

MIT Open Access Articles

House Age, Price and Rent: Implications from Land-Structure Decomposition

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Xu, Yangfei, et al. "House Age, Price and Rent: Implications from Land-Structure Decomposition." *The Journal of Real Estate Finance and Economics*, vol. 56, no. 2, Feb. 2018, pp. 303–24.

As Published: <http://dx.doi.org/10.1007/s11146-016-9596-6>

Publisher: Springer US

Persistent URL: <http://hdl.handle.net/1721.1/115554>

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Terms of use: Creative Commons Attribution-Noncommercial-Share Alike



House Age, Price and Rent: Implications from Land-Structure Decomposition

Yangfei Xu

Hang Lung Center for Real Estate and Department of Construction Management, Tsinghua
University, xuyf12@mails.tsinghua.edu.cn

Qinghua Zhang

Guanghua School of Management, Peking University, zhangq@gsm.pku.edu.cn.

Siqi Zheng (corresponding author)

Department of Urban Studies and Planning, Massachusetts Institute of Technology; and Hang
Lung Center for Real Estate and Department of Construction Management, Tsinghua University,
zhengsiqi@gmail.com. 86-18611045757

Guozhong Zhu

Alberta School of Business, University of Alberta, guozhong@ualberta.ca. 86-10-62767407.

Abstract

Big cities often witness land price outgrowing structure price. For such cities this paper derives two predictions regarding the dynamics between house prices, rent and structure age. First, older houses have a higher price growth rate than younger ones, even after controlling for location and other attributes; second, the age depreciation of house price, defined as the decline of house price with respect to house age, is slower than the similarly-defined age depreciation of rent. These hypotheses are supported by the micro-data on housing market in Beijing. These two inferences have implications for both real estate valuation and house price index construction.

Key words

Land price, structure price, house prices, rent, depreciation

Acknowledgments

Guozhong Zhu and Qinghua zhang acknowledge supports from National Science Foundation (#71473004 and #71273017). Yangfei Xu and Siqi Zheng acknowledge supports from National Natural Science Foundation of China (#71625004, #71273154, #71322307, #71533004), the national key research & development (R&D) plan of China (2016YFC0502804).

House Age, Price and Rent: Implications from Land-Structure Decomposition

Abstract

Big cities often witness land price outgrowing structure price. For such cities this paper derives two predictions regarding the dynamics between house prices, rent and structure age. First, older houses have a higher price growth rate than younger ones, even after controlling for location and other attributes; second, the age depreciation of house price, defined as the decline of house price with respect to house age, is slower than the similarly-defined age depreciation of rent. These hypotheses are supported by the micro-data on housing market in Beijing. These two inferences have implications for both real estate valuation and house price index construction.

1. Introduction

Researchers have become increasingly aware that house price should be decomposed into land price and structure price. Glaeser, Gyourko and Saks (2005) suggest that house price appreciation since the 1970s has been largely driven by increasing land costs in the U.S. Based on the land-structure decomposition, Davis and Heathcote (2007) infers the land price from data on house price and structure cost for US cities, which helps explain trends, fluctuations and regional variations in the price of housing. Similarly, Bostic, Longhofer and Redfearn (2007), and Bourassa, Hoesli, Scognamiglio, et al. (2011) focus on the land leverage of houses (the ratio of land to total value); both find that land value accounts for a large share of total home value, and that this decomposition can help explain price fluctuations, regional differences and other important issues in the real estate market.

In this paper we explore this land-structure decomposition from another perspective that complements the existing literature. It has been widely observed that in many big cities land price tends to increase in value faster than structure price due to land scarcity. We begin with the assumption that land price outgrows structure price, offer two theoretical predictions based on this assumption and empirically test them using data from Beijing city: First, older houses (houses with an older structure) have a higher price growth rate than younger ones, even after controlling for location and other important attributes. The intuition is that, as a house ages, the share of land value as a percentage of total value increases. Thus the price of an old house typically increases quickly provided that land price outgrows structure price. Second, the age depreciation of house price, defined as the decline of house price with respect to house age, is slower than the similarly-defined age depreciation of rent. Intuitively, house price more clearly reflects the investment value of a house, while rent represents only its consumption value. For an old house, the consumption value declines due to the depreciation of its structure, but at the same time, investment value depreciates to a lesser extent or even increases because of growing land price. Our theory also shows that if land price grows less quickly than structure price, then both of these predictions will be reversed.

For Chinese cities we are able to directly observe the sales price of new land parcels auctioned on the residential land market. Although land prices are available only for new land parcels, they are a reasonably good proxy of the price for adjacent residential land that has already been developed. Based on the quality-controlled land price (that is, based on the auction prices of new land parcels) and house price indices released by Tsinghua University's Hang Lung Center for Real Estate, in Beijing from 2006-2014, the average annual growth rate of land price was 25.2%, which is about 25% higher than the annual growth rate of house price (20.3%). This implies that land price has grown much faster than that of structure price, as is consistent with results seen in many other big cities with limited land supply (see Glaeser and Gyourko 2006, Davis and Heathcote 2007, Davis and Palumbo 2008 for cities in the US. For evidence in China, see Deng et al. 2012.). From this premise, we empirically test the above two theoretical predictions using a large micro-data set representative of Beijing's housing market. The dataset contains 55,706 second-hand housing sales and 210,600 rental transactions for the period of 2005-2012. Our transaction-level data on house price and rent are suitable for the test. In particular, for each residential complex¹, we have individual observations for both rental and re-sale transactions, facilitating a location-controlled comparison of the age depreciation of house price and rental rate.

Since we focus our study on house age, one empirical challenge is to effectively separate age, cohort and time effects (Coulson and McMillen 2008). Because we have a sufficiently large dataset, we adopt a methodology that controls for age, cohort and time effects in different functional forms, so as to avoid the multi-collinearity problem (McKenzie 2006) and obtain credible estimates of the age effects for both second-hand house sales and rental samples.

Our empirical results are consistent with the two theoretical predictions. Older houses have a higher price growth rate. This is true both at the level of individual houses and at the level of residential complexes. As for the second hypothesis, we estimate the depreciation rate with respect to age for both house price and house rent, and find that the depreciation rate of house rent is 25%-60% higher than that of house price (after controlling for both cohort effect and time trend),

¹ Housing development in many high-density cities in Mainland China occurs at a uniquely large scale and with a high degree of homogeneity in the units built within the typical residential "complex". In each complex, a number of buildings are constructed containing altogether hundreds or even thousands of units, all with essentially equivalent location, architectural design, structure, appliances and finishes.

and this difference is statistically significant.

Our theoretical framework and empirical findings shed light on the mechanism behind the distinct house price growth rates in two housing sub-markets – markets of older and younger housing units whose land leverage ratios differ. This paper also furthers our understanding of the distinct age depreciation rates in the housing sale and rental markets. These two inferences are useful in real estate valuation, for both growth trends and age depreciation rates are crucial factors in the commonly used valuation methodologies such as the hedonic pricing model. Another implication lies in the repeated-sales approach which is widely used to construct house price indices. Since age is not controlled for in the conventional repeated-sales regression, the estimated house price indices will depend on the age distribution in the repeated sales sample – everything else being equal, a larger share of older homes in the sample is likely to generate higher estimates for house price indices.

Housing is both consumption good and investment good (Henderson and Ioannides 1983). In the literature, a standard way to measure housing investment demand is through additionally owned houses other than primary living residence (Rosenthal and Ioniades 1994). However, due to financial friction which is prevalent in housing markets, households may have to divert their investment need into primary living residence when they are constrained from buying multiple houses. It is challenging to measure the investment need for owned-occupied houses. Dusanski and Koc (2007) and Cao, Chen and Zhang (2016) show that expected capital gains increase the investment need of owned-occupied houses. Our paper complements their work by offering another way to identify housing investment demand. We explore the decomposition of housing value into structure value and land value. Our theory suggests that because of the appreciation of land value, the expected capital gain of old houses is higher than that of new houses. Thus we expect a slower age depreciation rate for house prices (which may carry investment purposes) than that for rental prices. Our empirical analysis finds strong evidence supporting those predictions.

The remainder of this paper is organized as follows: Section 2 presents the theory and derives the two hypotheses. Section 3 introduces the data. Section 4 discusses the empirical strategies and reports the empirical results regarding the two hypotheses. The final section concludes.

2. Theory and Hypothesis

Let P_t denote the price of a housing unit at time t whose structure has an age of a years, it can be decomposed into

$$\begin{aligned} P_t &= L_t + S_t \\ &= q^l p_t^l + q_a^s p_t^s \end{aligned} \quad (1)$$

i.e., house price is the sum of land value L_t and structure value S_t , both land value and structure value are expressed as the product of quantity and price. Land price p_t^l and structure price p_t^s have subscript t , because prices change over time. The quantity of structure q_a^s is subscripted by age a , indicating that structure depreciates with age.

Further, let $G_t^L = L_{t+1} / L_t$ and $G_t^S = S_{t+1} / S_t$ be the growth factors of land value and structure value. We have

$$G_t^L = \frac{q^l p_{t+1}^l}{q^l p_t^l} = \frac{p_{t+1}^l}{p_t^l} \quad (2)$$

And

$$G_t^S = \frac{q_{a+1}^s p_{t+1}^s}{q_a^s p_t^s} = \tau \frac{p_{t+1}^s}{p_t^s} \quad (3)$$

Where $\tau = q_{a+1}^s / q_a^s < 1$ determines the depreciation of structure with age. For simplicity we assume it is time- and age-invariant.

Using equation (1), the growth factor of house price is

$$G_t^P = \frac{P_{t+1}}{P_t} = \frac{L_t}{L_t + S_t} G_t^L + \frac{S_t}{L_t + S_t} G_t^S \quad (4)$$

2.1 House Age and House Price Growth Rate

We are interested in how the growth rate of house price changes with age, i.e., what the sign of

$\frac{dG_t^P}{da}$ is. Before we proceed, we shall discuss what exactly $\frac{dG_t^P}{da}$ means. Theoretically, price of a

house is determined by a set of factors, including house age, local amenities, land supply in the surrounding area, employment opportunity in the city and macroeconomic policies and others. A simple view of the complex dynamics of house price is to summarize the time-varying factors into the time effect and age effect. The growth rate of house price, G_t^P , is essentially the derivative of

house price with respect to time. Therefore, $\frac{dG_t^P}{da}$ is essentially the cross derivative of house price

with respect to time and age. The effect of age on house price reflects the depreciation of structure. It is distinct in theory from the effect of fundamental changes in the economy and the change of people's expectation over time. However, empirically it is not easy to disentangle the two effects because time and age have perfect collinearity for a given house. We will return to this point in the empirical part of the paper.

It is reasonable to assume the growth factors of land price and structure price over time, as given in equation (2) - (3), are independent of house age. Rather, they are determined by fundamentals of the economy such as income growth, migration, labor cost and the cost of construction materials. Thus the derivatives of G_t^L and G_t^S with respect to age are both zero. From equation (4), we derive the following

$$\begin{aligned} \frac{dG_t^P}{da} &= \frac{(G_t^L - G_t^S)}{(L_t + S_t)^2} \left[S_t \frac{dL_t}{da} + L_t \left(-\frac{dS_t}{da} \right) \right] \\ &= \frac{(G_t^L - G_t^S)}{(L_t + S_t)^2} \left[S_t \frac{d(q^l p_t^l)}{da} + L_t \left(-\frac{d(q_a^s p_t^s)}{da} \right) \right] \end{aligned} \quad (5)$$

The term in brackets in the above equation is generally positive. First, land quantity and price should not change with structure age, so $\frac{d(q^l p_t^l)}{da} = 0$. Second, while structure quantity

depreciates over time ($\frac{dq_a^s}{da} < 0$), it is reasonable to assume structure price is independent of

structure age, therefore $\frac{d(q_a^s p_t^s)}{da} = p_t^s \frac{dq_a^s}{da} < 0$.² Hence

$$S_t \frac{dL_t}{da} + L_t \left(-\frac{dS_t}{da} \right) > 0 \quad (6)$$

Therefore, from equation (5) we reach the hypothesis (1).

Hypothesis (1): if land price grows faster than structure price, i.e., $G_t^L - G_t^S > 0$, then the prices of older houses have higher growth rates, i.e.,

$$\frac{dG_t^P}{da} > 0$$

Conversely, if $G_t^L - G_t^S < 0$, then prices of older houses will have lower growth rates.

Intuitively, as a housing unit ages, its price is composed more of land price than of structure price, and hence its growth rate increases and becomes closer and closer to the growth rate of land price.

2.2 Age Depreciation of House Price and Rent

Given a housing unit, let R_t be the rental rate at period t. Standard asset pricing theory gives rise to the following equation

² A more general condition for the bracketed term to be negative is that structure value is more elastic with respect to age than land value, i.e., $\left| \frac{dL_t / da}{L_t / a} \right| < \left| \frac{dS_t / da}{S_t / a} \right|$. This can be easily derived from equation (5).

$$P_t = R_t + E[D_{t+1}P_{t+1}] \quad (7)$$

Where D_{t+1} is the stochastic discount factor between period t and $t+1$ and E is the expectation operator.³ Here it is assumed that rent is collected in the beginning of the period. Notice that P_{t+1} contains all the relevant information about future rents and future discount factors, therefore equation (7) is essentially a Gordon growth model with time-varying discount rates. In the appendix we show in detail how to get from a version of Gordon growth model to equation (7).

We apply $P_{t+1} = P_t \times G_t^p$ to equation (7) to get

$$P_t = R_t + P_t \times E[D_{t+1}G_t^p] \quad (8)$$

This equation makes explicit that house price depends on rent and the future growth of house price. Recall that G_t^p is the growth factor of house price between period t and $t+1$, and hence is stochastic at time t .

Clearly P_t and R_t are functions of house age. On the other hand, the stochastic discount factor does not depend on house age, but instead depends on the marginal utility of consumption. Therefore, taking the logarithm of both sides of equation (8), and then taking derivatives with respect to age, we have

$$\frac{d \log P_t}{da} = \frac{\frac{d \log R_t}{da} R_t + \frac{d \log P_t}{da} P_t E[D_{t+1}G_t^p] + P_t E[D_{t+1} \frac{dG_t^p}{da}]}{R_t + P_t \times E[D_{t+1}G_t^p]} \quad (9)$$

Where the last expression in the numerator comes from the Leibniz Rule; that is,

³ In a consumption-based asset pricing model, D_{t+1} is the marginal rate of substitution between consumption bundles in period $t+1$ and period t . All the derivation holds true if the discount factor D_{t+1} is non-stochastic.

$$\frac{dE[D_{t+1}G_t^P]}{da} = E[D_{t+1}\frac{dG_t^P}{da}]$$

After multiplying both sides of equation (9) with $R_t + P_t E[D_{t+1}G_t^P]$ and rearranging terms, we have

$$\frac{d \log P_t}{da} = \frac{d \log R_t}{da} + \frac{P_t}{R_t} E[D_{t+1} \frac{dG_t^P}{da}] \quad (10)$$

It is natural to assume that the aging of a house negatively affects its price and rent, i.e.,

$$\frac{d \log P_t}{da} < 0 \text{ and } \frac{d \log R_t}{da} < 0, \text{ hence equation (10) leads to equation (11).}$$

$$\left| \frac{d \log R_t}{da} \right| - \left| \frac{d \log P_t}{da} \right| = \frac{P_t}{R_t} E \left[D_{t+1} \frac{dG_t^P}{da} \right] \quad (11)$$

where $|\cdot|$ denotes the absolute value operator. Therefore, the gap between age depreciation of rent and that of price depends on dG_t^P / da . If the prices of older houses reveal a higher growth rate, then we expect rent to decline more quickly with age than house price does.

Using equation (4), equation (11) becomes

$$\left| \frac{d \log R_t}{da} \right| - \left| \frac{d \log P_t}{da} \right| = \frac{P_t}{R_t} E \left[D_{t+1} \frac{(G_t^L - G_t^S)}{(L_t + S_t)^2} \left[S_t \frac{dL_t}{da} + L_t \left(-\frac{dS_t}{da} \right) \right] \right] \quad (12)$$

Based on (12), we have the following hypothesis.

Hypothesis (2): If $G_t^L - G_t^S > 0$, then rental rate depreciates more quickly with age than house price, i.e.,

$$\left| \frac{d \log R_t}{da} \right| > \left| \frac{d \log P_t}{da} \right| \quad (13)$$

Conversely, if $G_t^L - G_t^S < 0$, then rental rate depreciates more slowly with age than house price.

Intuitively, house price depends on the rental rate and the expected growth in house price in the future. As a house ages, it depreciates because the rental rate decreases due to the reduced utility flow from an older structure. However, the house price growth rate is higher for older houses given that land price grows faster than structure price, as stated in Hypothesis (1). This mitigates the effect of depreciation of rental rate caused by age, leading to house price depreciating more slowly.

3. Data

Our micro data on second-hand housing sales and rental transactions are from *WoAiWoJia*, a major real estate broker in Beijing, with a local market share of about 10%. The dataset contains 55,706 housing re-sales and 210,600 rental transactions for the period of 2005-2012. The housing unit transactions come from over 2,500 residential complexes distributed throughout Beijing's landscape and hence are representative of Beijing's housing market. It is worth mentioning that the large sample size of our dataset provides us a remarkable advantage in the empirical study. Because the sample size within each complex is large enough to run a complex-specific Hedonic regression, and also because there are a large number of complexes, we can first estimate the growth rate of house price for each complex, and then study how the growth rate of house price varies with house age at the complex level. Moreover, since the unobserved variables of housing units may be correlated within each complex, we cluster the standard errors by complex in our regressions at the housing unit level.

For each transaction, we have information about transaction price (*price* or *rent*), address, and physical attributes of housing such as unit size (*size*), level of decoration (*decoration*) and whether it is on the top floor (*top*). By geo-coding all sales and rental transactions on Beijing's GIS map, we

construct several location attribute measures for each residential complex, including distance from each residential complex to the city center (d_center), whether the complex is within the 2-km reach of a “key” primary school (*school*), of a “Grade-A” hospital (*hospital*) and within the 1-km reach of a subway stop (*subway*). Summary statistics are provided in Table 1.

** Figure 1 here **

** Table 1 here **

Figure 1 shows the geographical distributions for both the sales sample and rental sample in our data. A comparison of the left and right panels reveals that the two samples differ in their spatial distributions, although both samples contain housing units all over the city. Therefore, it is necessary to control for location attributes in our empirical analysis.

Table 1 shows other differences between sales and rental units. Sales units have an average size of 77.6 m² and an average price of 18,250 RMB/m². On average they are about 11.6 km away from the city center. Among these resale units, 52% are within 2km of the closest key primary school, 45% are within 2 km of the closest high-quality hospital, and 38% are within 1 km of the closest subway station.

In contrast, rental units have an average size of 64.3 m² and an average rent of 47.5 RMB/m² per month. Additionally, rental units are in slightly better locations than the sales units. On average they are 11.0 km away from the city center. The percentage of units within 2 km from the key primary school, the Grade-A hospital and within 1 km of the subway station are 64%, 57% and 42% respectively, all of which are higher than the corresponding percentages of sales units. In our empirical analysis, we control for both physical attributes and location attributes.

House age at the time of transaction (*age*) is a key variable in our study. Although the data do not have direct information about house ages\ for each transaction, we located the building year for each complex based on its address and complex name on the broker’s website, and this enables us to calculate house age. Most sales units were built during the period of 2000-2005. While most rental units were also built in about the same period, the variation of the building year for the

rental sample is larger than that for the sales sample⁴. The distribution of building year for both the sales sample and rental sample are reported in Figure 2. The average house age is 11.2 for the sales sample and 13.0 years for the rental sample.

** Figure 2 here **

As mentioned previously, the disentanglement of age effect, year effect and cohort effect is a crucial issue in our empirical study. To control for the cohort effect, we divide all housing transactions into 5-year cohort groups according to their construction year. Thus in total we have 7 cohort groups (that is, 7 cohort dummies). Table 2 shows the distribution of transactions among the 7 groups and the range of house *age* for each group for both the sales and rental samples. Thanks to our large data set, we observe that the variable *age* still has sufficient variation within each cohort. Thus we can control for both the cohort dummy and year effect, and still have a valid estimation of age depreciation in the empirical study.

** Table 2 here **

4. Empirical Analysis

In this section we test the two hypotheses identified in section 2 for the Beijing's housing market, on the premise that $G_t^L - G_t^S > 0$. The house price and land price indices recently released by Tsinghua University, as shown in Figure 3, indicate that land price has grown more quickly than house price from 2006-2014 in Beijing; this implies that the land growth rate is indeed larger than the structure growth rate .

** Figure 3 here **

4.1. Testing Hypothesis (1)

⁴ In order to have a strong comparison between the growth rate and depreciation rate of house price and rent, we only include houses built after 1980 in implementing our empirical equations.

Because only a very limited number of housing units have repeated sales information in our sample period, estimating the price growth rate for individual housing units is not feasible. To test Hypothesis (1), we estimate the average annual price growth rate for each residential complex and then study how the growth rate varies by the house age of each residential complex. We choose residential complexes with enough transactions; that is, with more than 50 or 20 transactions from 2005 to 2012.

Specifically, for the j -th residential complex, we estimate Equation (15) with a linear time trend t (in years).

$$\log(\text{price}_{it}) = \rho_j + \boldsymbol{\phi}_j \cdot \mathbf{X}_i + \gamma_j t + \xi_{it} \quad (15)$$

where the subscripts i, t denote the transaction of housing unit i in year t (the earliest year takes the value of 1), price_{it} denotes the unit price of per square meter for the housing unit i (in complex j) in year t , and \mathbf{X}_i includes the physical features of the housing unit excluding house age since it is perfectly correlated with t within a complex. The estimate of coefficient γ_j proxies the average annual price growth rate of complex j in our study period. Next, we examine how the estimated γ_j varies by the house age of residential complex j :

$$\gamma_j = \phi + \boldsymbol{\lambda} \cdot \mathbf{Y}_j + \eta \text{age}_j + u_j \quad (16)$$

Where age_j is the house age of residential complex j at the beginning of our sample period. \mathbf{Y}_j is a set of location attributes at the complex level, including d_center , $subway$, $hospital$, and $school$. The estimation results of equation (16) are reported in Table 3. The results are reported in Columns (1) and (2) respectively. The coefficients of age in the two columns are both significantly positive, indicating that price growth rate does increase with age.

** Table 3 here **

As an alternative test of Hypothesis (1), we also pool all sales transactions together and regress the

logarithm of house price on transaction year, age, interaction of transaction year and age, and other physical and location attributes at the housing unit level. The coefficient of the interaction term indicates whether the growth rate of house price over time varies by house age. If the coefficient is positive, this means older houses have higher house price growth rate. The regression specification is as in Equation (17).

$$\log(\text{price}_{ijt}) = \alpha + \delta \cdot \mathbf{W}_{ij} + S + \beta \text{age}_{ijt} + \theta t + \eta \text{age}_{ijt} * t + \gamma \cdot \text{cohort}_{ij} + v_{ijt} \quad (17)$$

where W_{ij} is a vector of control variables, including physical attributes of the house such as *size*, *decoration*, *floor*, *top*, and house location attributes such as $\log(d_center)$, *school*, *hospital*, *subway*, S denotes seasonal dummies⁵. Note that we analyze the age effect of house price based on repeated cross sectional data. A classical issue in this kind of regression is the identification of age effect, year effect and cohort effect. To extract the real age effect, it is necessary to filter out year and cohort effects. This is done by including time variable t and cohort vector cohort_j together in the regression as explanatory variables. To address the problem of multi-collinearity among age, cohort and year, we let t be the month in which the transaction happens (with 1 denoting the earliest month), while age_{jt} is still the house age in year. We also let cohort_j be a series of dummy variables indicating the 5-year group during which the transacted house was built⁶. Standard errors are clustered by complex.

The empirical results are reported in Table 4. The coefficients of physical and location attributes are all consistent with our expectations. Smaller houses with better decoration and on higher floors (but not the top floor) have higher prices per square meter, and houses in prime locations (near Central Business District, key primary schools and high-quality hospitals) also have higher prices. The coefficient of t is significantly positive, showing the growing trend for house price from 2006-2012. The coefficient of age is significantly negative, showing that house price decreases by 0.01% for every 1 year increase in housing age. The coefficient of the interaction term $\text{age} * t$ is significantly positive, indicating that older houses have higher growth rates of their prices, as is

⁵ Here we take month January to March as the default, so we have three dummies: month April to June, month July to September, month October to December.

⁶ While we admit that the cohort setting may be arbitrary, we make a robustness check without the cohort dummies in the appendix.

consistent with Hypothesis (1). In column (2) we replace the variable *age* with its logarithm $\log(\text{age})$ as a robustness check. Although the interaction term also has a positive sign, it is not statistically significant.

** Table 4 here **

4.2. Testing of Hypothesis (2)

We test Hypotheses (2) by running the following two regressions of house price and rent separately

$$\log(\text{price}_{ijt}) = \alpha_p + \delta_p \cdot \mathbf{W}_{ij} + S + \beta_p \text{age}_{ijt} + \theta_p T + \gamma_p \text{cohort}_{ij} + v_{p,ijt} \quad (18)$$

$$\log(\text{rent}_{ijt}) = \alpha_r + \delta_r \cdot \mathbf{W}_{ij} + S + \beta_r \text{age}_{ijt} + \theta_r T + \gamma_r \text{cohort}_{ij} + v_{r,ijt} \quad (19)$$

Where price_{ijt} and rent_{ijt} denote the unit sale price and rental price per square meter for the housing unit i in year t , respectively. β_p and β_r denote the depreciation rate of house price and house rent for every 1 year increase in housing age, respectively. Again, standard errors are clustered by complex in both regressions.

We also control for time and cohort effects (T and cohort). As discussed earlier, there is a multi-collinearity issue among age, cohort and time since $\text{age}_{jt} + \text{cohort}_j = t$. To deal with this issue, here we let age_{jt} be a continuous variable of housing age in years, and cohort_j be a series of dummy variables indicating the 5-year group during which the transacted house was built (we also use a 1-year group and 10-year group as robustness checks). We replace the continuous variable t with T , which is a polynomial vector of time; i.e., $T=(t, t^2, t^3, \dots)$ where t is a continuous variable denoting the month in which the transaction happens with the earliest month being 1⁷.

According to Hypotheses (2), we should have

⁷ The same as Table 4, we include a robustness check without cohort effect in the appendix. The results show that we would underestimate the age depreciation rate without controlling for cohort effect.

$$|\beta_p| < |\beta_r| \quad (20)$$

The empirical results are reported in Table 5. Physical attributes are included in the regressions but are not reported. Comparing the coefficients in column (1) with those in column (2), we can see that $|\beta_p| < |\beta_r|$, which is consistent with Hypothesis (2). The *Wald test*⁸ testing if these two coefficients are statistically different also confirms the significant difference between the depreciation rates of house price and house rent. As robustness checks, we also re-define our cohort dummies with 1-year and 10-year cohort groups. Columns (3) and (4) show house price and rent regressions with 1-year cohort dummies, respectively. Columns (5) and (6) present results with 10-year cohort dummies. The Wald test confirms the significant difference between the depreciation rates in the 10-year cohort regressions, although this gap is not that significant in the 1-year cohort regressions (perhaps due to the limited variation within each 1-year cohort).

** Table 5 here **

4.3. Robustness Check in terms of the land leverage hypothesis

In order to get a consistent conclusion in terms of the land leverage hypothesis, we also implement similar tests to other two structural variables *top* and *floor* as a robustness check. The regression results are reported as Table A3 and Table A4 in the Appendix. In Table A3, we include the floor level of house units in the Hedonic regression for both house price and rent. The results show that house units on higher floors have higher prices but lower growth rate of the price; and the positive effect of floor level on the rent is larger than that on the price. On the contrary, house units on top of a building have lower prices (due to the less comfort on the top) but higher price growth rate; and the negative effect on the rent is larger than that on the price. In Table A4, we firstly run the Hedonic regressions on the complex level and find that the variables *top* and *floor* have the same

⁸ As pointed out by the anonymous referee, the two estimated depreciation rates of house price and house rent are correlated with each other, since they are derived from the regressions for two highly related sub-markets. In this case we employ the Wald test, and use the command “*suest*” (designed for seemingly unrelated estimation) in STATA to test whether the two coefficients are statistically different, taking the covariance of the two coefficients into consideration.

impact pattern on house price and rent as in Table A3. Then we repeat what we have done in Table 3. We firstly estimate the growth rate of housing price based on the unit-level Hedonic regression of each complex (γ , not reported as a table), and then use *top* and *floor* to explain the difference in γ across complexes. The results are consistent with those of the tests of housing age, which confirms the hypotheses in terms of the land leverage theory.

5. Conclusion

By decomposing house price into separate prices for land and structure, this paper has studied an important factor in real estate valuation -- house age. The effect of house age on price, rent, and the growth rate of house price depends on the relative growth speed of land price and structure price. For cities where land price grows more quickly relative to structure price, we predict theoretically that older houses have higher growth rates in their house price, and that house rent has a larger depreciation rate with respect to age than that of house price. Both theoretical predictions are reversed if land price grows less quickly than structure price.

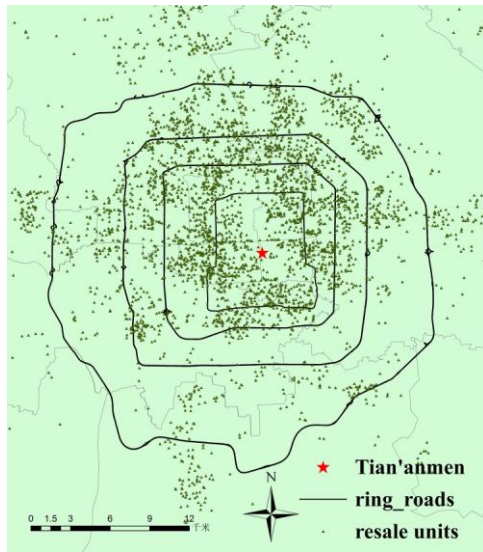
Using our unique micro data for the Beijing market, we find that both predictions are empirically supported. Our analysis sheds light on the distinct house price growth trends in markets with different land leverage ratios. It also explains the difference between age depreciation rates in the sales and rental markets. These two inferences have implications for both real estate valuation and house price index compilation.

While the growth rate of land price is observable for the Beijing market due to the unique land auction policy in China, it is usually not directly observable for markets in many advanced economies. For these markets, our theoretical predictions provide a novel way of testing the dynamic relationship between land price and structure price. -- researchers can infer whether land price outgrows structure price based on the age depreciation patterns of house price and rent which is more easily observable.

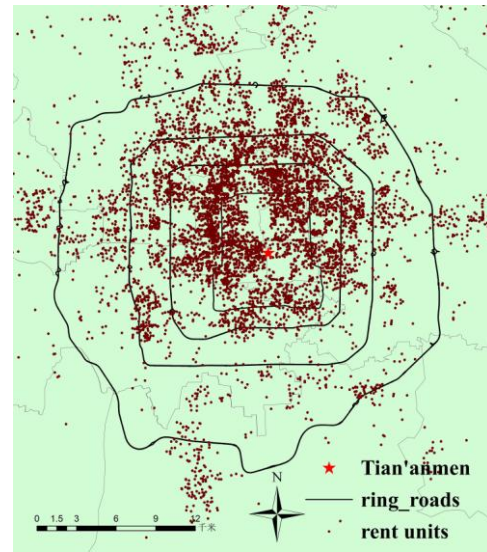
References

- [1] Bostic, R., Longhofer, S., Redfearn C. (2007). Land leverage: decomposing home price dynamics. *Real Estate Economics*, 35(2), 183-208.
- [2] Bourassa, S., Hoesli, M, Scognamiglio, D, et al. (2011). Land leverage and house prices. *Regional Science and Urban Economics*, 41(2), 134-144.
- [3] Cao, Y., Cheng, J. and Zhang, Q. (2016). Housing Investment in Urban China: Evidence from Chinese Household Survey. Working paper.
- [4] Coulson, N., McMillen, D. (2008). Estimating time, age and vintage effects in housing prices. *Journal of Housing Economics*, 17(2), 138-151.
- [5] Davis, M., Heathcote, J. (2007). The price and quantity of residential land in the United States. *Journal of Monetary Economics*, 54(8), 2595-2620.
- [6] Davis, M., Palumbo, M. (2008). The price of residential land in large US cities. *Journal of Urban Economics*, 63(1), 352-384.
- [7] Deng, Y., Gyourko, J., Wu, J. (2012). Land and house price measurement in China. *National Bureau of Economic Research*.
- [8] Dusansky, R., Koç, Ç. (2007). The capital gains effect in the demand for housing. *Journal of Urban Economics*, 61(2), 287-298.
- [9] Glaeser, E., Gyourko, J. (2005). Housing dynamics. *National Bureau of Economic Research*.
- [10] Glaeser, E., Gyourko, J., Saks R. (2005). Why have housing prices gone up? *American Economic Review*, 95(2), 329--333.
- [11] Henderson, J., Yannis, M. (1983). A model of housing tenure choice. *The American Economic Review*, 73(1), 98-113.
- [12] Hornstein, A., Greene, W. (2012). Usage of an estimated coefficient as a dependent variable. *Economics Letters*, 116(3), 316-318.
- [13] Ioannides, Y., Rosenthal, S. (1994). Estimating the consumption and investment demands for housing and their effect on housing tenure status. *The Review of Economics and Statistics*, 127-141.
- [14] McKenzie, D. (2006). Disentangling age, cohort and time effects in the additive model. *Oxford Bulletin of Economics and Statistics*, 68, 473–495.
- [15] Saxonhouse, G. (1976). “Estimated parameters as dependent variables.” *The American Economic Review*, 66(1), 178-183.

Fig.1 Spatial distribution of transaction samples

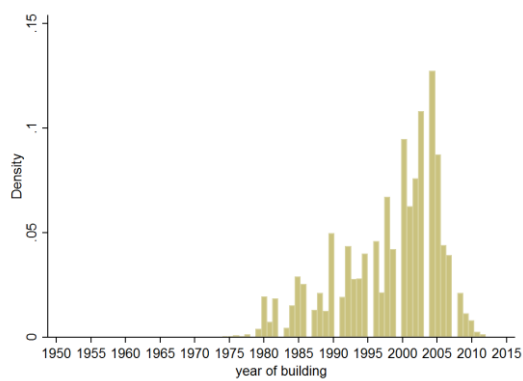


Spatial distribution of sales units

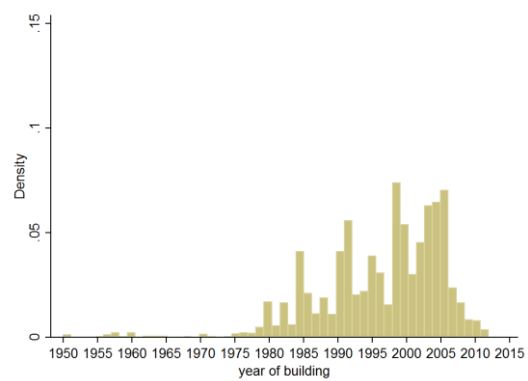


Spatial distribution of rental units

Fig.2 Distribution of building years

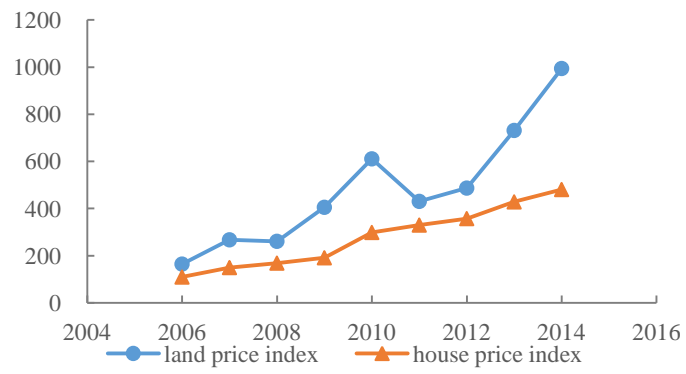


Sales units



Rental units

Fig.3 Hedonic land and house (land + structure) price indices in Beijing



Source: Tsinghua Hang Lung Center for Real Estate

Table 1 Summary statistics

Variable	Definition	Sales observations			Rental observations		
		Obs.	Mean	Std.	Obs.	Mean	Std.
<i>price/rent</i>	Transaction price/rent (RMB yuan per square meter for price and RMB yuan per square meter per month)	56411	18250.83	8798.92	254322	47.61	21.74
<i>size</i>	Housing unit size (square meter)	56411	77.56	30.96	254322	51.34	50.37
<i>age</i>	House age (years)	56411	11.42	7.37	216076	13.07	8.83
<i>decoration</i>	Decoration status (4=best; 1=worst)	56411	2.72	0.97	254322	1.80	1.62
<i>floor</i>	Floor number	56411	7.59	6.03	254322	7.04	5.79
<i>top</i>	Whether the unit is on the top floor (1=yes, 0=no)	56411	0.12	0.33	254322	0.12	0.33
<i>d_center</i>	Distance to the city center (km)	56411	11.63	5.85	254322	10.97	5.52
<i>school</i>	Whether the observation is within 2km of key primary school	56411	0.52	0.50	254322	0.64	0.48
<i>hospital</i>	Whether the observation is within 2km of hospital	56411	0.45	0.50	254322	0.57	0.49
<i>subway</i>	Whether the observation is within 1km of subway stop	56411	0.38	0.48	254322	0.42	0.49

Table 2 Sample distribution by cohort groups and within-cohort age range

Building year	Sales	Rentals
1981-1985	Sample size: 3379 Age range: 20-31	Sample size: 17402 Age range: 20-31
1986-1990	Sample size: 5539 Age range: 15-26	Sample size: 25977 Age range: 15-26
1991-1995	Sample size: 7203 Age range: 10-21	Sample size: 34571 Age range: 10-21
1996-2000	Sample size: 12351 Age range: 5-16	Sample size: 43883 Age range: 5-16
2001-2005	Sample size: 21006 Age range: 0-11	Sample size: 62469 Age range: 0-11
2006-2010	Sample size: 5626 Age range: 0-6	Sample size: 20842 Age range: 0-6
2011-2012	Sample size: 160 Age range: 0-1	Sample size: 937 Age range: 0-1

Table 3 House age and house price growth: complex-level regressions

γ	>20 transactions		>50 transactions	
	(1)	(2)	(3)	(4)
<i>age</i>	0.00123*** (0.000239)	0.00111*** (0.000181)	0.000941** (0.000363)	0.000969*** (0.000273)
$\log(d_center)$	0.00346 (0.00230)	0.00476** (0.00204)	0.00374 (0.00409)	0.00481 (0.00296)
<i>school</i>	0.0102 (0.00622)	0.0101** (0.00413)	0.0175** (0.00720)	0.0117** (0.00587)
<i>hospital</i>	0.00477 (0.00332)	-2.39e-05 (0.00282)	-0.00245 (0.00573)	-0.00444 (0.00418)
<i>subway</i>	0.00137 (0.00353)	0.000200 (0.00302)	0.00289 (0.00579)	0.00150 (0.00449)
Constant	0.170*** (0.00904)	0.173*** (0.00723)	0.171*** (0.0140)	0.176*** (0.0105)
Heteroscedasticity approach	OLS, Robust	WLS	OLS, Robust	WLS
Observations	607	607	194	194
R-squared	0.068	0.080	0.085	0.100

Note:

(1) First, we regress house price with time trend t (in month) using all the transaction data to get price growth rate γ , and then we regress growth rate γ with housing age. Here we choose those complexes with enough transactions to insure the regression. Column (1) and (2) show the results of regressions based on complexes with more than 20 transactions, column (3) and (4) show the results of regressions based on complexes with more than 50 transactions.

(2) According to Saxonhouse (1976), Hornstein, Greene (2012), et al., we use two different methods to mitigate the heteroskedasticity problem. In column (1) and (3) we use formula allowing for the presence of the heteroscedasticity in OLS regression and adjust the t statistic. In column (2) and (4) we use the weighted least square method, by the inverse of the estimated standard error of the predicted γ in the first stage.

(3) The coefficient of age in this table indicates that older houses have, in price per square meter, a higher growth rate.

(4) Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4 House age and house price growth: unit-level regressions

<i>log(price)</i>	(1)		(2)	
	Coefficient	Std. Err	Coefficient	Std. Err
<i>log(size)</i>	-0.107***	(0.0118)	-0.106***	(0.0117)
<i>decoration</i>	0.0128***	(0.00234)	0.0130***	(0.00234)
<i>top</i>	-0.0498***	(0.00523)	-0.0496***	(0.00525)
<i>floor</i>	0.00139***	(0.000489)	0.00138***	(0.000490)
<i>t</i>	0.0167***	(0.000368)	0.0174***	(0.000899)
<i>age</i>	-0.0123***	(0.00323)		
<i>age*t</i>	0.00813***	(0.00148)		
<i>log(age)</i>			-0.0966***	(0.0186)
<i>log(age)*t</i>			0.000149	(0.000293)
<i>log(d_center)</i>	-0.0305***	(0.00860)	-0.0305***	(0.00853)
<i>school</i>	0.0679***	(0.0112)	0.0676***	(0.0112)
<i>hospital</i>	0.0689***	(0.0114)	0.0696***	(0.0113)
<i>subway</i>	0.0667***	(0.00951)	0.0672***	(0.00942)
Constant	9.266***	(0.0893)	9.307***	(0.0774)
Cohort settings	Every 5 years		Every 5 years	
Observations	55,427		55,427	
R-squared	0.725		0.725	

Note:

(1) Here we use the traditional Hedonic function of house price to test Hypothesis (1). We add in an interaction term *age*t* to the Hedonic function, and its coefficient is significantly positive, indicating that older houses have a higher growth rate of their prices.

(2) We include cohort dummies in this regression, which are lumped into groups according to 5 year intervals, based on year of construction.

(3) Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table 5 Age depreciation of house price and rent

VARIABLES	(1)	(2)	Wald test	(3)	(4)	Wald test	(5)	(6)	Wald test
cohort setting	log(price)	log(rent)	$ \beta_p < \beta_r $	log(price)	log(rent)	$ \beta_p < \beta_r $	log(price)	log(rent)	$ \beta_p < \beta_r $
	5years	5years		1year	1year		10years	10years	
<i>age</i>	β_p -0.00693** (0.00296)	β_r -0.0112*** (0.00274)	Chi2=18.25 P=0.0000	β_p -0.0286 (0.0223)	β_r -0.0376*** (0.0120)	Chi2=0.20 P=0.6553	β_p -0.00622*** (0.00174)	β_r -0.00772*** (0.00141)	Chi2=7.56 P=0.0060
<i>t</i>	0.0239*** (0.000772)	0.00638*** (0.000408)		0.0257*** (0.00180)	0.00852*** (0.00107)		0.0239*** (0.000727)	0.00603*** (0.000385)	
<i>t</i> ²	-5.57e-05*** (5.33e-06)	1.66e-05*** (2.72e-06)		-5.57e-05*** (5.14e-06)	1.70e-05*** (2.70e-06)		-5.61e-05*** (5.14e-06)	1.72e-05*** (2.78e-06)	
Constant	10.26*** (0.116)	6.194*** (0.104)		10.84*** (0.570)	6.879*** (0.307)		10.26*** (0.108)	6.112*** (0.0850)	
Cohort settings	Every 5 years			Every 1 year			Every 10 years		
Physical attributes	YES	YES		YES	YES		YES	YES	
Locational attributes	YES	YES		YES	YES		YES	YES	
seasonality	YES	YES		YES	YES		YES	YES	
clustered by	complex	complex		complex	complex		complex	complex	
Observations	55,427	204,995		55,427	204,995		55,427	204,995	
R-squared	0.733	0.515		0.735	0.517		0.732	0.514	

Note:

(1) We test Hypothesis (2) with this regression. Here we include cohort dummies based on year of construction using 3 different settings: every 1 year, every 5 years and every 10 years.

(2) To mitigate the multi-collinearity problem, we use the following function form: *age* is a continuous variable denoting housing age, *cohort* is a series of dummy variables indicating the 5-year group during which the transacted house is built, in addition we use the polynomial function of time trend (t, t^2) where t is a continuous variable denoting the month in which the transaction occurs.

(3) The coefficient of age indicates the age depreciation rate of house price and rent respectively. We find that house price depreciates less than house rent, and the Wald test confirms the difference between them. The results in the 5-year and 10-year cohort setting both hold, showing the results are robust.

(4) Robust standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Appendix: Robustness check in terms of the cohort effect

Table A1: Robustness check in terms of the cohort effect for Table 4

VARIABLES	(1) log(<i>price</i>)	(2) log(<i>price</i>)	(3) log(<i>price</i>)	(4) log(<i>price</i>)
log(<i>size</i>)	-0.114*** (0.0121)	-0.103*** (0.0116)	-0.116*** (0.0120)	-0.105*** (0.0114)
<i>decoration</i>	0.00640*** (0.00240)	0.0189*** (0.00182)	0.00593** (0.00240)	0.0185*** (0.00183)
<i>top</i>	-0.0515*** (0.00532)	-0.0533*** (0.00505)	-0.0478*** (0.00532)	-0.0500*** (0.00503)
<i>floor</i>	0.00164*** (0.000504)	0.00171*** (0.000496)	0.00124** (0.000495)	0.00135*** (0.000490)
log(<i>d_center</i>)	-0.0265*** (0.00897)	-0.0278*** (0.00881)	-0.0285*** (0.00888)	-0.0296*** (0.00874)
<i>school</i>	0.0670*** (0.0111)	0.0647*** (0.0110)	0.0695*** (0.0110)	0.0670*** (0.0110)
<i>hospital</i>	0.0721*** (0.0116)	0.0735*** (0.0116)	0.0735*** (0.0116)	0.0748*** (0.0116)
<i>subway</i>	0.0666*** (0.00948)	0.0668*** (0.00923)	0.0667*** (0.00950)	0.0670*** (0.00926)
<i>t</i>	0.0234*** (0.000689)		0.0227*** (0.000723)	
<i>t</i> ²	-0.00612*** (0.000534)		-0.00608*** (0.000536)	
<i>age</i>	-0.00794*** (0.00123)	-0.00756*** (0.00116)		
<i>age*t</i>	0.00924*** (0.00159)	0.00912*** (0.00149)		
log(<i>age</i>)			-0.0870*** (0.0127)	-0.0808*** (0.0122)
log(<i>age</i>)*<i>t</i>			0.000783*** (0.000202)	0.000765*** (0.000193)
Constant	8.960*** (0.0604)	8.917*** (0.0782)	9.084*** (0.0651)	9.030*** (0.0784)
Sale month dummies	NO	YES	NO	YES
Seasonality	YES	NO	YES	NO
Observations	55,427	55,427	55,427	55,427
R-squared	0.727	0.780	0.728	0.781

Note:

(1) Here we use the traditional Hedonic function of house price to test Hypothesis (1). We add in an interaction term *age*t* to the Hedonic function, and its coefficient is significantly positive, indicating that older houses have a

higher growth rate of their prices. We also take the logarithm function of age in Column (3) and (4) as a robustness check.

(2) This table acts as a robustness check for Table 4 by excluding the cohort dummies in Column (1) and (3). In Column (2) and (4), we replace the quadratic polynomial function of time trend by the sale month dummies.

(3) Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Table A2: Robustness check in terms of the cohort effect for Table 5

VARIABLES	(1)	(2)	Wald test	(3)	(4)	Wald test
	log(<i>price</i>)	log(<i>rent</i>)	$ \beta_p < \beta_r $	log(<i>price</i>)	log(<i>rent</i>)	$ \beta_p < \beta_r $
<i>age</i>	-0.00199*** (0.000667)	-0.00590*** (0.000622)	Chi2=345.56 P=0.0000	-0.00168** (0.000660)	-0.00594*** (0.000617)	Chi2=474.52 P=0.0000
<i>t</i>	0.0239*** (0.000725)	0.00635*** (0.000347)				
<i>t</i> ²	-0.00567*** (0.000527)	0.00149*** (0.000281)				
Constant	8.899*** (0.0589)	4.732*** (0.0363)		8.863*** (0.0786)	4.747*** (0.0405)	
Physical attributes	YES	YES		YES	YES	
Locational attributes	YES	YES		YES	YES	
Sale month dummies	NO	NO		YES	YES	
Seasonality	YES	YES		NO	NO	
Cluster by	complex	complex		complex	complex	
Observations	55,427	204,995		55,427	204,995	
R-squared	0.726	0.504		0.779	0.512	

Note:

(1) We test Hypothesis (2) with this regression. This table acts as a robustness check for Table 5 by excluding cohort dummies. In Column (3) and (4) we further replace the time trend function with sale month dummies. The regression results show consistence of the relationship between age depreciation of house price and rent (the age depreciation rate of house rent is larger than that of house price). While the coefficients show that the cohort effect lead to a under-estimation of age depreciation.

(2) Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Appendix: Robustness check in terms of the land leverage hypothesis

Table A3: Robustness check in terms of the land leverage hypothesis – unit level

VARIABLES	(1) log(<i>price</i>)	(2) log(<i>price</i>)	(3) log(<i>price</i>)	(4) log(<i>rent</i>)
log(<i>size</i>)	-0.104*** (0.00295)	-0.119*** (0.00329)	-0.119*** (0.00330)	-0.361*** (0.00126)
log(<i>age</i>)	-0.0363*** (0.00185)	-0.0399*** (0.00206)	-0.0399*** (0.00206)	-0.0731*** (0.00118)
<i>decoration</i>	0.0190*** (0.00120)	0.0117*** (0.00127)	0.0117*** (0.00127)	0.0149*** (0.000465)
<i>top</i>	-0.0490*** (0.00323)	-0.0868*** (0.0102)	-0.0473*** (0.00362)	-0.0818*** (0.00228)
<i>floor</i>	0.00125*** (0.000183)	0.00116*** (0.000205)	0.00382*** (0.000577)	0.00665*** (0.000132)
<i>trend</i>		0.0173*** (5.94e-05)	0.0177*** (8.52e-05)	
<i>trend*top</i>		0.000642*** (0.000155)		
<i>trend*floor</i>			-4.23e-05*** (8.61e-06)	
log(<i>d_center</i>)	-0.0293*** (0.00224)			-0.0438*** (0.00156)
<i>school</i>	0.0673*** (0.00281)	0.0782*** (0.00303)	0.0781*** (0.00303)	0.0777*** (0.00198)
<i>hospital</i>	0.0758*** (0.00293)	0.0781*** (0.00326)	0.0779*** (0.00326)	0.0726*** (0.00196)
<i>subway</i>	0.0681*** (0.00230)	0.0740*** (0.00255)	0.0740*** (0.00255)	0.0638*** (0.00161)
Constant	8.938*** (0.0720)	9.089*** (0.0166)	9.065*** (0.0169)	4.853*** (0.0154)
Transaction monthly dummies	YES	NO	NO	YES
Seasonality	NO	YES	YES	NO
Observations	55,706	55,706	55,706	210,286
R-squared	0.780	0.723	0.723	0.514

Note:

(1) Column (1) is the Hedonic function for housing price, we control for transaction monthly dummies in the regression. Column (4) is the similar regression for housing rent, compare the coefficients of top and floor, being on the top of a building has a negative impact on housing price, but is smaller than the negative impact on housing

rent. While being on higher floors would have a positive impact on housing price, and also smaller than the positive impact on housing rent.

(2) In column (2) and (3) we employ the interaction term $trend*top$ and $trend*floor$ to test the growth rate of housing price with different structural features. $Trend$ refers to a continuous variable showing the transaction month. Their coefficients show that units on higher floors have lower growth rate in their prices, but those on the top have higher growth rate.

(3) Standard errors in parentheses; *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table A4: Robustness check in terms of the land leverage hypothesis – complex level

VARIABLES	(1) $\log(price)$	(2) $\log(rent)$	(3) γ	(4) γ
$\log(size)$	-0.0795*** (0.00428)	-0.314*** (0.00224)		
$\log(age)$	-0.0356*** (0.00251)	-0.0715*** (0.00169)		
<i>decoration</i>	0.0194*** (0.00178)	0.0146*** (0.000814)		
<i>top</i>	-0.0547*** (0.00453)	-0.0956*** (0.00379)	0.0810*** (0.0233)	
<i>floor_mode</i>	0.000666* (0.000342)	0.00541*** (0.000256)		-0.00111** (0.000493)
$\log(d_center)$	-0.0260*** (0.00312)	-0.0437*** (0.00237)	0.00278 (0.00299)	0.00389 (0.00351)
<i>school</i>	0.0788*** (0.00377)	0.0961*** (0.00284)	0.0152*** (0.00575)	0.0107* (0.00642)
<i>hospital</i>	0.0814*** (0.00389)	0.0875*** (0.00285)	-0.00367 (0.00418)	-1.72e-05 (0.00484)
<i>subway</i>	0.0609*** (0.00311)	0.0635*** (0.00235)	0.00259 (0.00450)	0.000590 (0.00517)
Constant	8.828*** (0.0725)	4.655*** (0.0216)	0.179*** (0.0103)	0.197*** (0.0126)
Transaction monthly dummies	YES	YES	NO	NO
Heteroscedasticity approach	--	--	WLS	WLS
Observations	25,215	74,019	194	194
R-squared	0.827	0.579	0.098	0.057

Note:

(1) In column (1) – (2), we simply run a hedonic function for housing price and rent separately on the complex level, where the variable *floor_mode* here refers to the mode of floor in each complex, while other structural

feature variables are the mean value for units in the same complex. The coefficients of *top* and *floor* show the same pattern as in column (1) and (4) in Table A3.

(2) In column (3) – (4), we repeat the tests as in Table 3. γ is the coefficient of time trend in the Hedonic regression of unit-level for each complex (we also choose those complexes with more than 50 transactions). And then we use the structural and locational features to explain the difference in γ , which is the growth rate of housing price. The coefficients of *top* and *floor_mode* show the same pattern as in Column (2) and (3) in Tale A3.

(3) Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

Appendix: From the Gordon growth model to equation (7)

Let R_t and D_{t+1}^t denote the rental rate at period t and the discount factor between period t and $t+1$, then house price at period t can be written as the discounted present value of rents from the current and future periods, as in the following equation.

$$P_t = R_t + E_t \left[\sum_{j=1}^{\infty} \left(\prod_{k=1}^j D_{t+k}^{t+k-1} \right) R_{t+j} \right] \quad A(1)$$

where E_t is the expectation operator that take all the information available at time t into account.

This equation is the Gordon growth model with time-varying stochastic discount rate.

In period $t+1$, we have the following

$$\begin{aligned} P_{t+1} &= R_{t+1} + E_{t+1} \left[\sum_{j=1}^{\infty} \left(\prod_{k=1}^j D_{t+k+1}^{t+k} \right) R_{t+j+1} \right] \\ &= R_{t+1} + E_{t+1} \left[\sum_{j=2}^{\infty} \left(\prod_{k=2}^j D_{t+k}^{t+k-1} \right) R_{t+j} \right] \end{aligned} \quad A(2)$$

We multiply both sides of the above equation with D_{t+1}^t , and then take expectation using E_t . This leads to

$$\begin{aligned} E_t[D_{t+1}^t P_{t+1}] &= E_t[D_{t+1}^t R_{t+1}] + E_t \left\{ D_{t+1}^t E_{t+1} \left[\sum_{j=2}^{\infty} \left(\prod_{k=2}^j D_{t+k}^{t+k-1} \right) R_{t+j} \right] \right\} \\ &= E_t[D_{t+1}^t R_{t+1}] + E_t \left\{ \left[\sum_{j=2}^{\infty} \left(\prod_{k=1}^j D_{t+k}^{t+k-1} \right) R_{t+j} \right] \right\} \\ &= E_t \left[\sum_{j=1}^{\infty} \left(\prod_{k=1}^j D_{t+k}^{t+k-1} \right) R_{t+j} \right] \end{aligned} \quad A(3)$$

In equation A(3) we used the law of iterated expectations, i.e., $E_t[E_{t+1}(\cdot)] = E_t(\cdot)$. Plugging A(3) into A(1), we have

$$P_t = R_t + E_t[D_{t+1}^t P_{t+1}] \quad A(4)$$

Equation A(4) is exactly equation (7) in the paper, except that we omitted the superscript t in the discount factor.