Revisiting the Chinese Housing Boom

Zhihan Yu[*]

In the 2008 global financial crisis, the Chinese government launched a fiscal stimulus of 4 trillion RMB package coupled with easy monetary policy to buffer domestic economy from external shocks. As a result, house prices boomed. Many studies in subsequent years identified the above as the key drivers to the real-estate price hike. Yet, this paper stresses the importance of population migration as a long-term fundamental factor affecting regional house prices, which are largely overlooked in the literature. Constructing a dynamic spatial model with borrowing constraints, it shows that household migration decisions have significant influences over municipal house prices, expectations of future house prices will not only spur speculative investments, but also incentivize households migration, and the impact of speculative investment on house prices may not be as dominating a factor as previously believed, especially against the backdrop of the Chinese urbanization.
Keywords: Chinese housing price, migration, expectations, collateral constraint

1. Introduction

Rapid increase in real estate prices is a major concern to both academics and policy workers. Continuous high real estate prices, on the one hand, may undermine the basis of social stability due to a lack of affordability. On the other hand, it may draw disproportionate amounts of social resources and hollow out the economy, thereby corrupting the foundations for future economic growth; on top of these, a sudden collapse in house prices may translate into a long-lasting economic recession, affecting not only domestic regional economies, but also disrupting global markets. Examples of boom-bust cycles in house prices are ample. Two recent housing boom-bust events are the Japan Housing Bubble of 1990 and the US Housing Bubble of 2007.

As we have learned from past events of the Japanese and US housing booms, it is the combined effects of cheap credit and market exuberance that leads to ballooning house prices. In this paper, I would show that while the Chinese government may have practiced loose monetary policies, cheap credit alone does not single-handedly explain the rocketing house prices. Market exuberance and the notion of "never falling house

[*] Zhihan Yu (email: zhihanyu@uchicago.edu). I am grateful for Veronica Guerrieri, Greg Kaplan, Alexandra Gonzalez, Haishi Li, and numerous seminar audiences in the Economics Department at the University of Chicago for their insights and guidance. All the errors and omissions are mine.

prices" is prolonged and reinforced by mass urbanization trends across the country. The intuition behind China's housing boom is relatively simple: individuals migrate to large cities in search of better living qualities and higher wages. Yet, migration from city to city is costly—both the physical costs of migrating and the costs associated with living in cities with higher price levels. With cheaper credit, individuals who were originally unable to afford a change in cities may now do so. The direct results are megacities where hundreds of thousands of migrant households reside. This migratory pattern pushes the demand for housing within the destination cities and reinforces the growing house prices. Investors, now faced with cheaper credit, observe the growing real estate prices and speculate in the market. Productive firms within destination cities experience positive spillover effects due to the spatial agglomeration of physical capital, consumers, and labor. It directly results in an increase in real wages, thereby drawing previously indifferent households into the megacity and establishing the feedback loop for market exuberance.

This paper is organized as follows. Section 2 provides the background information, describing China's urbanization and reviewing the literature. Sections 3 presents a generalized framework with full environment. Section 4 simplifies the model to capture key insights. Section 5 presents the empirical verification through regression analysis. Finally, Section 6 concludes the paper.

2. Background

It is almost as if China's urban sprawl sprouted overnight. From an urbanization rate of 17.92% in 1978, China's urbanization rate exceeded the world average of 54.74%, reaching to levels of 58.52% in 2017 (Xiao *et al*., 2018). According to a 2009 report published by McKinsey Global Institute, should urbanization rates in China continue at the levels observed at the beginning of the century? Chinese urban population in 2030 is projected to reach over 1 billion (Woetzel *et al*., 2009). Furthermore, the report points out that by 2025, an additional 350 million people will become urban dwellers, among which 240 million will be migrants.

Land and houses are often used as collateral for borrowing agents. Since the amount agents can borrow are related to the total market value of the collateralized asset, changes in asset prices will heavily affect an agent's ability to borrow. Recent scholars have studied the role of housing collaterals and its impact on the economy. Zhao (2013) explored the dynamics of how different levels of collateral constraints can induce rational housing bubbles; Berger *et al.* (2018) examined how house prices affect consumer spending under an economy with a collateral constraint. Finally, there is a large body of Chinese literature that identify loose financial market to cause inflated home prices (Zhu, 2010; Han, 2016; Tan *et al*., 2018; Wu, 2018). This paper incorporates this source of financial friction to understand how the existence and the levels of collateral constraints affect house prices within China's economy.

As an investment asset, house prices are largely susceptible to economy-wide expectations. There was a large literature that study the role of heterogenous expectation of fundamentals in an economy (Harrison and Kreps, 1978; Scheinkman and Xiong, 2003; Acemoglu *et al*., 2006; Piazzesi and Schneider, 2009; Dumas *et al*., 2009; Geanakoplos, 2010). Burnside *et al*. (2016) developed a theoretical model of agents with heterogenous expectations about long-run fundamentals that affect house prices and illustrated how boom-bust cycles are thus generated. Kaplan *et al.* (2017) showed, through a series of decompositions, that a change in economy-wide expectation was the main driver for the collapse in house prices during the Great Recession. Like its predecessors, this paper incorporates the role of expectation in studying its impact on household decisions and subsequent effects on Chinese house prices.

In sum, the paper delivers a tractable framework to understand China's rocketing house prices at the granularity of the municipal level. It investigates how the expectations of individual agents and the existence of financial constraints affect their consumptions, investments, and migratory decisions that ultimately impact observed house prices across cities in China. By incorporating all the afore-mentioned properties of housing units, this paper argues that to fully understand the Chinese house price dynamics, academics must analyze the financial environment alongside the spatial features of the economic setting.

3. Generalized Model

This section will first present the full theoretical model describing the economic environment that can be used for structural analysis. In the second part, a simplified version of the generalized model is used to provide intuition of relevant mechanisms at play. In the general model, there is a total of J cities, each populated with $N_{j,t}$ number of households of overlapping generations at time t. Each city has two sectors of production: production of non-durable consumption goods (denoted $Y_{j,t}$) and production of housing units (denoted $H_{j,t}$). Aside from the real-production markets, there is a national financial market where households can purchase/sell risk-free assets that yield some market-determined return in the following period. Households have heterogeneous preferences and form expectations of future period house-prices for each city. Goods are traded across cities, subject to trade cost, and individual households can migrate from city to city subject to migration friction, modeled as a positive cost in the budget constraint (as opposed to a disutility). Again, the purpose of this paper is to illustrate how spatial heterogeneity with the presence of collateral constraints promote migration that ultimately aerates house prices.

3.1. Cities and Production

In the model, cities are platforms where economic activities between households and firms take place. There is a total of J cities, each consisting of two sectors: the non-durable sector and the real-estate sector. For simplicity, it is assumed that, in each city, there is only one firm that produces non-durable goods and one firm that produces housing units. Furthermore, non-durable goods from different cities are viewed as heterogeneous goods, sold at market-determined price of P_j. Observed productivity (Ψ) of a city is determined by two fundamental factors: the raw efficiency (z_j); and the agglomeration effect of population size. Following Gaubert (2014) and Xuan (2018), the relationship between observed productivity and agglomeration is described as follow:

$$\log\Psi_{j,t}\left(z_j, N_{j,t}\right) = \sigma\log N_{j,t} + \log\left(z_j\right) \tag{1}$$

The parameter σ measures the classic log-linear agglomeration externalities. Specifically, this agglomeration effect captures the fact that when two cities have equal raw productivities, the observed productivity is higher for the city with the larger population size. Non-durable production at time t utilizes land ($\xi^y_{j,t}$) and labor ($l^y_{j,t}$) and exhibits constant returns to scale:

$$Y_{j,t} = \Psi_j\left(\xi^y_{j,t}\right)^\alpha\left(l^y_{j,t}\right)^{1-\alpha} \tag{2}$$

Similarly, production of housing units at time t utilizes land ($\xi^h_{j,t}$) and labor ($l^h_{j,t}$) and exhibits constant returns to scale, but does not enjoy an agglomeration effect:

$$H_{j,t} = \left(\xi^h_{j,t}\right)^\beta\left(l^h_{j,t}\right)^{1-\beta} \tag{3}$$

Within the city's housing sector, there is a housing rental market represented by a single firm making zero profits. The firm supplies housing in the rental market by frictionlessly combining housing units and offer them in the market for a fixed ratio of sale price. The leasing agreement (the combination of rental housing size and price) changes at the end of every period.

3.2. Financial Market

In the economy, there is a financial market accessible by all households at the national level. In the financial market, only one type of risk-free asset (i.e. bonds) is traded and ownership houses serve as the only form of collateral for borrowing. For every unit of asset a_{t+1} purchased in period t, and the value of the asset next period is

$R_{t+1}a_{t+1}$. Both lenders and borrowers are subject to the same real interest rate (R_{t+1}). All households are subject to the following borrowing constraint:

$$a_{t+1} \geqslant -(1-\theta)\, h_{t+1} P_t^H \tag{4}$$

where $h_{t+1} \in R_+^j$ denotes the household's next period housing stock; $P_t^H \in R_+^j$ denotes the current house prices, the down payment ratio $\theta \in (0,1)$. Another way to interpret Equation (4) is to see housing as ATM's, where the maximum amount of withdrawal (determined by θ) is contingent on the households' possession of total housing assets at market value.

3.3. Households

3.3.1. Preferences and Utility

Households have heterogeneous preferences for non-durable goods and identical preference for housing units. I model household utility in two periods, where households acquire utility through the consumption of non-durable goods and the enjoyment of housing services of the city in which they reside. The utility of individual i living in location (j_t^i) at time t is described as follows:

$$u_t^i\left(c_t, h_t, j_t^i\right) = \zeta \ln c_t + \zeta_H \ln s_t^{j_t^i} \tag{5}$$

where ζ denotes the Cobb-Douglas taste constants for goods from city j and ζ_H denotes the Cobb-Douglas taste constant for housing services, s_t; $\zeta_j \geqslant 0$ for $j = \{1,\ldots,J\}$, $\zeta_H \geqslant 0$. I further assume that the utility function is homogeneous of degree one.

Following Kaplan *et. al.* (2019), housing services come in two tenure types: rental and owner. Rental housing generates housing services one-for-one with the size of the house, i.e. $s_{rentel} = \rho$; owner-occupied housing generates $s_{owner} = \kappa h$, with $\kappa > 1$.[1] It should be noted that that when an individual living in city j, owning housing stock $h_j > 0$, the household is barred from the housing rental market of city j, $\rho_j = 0$. In other words, households do not enter the housing rental market for the city in which they have positive housing stock.

3.3.2. Expectations of Prices

Households do not have perfect information in the modeled economy. They can

[1] Both ρ and h have the same units and represent housing quantities, but they are denoted differently to reflect the difference in tenure types.

only perfectly observe migration costs and the prices of all the transacted goods of the current period. In order to make decisions across time, households must form reasonable expectations with regards to future prices. Specifically, individuals naively believe that next period wages and non-durable prices remain at the same levels as they are in the current period. They form this belief because they do not anticipate migratory changes of other households and by recognizing the fact that their own migration decisions are not significant enough to affect local markets. Yet, households have priors with regards to house price changes. Based on observed changes in house prices, they update their priors in a Bayesian fashion. The following table summarizes household beliefs:

Table 1. Expectations and Beliefs

	Source of heterogeneity	Beliefs
Next period city *j* wages ($\omega_{j,t+1}$)	No	$\omega_{j,t}$
Next period city *j* non-durable goods price vector observed at location *l* ($P_{j,t+1}(l)$)	No	$P_{j,t+1}(l)$
Growth rate of city *j* house price – mutually independent priors ($\frac{P^H_{j,t+1}}{P^H_{j,t}}$)	Yes	Drawn from distribution

3.3.3. The Household Optimization Problem

This paper model household decision as a two-period decision: The households choose non-durable consumption bundles in two periods (c_t, c_{t+1}),housing stock adjustments (τ), asset investments (a_{t+1}), rental decision (ρ_{t+1}), and location decisions for the next period (j_{t+1}). Even though individuals live in a multiperiod economy, their current period decisions can be identified by a two-period model based on their expectations for future prices. At the beginning of the second period, households receive an "information shock" and update their beliefs to adjust their current (and subsequent) period(s) decisions. Thus, a household can be characterized by their next-period housing tenure types: renters or homeowners. The homeowner-households' decision can be described as

util (homeowners):

$$\begin{aligned}
&\text{maximize} && u\left(c_t, s_t, j_t\right)+\beta u\left(c_t, \kappa h_{t+1}, j_{t+1}\right)\\
&c_t,\ c_{t+1},\ \tau,\ a_{t+1},\ j_{t+1} &&\\
&\text{subject to} && P\left(j_t\right) c_t+P_t^H\ \tau+a_{t+1}=\omega\left(j_t\right)+P_t^H\ h_t\\
& && P\left(j_{t+1}\right) c_{t+1}=\omega\left(j_{t+1}\right)+R_{t+1} a_{t+1}+P_{t+1}^H\ h_{t+1}\\
& && h_{t+1}=(1-\delta) h_t+\tau\\
& && a_{t+1} \geqslant -(1-\theta) h_{t+1}\ P_t^H\\
& && c_t, c_{t+1}, h_{t+1} \geqslant 0\\
& && \rho_{t+1}=0
\end{aligned} \tag{6}$$

where $\beta \in (0,1)$ is the discount factor, $\delta \in (0,1)$ is the depreciation of housing stock; $P(j_t)$ is the non-durable prices observed in city j, $\omega(j)$ is the wages observed in city j, P_t^H is the observed house prices at time t, P_{t+1}^H is the expected housing prices at time t+1; h_t and h_{t+1} are housing stocks, T is the upper bound to housing purchases in each city.

Similarly, the renter-households' utility can be described as

util (renters):

$$\begin{aligned}
&\text{maximize} && u\left(c_t, s_t, j_t\right)+\beta u\left(c_t, \rho_{t+1}, j_{t+1}\right)\\
&c_t,\ c_{t+1},\ \tau,\ a_{t+1},\ j_{t+1} &&\\
&\text{subject to} && P\left(j_t\right) c_t+P_t^H\ \tau+a_{t+1}=\omega\left(j_t\right)+P_t^H\ h_t\\
& && P\left(j_{t+1}\right) c_{t+1}=\omega\left(j_{t+1}\right)+R_{t+1} a_{t+1}+P_{t+1}^H\ h_{t+1}\\
& && h_{t+1}=(1-\delta) h_t+\tau\\
& && a_{t+1} \geqslant -(1-\theta) h_{t+1}\ P_t^H\\
& && c_t, c_{t+1}, h_{t+1} \geqslant 0\\
& && \rho_{t+1} \neq 0
\end{aligned} \tag{7}$$

3.4. Market Clearing

For each period, the market clearing price vector, $\pi_t = (P_t^*, P_t^{H*}, R_{t+1}^*, \omega_t^*)$, is composed of prices for non-durable goods (at origin city), house prices, the national interest rate, and wages for different cities. Each period, the market clearing price vector can be identified by the clearing of all the markets in the nation. Specifically, given the market clearing price vector π_t, a measure $\left\{\mu_{j,t}^i\right\}$ of households in City j, the measure of households in the nation $\left\{M_t^i\right\} = \{U_{j=1}^J \mu_{j,t}^i\}$,

(1) The non-durable goods markets clear at prices π_t, specifically P_t

$$\int c^i_{j,t} d\mu^i_{j,t} = Y_{j,t} = \Psi_j \left(\xi^y_{j,t}\right)^\alpha \left(l^y_{j,t}\right)^{1-\alpha} \qquad \forall j = \{1, 2, \ldots, J\}$$

(2) The housing markets clear at prices π_t, specifically P^{H*}_t

$$\int \tau^i_{j,t} d\mu^i_{j,t} + \int \rho^i_{j,t} d\mu^i_{j,t} = H_{j,t} = \left(\xi^h_{j,t}\right)^\beta \left(l^h_{j,t}\right)^{1-\beta} \qquad \forall j = \{1, 2, \ldots, J\}$$

(3) The national assets market clears at prices π_t, specifically R_{t+1}

$$\int a^i_{t+1} dM^i_t = 0$$

(4) The labor markets clear at prices π_t, specifically ω^*_t

$$N_{j,t} = \xi^y_{j,t} + \xi^h_{j,t} \qquad \forall j = \{1, 2, \ldots, J\}$$

4. Simplified Theory for Intuition

In this section, the paper provides a simplified version of the generalized model to underscore the key insights. Everything follows through from the generalized model, but the paper makes the following simplifying assumptions:

(1) There are only two cities: City *A* and City *B*.

(2) There is a measure one of individuals living in City *A*, and a greater measure of people living in City *B*.

(3) Market friction in the rental market is infinitely high.[1]

(4) The preference for housing services is universal for all households set at $\zeta_H > 0$.

(5) Migration cost between the two cities is symmetric and fixed at *m*.

(6) Individuals in City *A* have heterogeneous expectations with regard to house prices in the two cities:[2]

For individual *i*, $\frac{P^{i,H}_A}{P^{i,H}_A}, \frac{P^{i,H}_B}{P^{i,H}_B} \quad Unif\left(0, 2\right)$

(7) Prices in the non-durable goods market are fixed, trade costs are well-defined.

The price vector for non-durable Good *A* is $P_A = \left(P_{AA}, P_{AB}\right)$, where the first scalar is the price of Good *A* observed in City *A*, and the second scalar is the price of Good *A*

[1] As Zhao (2013) proved in his paper, "as long as the rental market friction is high enough," a bubble economy will exist where house prices exceed its fundamental values. Thus, it is safe to say that assuming away the rental markets will not alter our understanding so much that it changes the direction of our analysis.

[2] To simplify notation, I drop the time subscripts when doing so does not raise excessive confusion.

observed in City B (including trade costs) in the presence of trade costs. It follows that $P_{AB} \geqslant P_{AA}$.

Similarly, non-durable Good B is $P_B = (P_{BA}, P_{BB})$, and $P_{BA} \geqslant P_{BB}$.

(8) The interest rate is determined in a larger market and taken as exogenous to this model as R_{t+1}.

(9) The agglomeration effect is still present and is manifested through wages. Specifically, $\omega_{j,t} = K \times \Psi_{j,t} \quad \forall j = \{A, B\}, K > 0$

From Equation (1), we see that

$$\frac{\partial \omega_{j,t}}{\partial N_{j,t}} = Kz\sigma N_{j,t}^{\sigma-1} > 0 \tag{8}$$

(10) Everyone living in City A are endowed with zero housing units; everyone living in City B are endowed with positive amounts of housing units such that given their preferences and expectations, they do not have incentive to migrate to City A.[1]

(11) There is no upper bound to housing purchases in either of the cities (i.e. $T_j \to \infty, j = \{A, B\}$).

(12) There is an exogenous housing stock of H_A and H_B created by the real-estate industry in each of the Cities in each period.

Under this setting, all households are homeowners and the objective function for the household problem of City A can be rewritten as the following:

$$V = \underset{c_t, c_{t+1}, \tau, a_{t+1}, j_{t+1}}{\text{maximize}} \; u\left(c_t, h_t, j_t = A\right) + \beta u\left(c_{t+1}, h_{t+1}, j_{t+1}\right)$$

And since the next period location decision is a finite discrete choice, the problem can be further simplified:

$$V = \max\{U(A), U(B)\}$$

where

$$\textit{Non-migratory}: U(A) = \underset{c_t, c_{t+1}, \tau, a_{t+1}}{\text{maximize}} \; u\left(c_t, h_t, j_t = A\right) + \beta u\left(c_{t+1}, h_{t+1}, j_{t+1} = A\right)$$

$$\textit{Migratory}: U(B) = \underset{c_t, c_{t+1}, \tau, a_{t+1}}{\text{maximize}} \; u\left(c_t, h_t, j_t = A\right) + \beta u\left(c_{t+1}, h_{t+1}, j_{t+1} = B\right)$$

The non-migratory household problem can be characterized as the following maximization problem:

[1] The purpose of this assumption suppresses bilateral migration. In the case of bilateral migration, it complicates the point of interest in the paper by requiring us to analyze the net effect of migration. In the analytical setting described above, this will obfuscate the matter with no additional insights.

$$\underset{c_t, c_{t+1}, \tau, a_{t+1}}{\text{maximize}} \quad \zeta_A \ln\left(c_t^A\right) + \zeta_B \ln\left(c_t^B\right) + \beta\left[\zeta_A \ln\left(c_{t+1}^A\right) + \zeta_B \ln\left(c_{t+1}^B\right) + \zeta_H \ln\left(\tau_A\right)\right]$$

$$\text{subject to} \quad \begin{aligned} & P_{AA} c_t^A + P_{BA} c_t^B + P_A^H \tau_A + P_B^H \tau_B + a_{t+1} = \omega_A \\ & P_A c_{t+1}^A + P_B c_{t+1}^B = \omega_A + P_A^H \tau_A + P_B^H \tau_B + R_{t+1} a_{t+1} \\ & a_{t+1} \geqslant -(1-\theta)\left[P_A^H \tau_A + P_B^H \tau_B\right] \\ & c_t^A, c_{t+1}^A, c_t^B, c_{t+1}^B, \tau_A, \tau_B \geqslant 0 \end{aligned} \tag{9}$$

Where all prices that have tildes (~) are expected prices of the next period. By construction, for the non-migratory households, $P_A = P_{AA}$, $P_B = P_{BA}$, $\omega_A = \omega_A$.

And similarly, migratory household problem can be characterized by the following maximization problem:

$$\underset{c_t,\ c_{t+1},\ \tau,\ a_{t+1}}{\text{maximize}} \quad \zeta_A \ln\left(c_t^A\right) + \zeta_B \ln\left(c_t^B\right) + \beta\left[\zeta_A \ln\left(c_{t+1}^A\right) + \zeta_B \ln\left(c_{t+1}^B\right) + \zeta_H \ln\left(\tau_B\right)\right]$$

$$\text{subject to} \quad \begin{aligned} & P_{AA} c_t^A + P_{BA} c_t^B + P_A^H \tau_A + P_B^H \tau_B + a_{t+1} + m = \omega_A \\ & P_A c_{t+1}^A + P_B c_{t+1}^B = \omega_B + P_A^H \tau_A + P_B^H \tau_B + R_{t+1} a_{t+1} \\ & a_{t+1} \geqslant -(1-\theta)\left[P_A^H \tau_A + P_B^H \tau_B\right] \\ & c_t^A, c_{t+1}^A, c_t^B, c_{t+1}^B, \tau_A, \tau_B \geqslant 0 \end{aligned} \tag{10}$$

By construction, for the migratory households $P_A = P_{AB}$, $P_B = P_{BB}$, $\omega_B = \omega_B$.

After solving for Equations (9) and (10), one would arrive at propositions 1 and 2, respectively. They identify the optimal decisions of households, given migratory decisions.

Proposition 1. Given observed costs and prices P_A, P_B, ω_A, and individual expectation, there are four types of non-migratory agents and their optimal decisions are:[1]

(1) Unconstrained real-estate investors (denote this group as AA-I-U)[2]

$$c_t^A = \frac{\omega_A}{P_{AA}}\left(1 + \frac{1}{R_{t+1}}\right)\left(\frac{\zeta_A}{\zeta_A + \zeta_B + \beta}\right) \qquad c_t^B = \frac{\omega_A}{P_{BA}}\left(1 + \frac{1}{R_{t+1}}\right)\left(\frac{\zeta_B}{\zeta_A + \zeta_B + \beta}\right)$$

$$c_{t+1}^A = \frac{\beta\omega_A}{P_{AA}}\left(1 + R_{t+1}\right)\left(\frac{\zeta_A}{\zeta_A + \zeta_B + \beta}\right) \qquad c_{t+1}^B = \frac{\beta\omega_A}{P_{BA}}\left(1 + R_{t+1}\right)\left(\frac{\zeta_B}{\zeta_A + \zeta_B + \beta}\right)$$

$$\tau_A = \frac{\beta\omega_A}{P_A^H}\left(\frac{1 + R_{t+1}}{R_{t+1} - \dfrac{P_A^H}{P_A^H}}\right)\left(\frac{\zeta_H}{\zeta_A + \zeta_B + \beta}\right)$$

---

[1] Please contact the author for the Proofs Appendix.

[2] "AA" corresponds to locations in the two periods; "I" corresponds to "Investors"; "U" Corresponds to "Unconstrained".

$$\tau_B = \frac{\omega_A}{P_B^H}\left[1-\left(1+\frac{1}{R_{t+1}}\right)\left(\frac{\zeta_A+\zeta_B}{\zeta_A+\zeta_B+\beta}\right)-\frac{\beta\left(1+R_{t+1}\right)}{\left(R_{t+1}-\frac{P_A^H}{P_A^H}\right)}\left(\frac{\zeta_H}{\zeta_A+\zeta_B+\beta}\right)\right]-\frac{a_{t+1}}{P_B^H}$$

$$a_{t+1} > -\left(1-\theta\right)P_A^H\tau_A-\left(1-\theta\right)P_B^H\tau_B$$

(2) Constrained real-estate investors (denote this group as AA-I-C)

$$c_t^A = \frac{\omega_A}{P_{AA}}\frac{\zeta_A\left(\gamma_B-\gamma_A\right)}{\beta g\zeta_H}\left(\frac{1}{\gamma_B}+\frac{G}{P_B^H}\right) \qquad c_t^B = \frac{\omega_A}{P_{BA}}\frac{\zeta_B\left(\gamma_B-\gamma_A\right)}{\beta g\zeta_H}\left(\frac{1}{\gamma_B}+\frac{G}{P_B^H}\right)$$

$$c_{t+1}^A = \frac{\omega_A}{P_{AA}}\frac{\zeta_A\left(\gamma_B-\gamma_A\right)}{g\zeta_H}\left(1+\frac{\gamma_B G}{P_B^H}\right) \qquad c_{t+1}^A = \frac{\omega_A}{P_{BA}}\frac{\zeta_B\left(\gamma_B-\gamma_A\right)}{g\zeta_H}\left(1+\frac{\gamma_B G}{P_B^H}\right)$$

$$\tau_A = \frac{\omega_A}{P_A^H\theta g}\left(1+\gamma_B G\right) \qquad \tau_B = \frac{\omega_A}{P_B^H\theta}G \qquad a_{t+1} = \frac{-\left(1-\theta\right)\omega_A}{\theta}\left(\frac{1+\gamma_B G}{g}+G\right)$$

(3) Unconstrained homeowners (denote this group as AA-H-U)

$$c_t^A = \left(\frac{\zeta_A}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_A\left(1+R_{t+1}\right)}{P_{AA}R_{t+1}} \qquad c_t^B = \left(\frac{\zeta_B}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_A\left(1+R_{t+1}\right)}{P_{BA}R_{t+1}}$$

$$c_{t+1}^A = \left(\frac{\beta\zeta_A}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_A\left(1+R_{t+1}\right)}{P_{AA}} \qquad c_{t+1}^B = \left(\frac{\beta\zeta_B}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_A\left(1+R_{t+1}\right)}{P_{BA}}$$

$$\tau_A = \left(\frac{\beta\zeta_H}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_A\left(1+R_{t+1}\right)}{P_A^H\left[R_{t+1}-\frac{P_A^H}{P_A^H}\right]} \qquad \tau_B = 0$$

$$a_{t+1} = \omega_A\left[1-\frac{\left(1+R_{t+1}\right)}{\zeta_A+\zeta_B+\beta}\left[\frac{\zeta_A+\zeta_B}{R_{t+1}}+\frac{\beta\zeta_H}{R_{t+1}-\left(\frac{P_A^H}{P_A^H}\right)}\right]\right]$$

(4) Constrained homeowners (denote this group as AA-H-C)

$$c_t^A = \frac{\zeta_A}{\left(\zeta_A+\zeta_B\right)P_{AA}}\left[\omega_A-\frac{\omega_A\left[\beta\gamma_A-\left(\beta+\zeta_A+\zeta_B-\beta\left(\zeta_A+\zeta_B\right)\right)\right]+\Phi_A}{2\gamma_A\left(\beta+\zeta_A+\zeta_B\right)}\right]$$

$$c_t^B = \frac{\zeta_B}{\left(\zeta_A+\zeta_B\right)P_{BA}}\left[\omega_A-\frac{\omega_A\left[\beta\gamma_A-\left(\beta+\zeta_A+\zeta_B-\beta\left(\zeta_A+\zeta_B\right)\right)\right]+\Phi_A}{2\gamma_A\left(\beta+\zeta_A+\zeta_B\right)}\right]$$

$$c_{t+1}^A = \frac{\zeta_A}{\left(\zeta_A+\zeta_B\right)P_{AA}}\left[\omega_A+\frac{\omega_A\left[\beta\gamma_A-\left(\beta+\zeta_A+\zeta_B-\beta\left(\zeta_A+\zeta_B\right)\right)\right]+\Phi_A}{2\left(\beta+\zeta_A+\zeta_B\right)}\right]$$

$$c_{t+1}^B = \frac{\zeta_B}{\left(\zeta_A+\zeta_B\right)P_{BA}}\left[\omega_A+\frac{\omega_A\left[\beta\gamma_A-\left(\beta+\zeta_A+\zeta_B-\beta\left(\zeta_A+\zeta_B\right)\right)\right]+\Phi_A}{2\left(\beta+\zeta_A+\zeta_B\right)}\right]$$

$$\tau_A = \frac{\omega_A \left[\beta\gamma_A - \left(\beta + \zeta_A + \zeta_B - \beta\left(\zeta_A + \zeta_B\right)\right)\right] + \Phi_A}{2\theta\gamma_A\left(\beta + \zeta_A + \zeta_B\right)P_A^H}$$

$$\tau_B = 0$$

$$a_{t+1} = -\left(1-\theta\right)\frac{\omega_A \left[\beta\gamma_A - \left(\beta + \zeta_A + \zeta_B - \beta\left(\zeta_A + \zeta_B\right)\right)\right] + \Phi_A}{2\theta\gamma_A\left(\beta + \zeta_A + \zeta_B\right)}$$

where

$$\gamma_A = \frac{P_A^H - \left(1-\theta\right)P_A^H R_{t+1}}{\theta P_A^H} \qquad \gamma_B = \frac{P_B^H - \left(1-\theta\right)P_B^H R_{t+1}}{\theta P_B^H}$$

$$g = \frac{\left(1-\zeta_H\right)\gamma_B - \gamma_A}{\zeta_H} \qquad G = \frac{\left[g - \frac{\left(\zeta_A + \zeta_B\right)\left(\gamma_B - \gamma_A\right)}{\zeta_H\beta\gamma_B} - 1\right]}{\left[g + \frac{\left(\zeta_A + \zeta_B\right)\left(\gamma_B - \gamma_A\right)}{\zeta_H\beta} + \gamma_B\right]}$$

$$\Phi_A = \sqrt{\omega_A^2\left[\left(\beta + \zeta_A + \zeta_B - \beta\left(\zeta_A + \zeta_B\right)\right) - \beta\gamma_A\right]^2 + 4\beta\gamma_A\omega_A^2\left[\zeta_H\left(\beta + \zeta_A + \zeta_B\right)\right]}$$

Proposition 2. Given observed costs and prices $P_A, P_B, \omega_A, \omega_B$ and individual expectations, there are four types of migratory agents and their optimal decisions are:

(1) Unconstrained real-estate investors (denote this group as AB-I-U)

$$c_t^A = \left(\frac{\omega_A}{P_{AA}} + \frac{\omega_B}{P_{AA}R_{t+1}} - \frac{m}{P_{AA}}\right)\left(\frac{\zeta_A}{\zeta_A + \zeta_B + \beta}\right) \qquad c_t^B = \left(\frac{\omega_A}{P_{BA}} + \frac{\omega_B}{P_{BA}R_{t+1}} - \frac{m}{P_{AA}}\right)\left(\frac{\zeta_B}{\zeta_A + \zeta_B + \beta}\right)$$

$$c_{t+1}^A = \left(\frac{\omega_A R_{t+1}}{P_{AB}} + \frac{\omega_B}{P_{AB}} - \frac{m}{P_{AA}}\right)\left(\frac{\beta\zeta_A}{\zeta_A + \zeta_B + \beta}\right) \qquad c_{t+1}^B = \left(\frac{\omega_A R_{t+1}}{P_{BB}} + \frac{\omega_B}{P_{BB}} - \frac{m}{P_{AA}}\right)\left(\frac{\beta\zeta_B}{\zeta_A + \zeta_B + \beta}\right)$$

$$\tau_A = \frac{\omega_A}{P_A^H} - \left(\omega_A + \frac{\omega_B}{R_{t+1}}\right)\left(\frac{\zeta_A + \zeta_B}{P_A^H\left(\zeta_A + \zeta_B + \beta\right)}\right) - \left(\frac{\omega_A R_{t+1} + \omega_B}{R_{t+1} - \frac{P_A^H}{P_A^H}}\right)\left(\frac{\beta\zeta_H}{P_A^H\left(\zeta_A + \zeta_B + \beta\right)}\right) - \frac{a_{t+1} + m}{P_A^H}$$

$$\tau_B = \frac{\left(\omega_A R_{t+1} + \omega_B - mR_{t+1}\right)}{P_B^H\left(R_{t+1} - \frac{P_B^H}{P_B^H}\right)}\left(\frac{\beta\zeta_H}{\zeta_A + \zeta_B + \beta}\right) \qquad a_{t+1} = -\left(1-\theta\right)P_A^H\tau_A - \left(1-\theta\right)P_B^H\tau_B$$

(2) Constrained real-estate investors (denote this group as AB-I-C)

$$c_t^A = \frac{\zeta_A}{P_{AA}}\frac{\omega_B\left(\gamma_A - \gamma_B\right)}{\zeta_H f\beta}\left(\frac{1}{\gamma_A} + \frac{F}{P_A^H}\right) \qquad c_t^B = \frac{\zeta_B}{P_{BA}}\frac{\omega_B\left(\gamma_A - \gamma_B\right)}{\zeta_H f\beta}\left(\frac{1}{\gamma_A} + \frac{F}{P_A^H}\right)$$

$$c_{t+1}^A = \frac{\zeta_A}{P_{AB}}\frac{\omega_B\left(\gamma_A - \gamma_B\right)}{\zeta_H f}\left(1 + \frac{\gamma_A F}{P_A^H}\right) \qquad c_{t+1}^B = \frac{\zeta_B}{P_{BB}}\frac{\omega_B\left(\gamma_A - \gamma_B\right)}{\zeta_H f}\left(1 + \frac{\gamma_A F}{P_A^H}\right)$$

$$\tau_A = \frac{\omega_B}{P_A^H\theta}F \qquad \tau_B = \frac{\omega_B}{P_B^H\theta f}\left(1 + \gamma_A F\right) \qquad a_{t+1} = \frac{-\left(1-\theta\right)\omega_B}{\theta}\left(\frac{1 + \gamma_A F}{f} + f\right)$$

(3) Unconstrained homeowners (denote this group as AB-H-U)

$$c_t^A = \left(\frac{\zeta_A}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_B+\omega_A R_{t+1}-mR_{t+1}}{P_{AA}R_{t+1}} \qquad c_t^B = \left(\frac{\zeta_B}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_B+\omega_A R_{t+1}-mR_{t+1}}{P_{BA}R_{t+1}}$$

$$c_{t+1}^A = \left(\frac{\beta\zeta_A}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_B+\omega_A R_{t+1}-mR_{t+1}}{P_{AB}} \qquad c_{t+1}^B = \left(\frac{\beta\zeta_B}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_B+\omega_A R_{t+1}-mR_{t+1}}{P_{BB}}$$

$$\tau_A = 0 \qquad \tau_B = \left(\frac{\beta\zeta_H}{\zeta_A+\zeta_B+\beta}\right)\frac{\omega_B+\omega_A R_{t+1}-mR_{t+1}}{P_B^H\left[R_{t+1}-\frac{P_B^H}{P_B^H}\right]}$$

$$a_{t+1} = \omega_A - \frac{\omega_B+\omega_A R_{t+1}-mR_{t+1}}{\zeta_A+\zeta_B+\beta}\left[\frac{\zeta_A+\zeta_B}{R_{t+1}}+\frac{\beta\zeta_H}{R_{t+1}-\frac{P_B^H}{P_B^H}}\right]$$

(4) Constrained homeowners (denote this group as AB-H-C)

$$c_t^A = \frac{\zeta_A}{(\zeta_A+\zeta_B)P_{AA}}\left[(\omega_A-m)-\frac{\beta(\omega_A-m)\gamma_B-\omega_B(\beta+\zeta_A+\zeta_B-\beta(\zeta_A+\zeta_B)+\Phi_B}{2\gamma_B(\beta+\zeta_A+\zeta_B)}\right]$$

$$c_t^B = \frac{\zeta_B}{(\zeta_A+\zeta_B)P_{BA}}\left[(\omega_A-m)-\frac{\beta(\omega_A-m)\gamma_B-\omega_B(\beta+\zeta_A+\zeta_B-\beta(\zeta_A+\zeta_B)+\Phi_B}{2\gamma_B(\beta+\zeta_A+\zeta_B)}\right]$$

$$c_{t+1}^A = \frac{\zeta_A}{(\zeta_A+\zeta_B)P_{AB}}\left[(\omega_A-m)-\frac{\beta(\omega_A-m)\gamma_B-\omega_B(\beta+\zeta_A+\zeta_B-\beta(\zeta_A+\zeta_B)+\Phi_B}{2\gamma_B(\beta+\zeta_A+\zeta_B)}\right]$$

$$c_{t+1}^B = \frac{\zeta_B}{(\zeta_A+\zeta_B)P_{BB}}\left[(\omega_A-m)-\frac{\beta(\omega_A-m)\gamma_B-\omega_B(\beta+\zeta_A+\zeta_B-\beta(\zeta_A+\zeta_B)+\Phi_B}{2\gamma_B(\beta+\zeta_A+\zeta_B)}\right]$$

$$\tau_A = 0$$

$$\tau_B = \frac{\beta(\omega_A-m)\gamma_B-\omega_B(\beta+\zeta_A+\zeta_B-\beta(\zeta_A+\zeta_B)+\Phi_B}{2\gamma_B(\beta+\zeta_A+\zeta_B)P_B^H}$$

$$a_{t+1} = -(1-\theta)\frac{\beta(\omega_A-m)\gamma_B-\omega_B(\beta+\zeta_A+\zeta_B-\beta(\zeta_A+\zeta_B)+\Phi_B}{2\gamma_B(\beta+\zeta_A+\zeta_B)}$$

where

$$\gamma_A = \frac{P_A^H-(1-\theta)P_A^H R_{t+1}}{\theta P_A^H} \qquad \gamma_B = \frac{P_B^H-(1-\theta)P_B^H R_{t+1}}{\theta P_B^H} \qquad f = \frac{(1-\zeta_H)\gamma_A-\gamma_B}{\zeta_H}$$

$$F = \frac{\left[f\left(\frac{\omega_A-m}{\omega_B}\right)-\frac{(\zeta_A+\zeta_B)(\gamma_B-\gamma_A)}{\zeta_H\beta\gamma_A}-1\right]}{\left[f+\frac{(\zeta_A+\zeta_B)(\gamma_B-\gamma_A)}{\zeta_H\beta}+\gamma_{AA}\right]}$$

$$\Phi_A = \sqrt{\omega_A^2\left[\left(\beta+\zeta_A+\zeta_B-\beta\left(\zeta_A+\zeta_B\right)\right)-\beta\gamma_A\right]^2+4\beta\gamma_A\omega_A^2\left[\zeta_H\left(\beta+\zeta_A+\zeta_B\right)\right]}$$

Definition 1. Given a measure of population with a distribution of house price expectations living in City A $\{\mu^i\}$, the equilibrium at time t consists of a vector of prices $\{R_{t+1}, P^H_{A,t}, P^H_{B,t}\}$ and allocations $\left\{c^A_t, c^B_t, c^A_{t+1}, c^B_{t+1}, \tau_A, \tau_B, a_{t+1}\right\}$ such that:

(1) The allocations satisfy proposition 1 and proposition 2.

(2) The housing markets clear: $\int \tau^i_A d\mu^i_t = H_A$ and $\int \tau^i_B d\mu^i_t = H_B$

After solving Equations (9) and (10), one could separate the population and their optimal decisions based on their expectations for future house prices. The five types of households are summarized in Proposition 3.

Proposition 3. Given the market clearing interest rate (R_t+1), non-migratory and migratory household-types are determined jointly by the borrowing constraint (θ) and their expectation of house prices: the following summarizes the separations for different household types.

Table 2. The Separations for Different Household Types

Types	Expectations	Non-migratory H.H.	Migratory H.H
Type I	$\frac{P^H_A}{P^H_A} < \frac{P^H_B}{P^H_B} = R_{t+1}$	AA-I-U	AB-H-C
Type II	$\frac{P^H_A}{P^H_A} < R_{t+1} \leqslant \frac{P^H_B}{P^H_B}$	AA-I-C	AB-H-C
Type III	$\frac{P^H_A}{P^H_A} < R_{t+1}$ $\frac{P^H_B}{P^H_B} < R_{t+1}$	AA-H-U or AA-H-C	AB-H-U or AB-H-C
Type IV	$\frac{P^H_A}{P^H_A} = R_{t+1} > \frac{P^H_B}{P^H_B}$	AA-H-C	AB-I-U
Type V	$\frac{P^H_A}{P^H_A} \geqslant R_{t+1} > \frac{P^H_B}{P^H_B}$	AA-H-C	AB-I-C

Proposition 4. Let $\Omega = \omega_B / (\omega_A - m)$ denote wage gaps. Under the strict assumption of identical goods in the two cities, there are thresholds of wage gaps (Ω_1, Ω_2) such that whenever:

(1) $\hat{\Omega} < \Omega_1$: among the migrating households from City A to City B, one would only observe households of Expectation Types I and II.

(2) $\hat{\Omega} < \Omega_2$: among the migrating households from City *A* to City *B*, one may observe households of Expectation Types I, II, and III.

(3) $\hat{\Omega} > \Omega_2$: among the migrating households from City *A* to City *B*, one would observe households from any Expectation Types.

Proposition 5. At the same price levels, Type I and II households will demand a greater amount of City *B* housing τ_B, should they choose to migrate. At the same price levels, Type IV and V households will demand a lesser amount of City *A* housing.

Propositions 4 and 5 provide a very intuitive understanding to the problem. From Proposition 4, we see that at a fixed wage gap, as more households believe that the return from investing in City *B* housing is superior to other investment options, there will be a greater amount of migration towards City *B*. And when migration does occur, from Proposition 5, the migrating families will either demand more of City *B* housing or less of City *A* housing. This migration dynamics will change prices to reinforce the optimism in City *B* housing returns. In fact, Proposition 4 illustrates the fact that this said optimism can also be generated through market mechanisms described in the model. In the most conservative assumption, Type V investors dominate the economy (i.e. households believe that the return of City *A* housing is superior). Yet, as Proposition 4 suggests, if the wage gap between City *A* and City *B* are sufficiently large, the same migratory patterns and house price changes can be observed. As a result of updating priors, households become a little more optimistic about City *B* housing.

Proposition 6. Under the prescribed conditions of the model, at any time period, those already dwelling in City *B* will not have an incentive to migrate to City *A*.

In conclusion of Propositions 1−6, by studying the simplified model and generalizing the assumptions, the following conclusions can be reached.

(1) House prices in are affected by expectations for future prices.

(2) Given the economic environment and individual endowments, household expectations of future house prices significantly influence migratory decisions.

(3) Against the Chinese socioeconomic backdrop, the impact of investor-driven on house price hike is small.

5. Empirical Verification

5.1. Data Source

The data used in the reduced form analysis come from two databases: WIND Financial Terminal and the CEIC Database. Both data platforms gather publicly available data from official government publications. Finally, house supply (price) elasticities are taken from Wang *et al*. (2012). The following table (Table 3) presents a brief summary of data sources.

Table 3. Brief Summary of Data and Sources

	Cities	Time span	Data source
GDP total	270	2010−17	CIEC
GDP second sect	270	2010−17	CIEC
City wages	270	2010−17	CIEC
House price	270	2010−17	CIEC
Debt vs GDP	N.A.	2010−17	CIEC
Registered population	270	2010−17	WIND
Residential population	270	2010−17	WIND
Government revenue	270	2010−17	WIND
Supply elasticity	34	N.A.	(Wang et al., 2012)

Note: Please contact the author for the Data Appendix.

An important subtlety is the distinction between "Registered Population" (户籍人口) and "Residential Population" (常住人口). Referencing the Sixth China Population Census, "Registered Population" records the number of individuals who have their hukou registered in the city; "Residential Population" are identified as individuals who reside in a city for at least five months.

From population data, this paper defines migration for city j at time t as the detrended population growth rate: $Migration_{j,t} = \Delta Population_{j,t} - \Delta Population_{rural,t}$

where Δ denotes percent change. In other words, a city's migration is the city's change in population, subtracting the national population growth rate. Furthermore, there are two ways to measure the said migration rates, measured through either registered population or residential population. This gives rise to two definitions of migration: "Detrended Residential Population Growth" and "Detrended Registered Population Growth".

5.2. Reduced Form Analysis

In this section, this paper will first verify that the findings derived from the simplified theoretical model through a series of regression analysis. Recall that from the simplified theory, we arrived at three findings. First, migration and House Prices: House prices are affected by migratory decisions of households. Second, the Role of Expectations: Given the economic environment and individual endowments, household expectations of future house prices significantly influence migratory decisions. Third, the Role of Speculative Investments: Against the Chinese socioeconomic backdrop, the impact of investor-driven house price increase is small.

5.2.1. Migration and House Prices

Even at the intuitive level, migration must be a major component to determining house price movements—as more people migrate to a city, the demand for housing will naturally increase and prices must also increase as a result. This intuition was corroborated through the market clearing condition. To test this hypothesis, this paper employs the following identification:

$$\begin{aligned} HousePrice_{j,t+1} = \alpha_0 + \alpha_1 SupplyElasticity_j + \alpha_2 Migration_{j,t} + \alpha_3 LogPopulation_{j,t} \\ + \alpha_4 LogWages_{j,t} + \alpha_5 DebtGdpRatio_t + \in \end{aligned} \tag{11}$$

Where migration is either "Detrended Residential Population Growth" or "Detrended Registered Population Growth", defined in the previous section. The following regression table uses the identification illustrated in Equation (11); Models 1 and 2 use detrended registered population growth as proxy for migration; Model 3 and 4 use detrended resident population growth as proxy for migration (Table 4).

Table 4. Regression with Respect to Migration: Dependent Variable−House Price

Variable	Model 1	Model 2	Model 3	Model 4
Migration				
Registered	1.529 (1.004)	0.500^{*} (0.283)		
Residential			1.622^{**} (0.758)	0.115 (0.280)
Log Population				
Registered	44444.0170 (0.0280)	0.617^{***} (0.150)		
Residential			0.377^{**} (0.0305)	0.581^{***} (0.164)
Log Wages	1.461^{***} (0.0690)	0.464^{***} (0.122)	1.411^{***} (0.0798)	0.402^{***} (0.135)
Supply elasticity	-0.0336^{***} (0.0047)		-0.0317^{***} (0.0048)	
Debt GDP ratio	-2.295^{***} (0.293)	0.795* (0.418)	-2.134^{***} (0.346)	0.572 (0.476)
Year		0.0070 (0.021)		0.0167 (0.0233)
N	339	339	305	305
R^2	0.714	0.870	0.724	0.860
Treatment	Pooled	Entity F.E.	Pooled	Entity F.E.

Note: $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Immediately from Table 4, I see that wage level within a city has significant positive impact on house prices. The leverage within a city also has significant impacts on local house prices. These are in line with established theory and literature. What is interesting is to see that impact of population and migration on house prices. I see that in the entity fixed effect for residential population (Model 2), both migration and the level of population matters significantly, while this is not true for the pooled analysis (Model 1). Seeing that an individual may be registered in a city and reside elsewhere, it is very likely that he does not demand housing from the city in which he is registered. Thus, not controlling for the differences between cities, the impact of population will not affect observed house prices. On the other hand, when using residential population data, I see that the pooled analysis yields significant impacts on house price. Again, I believe this discrepancy is due to how residents are defined. Because the definition of "resident population" most closely match the notion of a city's population, I will use "residential population" data onward.

5.2.2. Role of Expectation

The next step is to investigate if and how expectations play a role in affecting house prices. As illustrated in the simplified theoretical model, household expectations play major roles in household migratory decisions. Specifically, should a household expect the house price of a city to increase substantially, then he will have more incentive to migrate there at the end of the period to enjoy a "cheaper" house. To verify that migration is indeed affected by house price expectations, I regress the following.

$$\begin{aligned} HousePrice_{j,t+1} = {} & \alpha_0 + \alpha_1 SupplyElasticity_j + \alpha_2 Migration_{j,t} \mathrm{II}_{exp} \\ & + \alpha_3 Migration_{j,t} \mathrm{II}_{gdp} + \alpha_4 Migration_{j,t} \mathrm{II}_{exp} \mathrm{II}_{gdp} \\ & + \alpha_5 LogPopulation_{j,t} + \alpha_6 LogWages + \alpha_7 DebtGdpRatio_t + \in \end{aligned} \tag{12}$$

where migration is defined solely as residential migration. The indicator function for GDP is such that

$$\mathrm{II}_{GDP} \equiv \mathrm{II}\,[GDP_{j,t} \geq First\ Qantile\{GDP_t\}_j]$$

And GDP is measured as either total GDP or Secondary Sector GDP. GDP data is used here to control for other unidentified economic features. For example, a healthy GDP growth could indicate a stronger economy in the next period and thus, higher living standards. I use both total GDP and Secondary Sector GDP because much of China's growth have stemmed from the secondary sector of the economy. Hence, separately looking at the secondary sector may yield a stronger conclusion.

The indicator function for expectations is such that

$$\mathrm{II}_{exp} \equiv \mathrm{II}\,[EXP_{j,t} \geqslant First\ Qantile\{EXP_t\}_j]$$

And expectation is simply defined as the two-period lagged growth rate of house prices.[1]

From Table 5, we see that whenever the expectation for a city's house price to be in the top 25 percentile, migration has significant (at the 1% level) impact on house prices. Other effect encapsulated by the GDP control variable is dulled out. As established earlier, supply elasticity of housing, city wages, and leverage all have significant impacts on house prices.

Table 5. Regression with Respect to Expectations: Dependent Variable−House Price

Variable	Total GDP	Secondary Sector GDP
Res. Migration	1.698 (1.596)	1.0240 (1.567)
Mig * Π_{exp}	10.233*** (2.995)	15.228*** (3.723)
Mig * Π_{GDP}	−0.225 (1.816)	0.720 (1.790)
Mig * Π_{GDP}* Π_{exp}	−10.861*** (3.233)	−15.904*** (3.919)
Log Res. Population	0.0345 (0.0891)	0.0338 (0.0304)
Log Wages	1.384*** (0.0891)	1.449*** (0.0900)
Supply elasticity	−0.030*** (0.0045)	−0.0295*** (0.0044)
Debt GDP ratio	−2.088*** (0.361)	−2.376*** (0.357)
N	305	305
R^2	0.749	0.753
Treatment	Pooled	Pooled

Note: $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

Finally, to capture the overall effect of migration on house prices, I use GDP growth rates, city wage growth rates, and house price expectation as instruments for migration levels. I obtain regression Table 5 through the Two Stage Least Squares (TSLS) identified in Equation (13).

$$\begin{aligned} HousePrice_{j,t+1} = \alpha_0 + \alpha_1 SupplyElasticity_j + \alpha_2 Migration_{j,t} + \alpha_3 LogPopulation_{j,t} \\ + \alpha_4 LogWages_{j,t} + \alpha_5 DebtGdpRatio_t + \in \end{aligned} \tag{13}$$

$$\begin{aligned} Migration_{j,t+1} = \beta_0 + \beta_1 Growth_{GDP_{j,t}} + \beta_2 Growth_{Wages_{j,t}} \\ + \beta_3 Expectation_{HousePrice_{j,t+1}} + v \end{aligned} \tag{14}$$

[1] Using the one-period lagged growth rate of house prices to proxy for expectation yields similar results.

Referring to Table 6, we see that the level of population migration is statistically significant in determining house price. Moreover, the exact level of population count in the current period has minimal impact on next period house prices. This is in line with the intuition developed from the simplified theory outlined before.

Table 6. (TSLS) Dependent Variable−House Price

Variable	Total GDP	Secondary sector GDP
Population migration	3.763* (2.126)	4.334* (2.317)
Log population	0.0384 (0.0353)	0.0375 (0.0353)
Log wage	0.773*** (0.0571)	0.773*** (0.0575)
Debt GDP ratio	−0.0351 (0.388)	0.0049 (0.394)
Supply elasticity	−0.0457*** (0.0064)	−0.0461*** (0.0064)
N	258	258
R^2	0.998	0.999

Note: $^{***}p<0.01$, $^{**}p<0.05$, $^{*}p<0.1$.

5.2.3. The Role of Speculative Investment

The last question to investigate is how big a role does housing investors play in pushing up Chinese house prices? As depicted in the theoretical model, when an economy consists of households with a skewed distribution of wealth, the effect of speculative investment on house prices is less significant. To test this hypothesis, I make use of a major financial event that could have affected investment portfolios of many Chinese households: the Chinese bull market of 2015. I treat this event as a natural experiment and study the growth-rates of house prices before and after the 2015 Bull Market.

The 2015 Bull Market provides an opportunity to delineate real-estate investors' impact on house prices. In an underdeveloped financial market with few financial derivatives, real-estate investors' thus treat stocks and housing as substitute goods. Furthermore, the trading population in China differs significantly from those elsewhere in the world. The Chinese stock market trading activity is dominated by individual investors (close to 85%). According to data from the China's Securities Depository and Clearing Corp, in the first five months of 2015, more than 30 million new accounts were opened by individual investors. Thus, households that previously have not engaged in the stock market started to participate in the 2015 Bull Market. Following this logic, when investor households began to believe the return from stocks to be higher than the return from housing, they will reallocate money used for real-estate investment to the stock market, thereby slowing the growth rates for house prices.

To formalize the analysis, after controlling for population migration, leverage, and government revenue[1] for top (8%) expected house price growers, I compare the house price residuals (using the Kolmogorov–Smirnov test) from different periods to see if the residuals are fundamentally different. Specifically, I chunk my data into three time periods: before 2015 Bull Market (2012−2014), during 2015 Bull Market (2015), after 2015 Bull Market (2016−2017).[2] Using house price data, I obtain residuals ($\in_t$) through the following identification:

$$\begin{aligned}\Delta HousePrice_{j,t} &= \alpha_0 + \alpha_1 PopulationTotal_{j,t} + \alpha_2 Migration_{j,t} + \alpha_3 \Delta Migration_{j,t} \\ &+ \alpha_4 Wages_{j,t} + \alpha_5 \Delta Wages_{j,t} + \alpha_6 \Delta DebtGdpRatio_t + \in_t\end{aligned} \tag{15}$$

Where Δ is the annual percent change. Then, yearly residuals are grouped together via the definition:

$$\in_{before} = \{\in_t \,|\, t = 2012, 2013, 2014\} \quad \in_{during} = \{\in_t \,|\, t = 2015\} \quad \in_{after} = \{\in_t \vee t = 2016, 2017\} \tag{16}$$

Table 7. Kolmogorov-Smirnov Test Results

Samples tested	Test statistic	p-Value
$\in_{before}$ vs. $\in_{during}$	0.38312	0.05560
$\in_{during}$ vs. $\in_{after}$	0.42857	0.04530
$\in_{before}$ vs. $\in_{after}$	0.13766	0.84660

The results of the analysis show that house price growth in the years before the 2015 Bull Market is significantly different from that during the Bull Market (at the 10% level); house price growth in the years after the 2015 Bull Market is significantly different from that during the Bull Market (at the 5% level); the difference between house price growths before and after the Bull Market is not statistically significant. This result suggests that the Bull Market did have a significant impact on Chinese house price growth, after controlling for population levels, migratory patterns, wages, and leverage. This finding is congruent with established beliefs that investors play non-trivial roles in the Chinese housing market. Yet, it should also be noted that the confidence levels of the significant results are lower than the confidence level indicating the significance of migration in affecting house prices. While the reduced form analysis cannot fully identify the magnitudes of house price movements to the specific channels (either the migratory channel or the investor channel) affecting the climbing Chinese house prices, the fact that the confidence level for the migratory is

[1] In the previous analyses, I used housing "supply elasticity", but it severely limited the amount of data I could use, as Wang *et al*. (2012) only provided estimated elasticities for 34 cities. In their analysis, government revenue was a major predictor for elasticities, so it is included here as a control variable.

[2] I also investigated the possibility that the real-estate market may be slow to react to changes in the financial market. Both analyses yield similar results.

higher than that for the investor-channel may suggest that the magnitude of the impact from migratory patterns is greater than that from investor- driven speculations.

6. Conclusion

This paper reveals importance of incorporating population movement to analyze mechanisms of Chinese housing price hikes. While dominating Chinese literature explored various impacts of financial policies and financial market environments on house prices under the assumption that speculation driven investments to be the sole dominating factor on prices, this paper constructs a tractable dynamic spatial economic framework to show that expectations for future house prices may result in some degrees of speculative investment under loose financial constraints, it too will impact household migratory decisions. Because migrating households need to have shelter at destination cities, population migration inevitably will push house prices to higher levels. Furthermore, as illustrated through reasoning stemming from the model, the rapid urbanization observed in China can be identified as a channel through which market exuberance is initiated and sustained.

The empirical analysis presented in the paper corroborate the findings from the theoretical reasoning. It establishes that migration decisions of households are non-trivial influences on local house prices, household expectation of house prices affect migratory decisions, and lead to some levels of speculative investment.

In the current state, this paper provides a new angle to explore in-depth analysis of the complicated intricacies of factors generating the Chinese housing boom. A natural extension of the analysis is to conduct a structural estimation to identify relative magnitudes of the different channels through which local house prices are influenced. For example, through a calibrated structural model, one could identify whether a "bubble" exist in the Chinese real estate market. Furthermore, should a bubble exist, a calibrated model account for spatial dynamics between cities can yield constructive policy guidance for government at both the national and municipal levels. A key to preventing bubble-burst events in the housing market, like those of Japan in the 1980s and USA in 2007, is to understand how real-estate prices climb to such gargantuan levels. This paper suggests that China's inflated house prices is not just the results of high financial leverage, but also, and perhaps more important, the consequences of rapid urbanization and household migrations to megacities in search for better lives.

References

Acemoglu, D., Chernozhukov, V., & Yildiz, M. (2006). Learning and Disagreement in an Uncertain World. NBER Working Papers 12648.

Berger, D., Guerrieri, V., Lorenzoni, G., & Vavra, J. (2018). House Prices and Consumer Spending. *Review of Economic Studies*, 85(3), 1502–1542.

Bilal, A., & Rossi-Hansberg, E. (2018). Location as an Asset. Technical Report.

Burnside, C., Eichenbaum, M., & Rebelo, S. (2016). Understanding Booms and Busts in Housing Markets. *Journal of Political Economy*, 124(4), 1088–1147.

Desmet, K., Nagy, D. K., & Rossi-Hansberg, E. (2018). The Geography of Development. *Journal of Political Economy*, 126(3), 903 – 983.

Dumas, B., Kurshev, A., & Uppal, R. (2009). Equilibrium Portfolio Strategies in the Presence of Sentiment Risk and Excess Volatility. *Journal of Finance*, 64(2), 579–629.

Gaubert, C. (2018). Firm sorting and Agglomeration. *American Economic Review*, 108(11), 3117–53.

Geanakoplos, J. (2010). The Leverage Cycle. *NBER Macroeconomics Annual 2009*, 24, 1–65.

Han, X. (2016). The Financial Leverage behind Rocketing Hose Prices. *Financial Planner(Licai Shi)*, 10, 16–17.

Harrison, J. M., & Kreps, D. (1978). Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations. *The Quarterly Journal of Economics*, 92(2), 323–336.

Kaplan, G., Mitman, K., & Violante, G. L. (2017). The Housing Boom and Bust: Model Meets Evidence. NBER Working Papers 23694.

Monte, F., Redding, S. J., & Rossi-Hansberg, E. (2018). Commuting, Migration, and Local Employment Elasticities. *American Economic Review*, 108(12), 3855–3890.

Scheinkman, J., & Xiong, W. (2003). Overconfidence and Speculative Bubbles. *Journal of Political Economy*, 111, 1183–1219.

Wang, S., Chan, S. H., & Xu, B. (2012). The Estimation and Determinants of the Price Elasticity of Housing Supply: Evidence from China. *Journal of Real Estate Research*, 34(3), 311–344.

Woetzel, J., Mendonca, L., Devan, J., Negri, S., Hu, Y., Jordan, L., Li, X., Maasry, A., Tsen, G., Yu, F., & *et al*. (2009). *Preparing for China's Urban Billion*. Mckinsey & Company.

Wu, J. (2018). *House Prices and Spending: Crowding out effect vs. wealth effect*. Ph. D. Thesis, University of International Business and Economics.

Xiao, Y., Song, Y., & Wu, X. (2018). How Far Has China's Urbanization Gone? *Sustainability*, 10(8), 1–17.

Xuan, F. (2018). Misallocation in the Chinese Land Market. Job Market Paper.

Zhao, B. (2013). Rational Housing Bubble. NBER Working Paper 19354.