fundamentals,” OECD Economics Department Working Papers 475. Paris: OECD.
- Glaeser, E. L., W. Huang, Y. Ma, and A. Shleifer (2017), “A Real Estate Boom with Chinese Characteristics,” *Journal of Economic Perspectives* 31(1), 93-116.
- Glaeser, E. L. and C. G. Nathanson (2017), “An extrapolative model of house price dynamics,” *Journal of Financial Economics* 126(1), 147-170.

- Guo, J., and X. Qu (2019), Spatial Interactive Effects on Housing Prices in Shanghai and Beijing, *Regional Science and Urban Economics* 76(1), 147-60.
- Gyourko, J., C. Mayer and T. Sinai (2013), "Superstar Cities," *American Economic Journal: Economic Policy* 5(4), 167-99.
- Herd, R. (2020), "Estimating Capital Formation and Capital Stock by Economic Sector in China: The Implications for Productivity Growth," Policy Research Working Paper 9317, World Bank.
- Hicks, J. R. (1946), *Value and Capital*, Second Edition. Oxford: Clarendon Press.
- Hill, R. J. (2013), "Hedonic Price Indexes for Residential Housing: A Survey, Evaluation and Taxonomy," *Journal of Economic Surveys* 27(5), 879-914.
- Hill, R. J. and D. Melser (2008), "Hedonic Imputation and the Price Index Problem: An Application to Housing," *Economic Inquiry* 46(4), 593-609.
- Hill, R. J. and I. Syed (2016), "Hedonic Price-Rent Ratios, User Cost, and Departures from Equilibrium in the Housing Market," *Regional Science and Urban Economics* 56, 60-72.
- Himmelberg, C., C. Mayer and T. Sinai (2005), "Assessing High House Prices: Bubbles, Fundamentals and Misperceptions," *Journal of Economic Perspectives* 19(4), 67-92.
- Kennedy, P. E. (1981), "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations." *American Economic Review* 71 (4), 801.
- Li, H., Y. D. Wei, and Y. Wu. (2019a), Analyzing the Private Rental Housing Market in Shanghai with Open Data, *Land Use Policy* 85 (June), 271-84.
- Li, H., Y. D. Wei, Y. Wu, and G. Tian (2019b), Analyzing Housing Prices in Shanghai with Open Data: Amenity, Accessibility and Urban Structure. *Cities* 91(1), 165-79.
- Li, H., J. Wu and M. Xu (2020), "Estimating the Leverage Condition of Chinas Urban Households: Evidence from the Housing Sector," Paper available at papers.ssrn.com.
- Li, V. J., A. W. W. Cheng, and T. S. Cheong (2017), Home Purchase Restriction and Housing Price: A Distribution Dynamics Analysis, *Regional Science and Urban Economics* 67 (July), 110.

- Lu, J. (2018), The Value of a South-Facing Orientation: A Hedonic Pricing Analysis of the Shanghai Housing Market, *Habitat International* 81 (November), 24-32.
- Malpezzi, S. (2003), "Hedonic Pricing Models: A Selective and Applied Review," in A. OSullivan and K. Gibb (eds.) *Housing Economics: Essays in Honor of Duncan MacLennan*, 67-89. Blackwell: Malder MA.
- Pancak, K. A. (2017), "Variation in Local House Price-Rent Ratios," *Journal of Housing Research* 26(1), 79-94.
- Poterba, J. (1984), "Tax subsidies to owner-occupied housing: An asset market approach," *Quarterly Journal of Economics* 99(4), 729-752.
- Rosen, S. (1974), "Hedonic prices and implicit markets: product differentiation in pure competition," *Journal of Political Economy* 82(1), 34-55.
- Shiller, R. J. (2000), *Irrational Exuberance*. Princeton NJ: Princeton University Press.
- Measurement, *Journal of Economic and Social Measurement*, 39(4),305-328.
- Silver, M. (2016), "How to Better Measure Residential Property Price Indexes," IMF Working Paper Series WP/16/213.
- Stiglitz, J. E. (1990), "Symposium on Bubbles," *Journal of Economic Perspectives* 4(2), 13-18.
- Tsai, I.-C., and S.-H. Chiang (2019), Exuberance and Spillovers in Housing Markets: Evidence from First- and Second-Tier Cities in China, *Regional Science and Urban Economics* 77, 75-86.
- World Bank (2020), *Purchasing Power Parities and the Size of World Economies*. Washington DC: World Bank Group. <https://openknowledge.worldbank.org/bitstream/handle/10986/33623/9781464815300.pdf>
- Wu, J., J. Gyourko, and Y. Deng (2012), "Evaluating Conditions in Major Chinese Housing Markets," *Regional Science and Urban Economics* 42, 531-543.
- Wu, J., J. Gyourko, and Y. Deng (2016), "Evaluating the risk of Chinese housing markets: What we need to know," *China Economic Review* 39, 91-114.
- Zhang, L., J. Chen, Q. Hao, and C. Z. Li (2018), Measuring the NIMBY Effect in Urban China: The Case of Waste Transfer Stations in Metropolis Shanghai, *Journal of*

Housing and the Built Environment 33(1), 1-18.

Zhang, M., and J. Chen (2018), Unequal School Enrollment Rights, Rent Yields Gap, and Increased Inequality: The Case of Shanghai, *China Economic Review* 49(3), 229-40.

Zhou, Z. (2016), Overreaction to Policy Changes in the Housing Market: Evidence from Shanghai, *Regional Science and Urban Economics* 58(1), 26-41.

Online Appendix

The Appendix is submitted separately as an additional file. It consists of estimated shadow prices and t-statistics of the 32 hedonic models from section 4.3. The Appendix runs to 155 pages.