Getting Rich in China: an Empirical and Structural Investigation of Wealth Mobility

Vincenzo Quadrini*, Qi Sun†, Junjie Xia‡, Ting Zeng§

August 2, 2023

Abstract

We study the properties of individual wealth growth and mobility in China using the China Household Finance Survey (CHFS). Our findings reveal that capital gains are the most significant factor contributing to wealth mobility, while individual savings play a minor role. A second finding is that housing wealth plays an important role in wealth mobility due to its high share in household portfolios and the significant cross-sectional dispersion in housing capital gains. The third finding is that wealth mobility is positively associated with households' debt. To further analyze the significance of these empirical features, we construct a general equilibrium model that we use to explore the implications of financial development and policies on wealth distribution and mobility.

*University of Southern California, CEPR and NBER, e-mail: quadrini@usc.edu
†Shanghai University of Finance and Economics, e-mail: sunqi@sufe.edu.cn
‡Central University of Finance and Economics and Peking University, e-mail: junjiexia@nsd.pku.edu.cn
§Southwestern University of Finance and Economics, e-mail: zeng-ting@hotmail.com

Introduction

It is well known that wealth is highly concentrated, much more concentrated than earnings, income, and consumption. However, the properties of wealth mobility—i.e., the change in individual wealth over time—are less well-known. This is because longitudinal data that tracks individual assets over time is more limited than cross-sectional data. As a result, many empirical studies limit the analysis to cross-sectional comparisons which, unfortunately, do not inform us about the movement of individual households within the distribution of wealth. In this paper, we take advantage of the longitudinal features of the China Household Finance Survey (CHFS) to study the wealth mobility properties of Chinese households.

The CHFS is similar to the Survey of Consumer Finances (SCF) for the United States but with the additional longitudinal feature which is absent in the SCF. The survey is conducted every two years, starting in 2011. There are several studies that used the cross-sectional dimension of the CHFS but very few explored the longitudinal dimension of the survey and characterized wealth mobility. One exception is Zeng and Zhu (2022), who documented facts on earning, income, and wealth mobility for the upper income groups. In our study, we go a step further and characterize additional facts that are especially important for outlining some of the driving forces for wealth mobility.

The results of our empirical analysis can be summarized in three main findings. First, heterogeneity in individual savings plays a relatively minor role in explaining wealth mobility. Although households with higher rates of savings experience, on average, a higher growth rate of wealth, the heterogeneity in savings across households explains only a small fraction of the variation in individual wealth growth. Instead, the most important force underlying the heterogeneity in individual wealth growth is the large dispersion in capital gains. The large dispersion in capital gains, which implies very heterogeneous returns on assets, indicates that households' wealth is very undiversified.

The second finding is that housing wealth is very important for wealth mobility. This is due to two features of the Chinese economy. First, household wealth is mostly invested in housing, with approximately 70 percent of household wealth in the form of housing. This is much larger than in the United States and is consistent with the literature that studies the role of housing market for the Chinese economy (Chen and Wen (2017), Han,

Han, and Zhu (2018), and Jiang, Miao, and Zhang (2022)).[1] Second, there is significant cross-sectional dispersion in capital gains in housing. While the importance of housing wealth for Chinese households is well known, the large dispersion in capital gains on housing and its importance for mobility is relatively new. These empirical findings—large share of housing wealth and large dispersion in housing capital gains—further indicate that households' wealth is very undiversified in China and households portfolios are exposed to significant idiosyncratic risks.

The third finding is the importance of households' debt for wealth mobility. Households that hold more debt relatively to their assets (higher leverage) tend to experience greater volatility of wealth growth. This is intuitive since leverage increases the volatility of net worth in the same way it does for a leveraged firm. Although household borrowing is not very diffuse in China, borrowers experience greater wealth growth volatility.

The empirical findings raise several questions. If housing ownership is so risky, why do Chinese households hold so much housing wealth? Why do they hold low shares of stocks in their wealth? A point also raised by Cooper and Zhu (2017). This, of course, depends on the availability of alternative investment assets such as corporate shares, which raises a related question: how would the privatization of state-owned enterprises affect portfolio holdings? What would be the impact of privatization on wealth distribution and mobility? Another question relates to the importance of household debt: how would greater borrowing accessibility affect wealth distribution and mobility?

To address these questions, we built a general equilibrium model in which households choose three types of assets: housing, stock market, and bonds (or debt when negative). Housing carries both aggregate and idiosyncratic risk, which is significant. Stock market investment carries only aggregate risk (since it is easier to diversify stock market investments compared to housing), while bonds have no risk. The statement that stock market investments can be more easily diversified than housing investments pertains to the idiosyncratic risk, not the aggregate risk.[2]

---

[1] See also Glaeser, Huang, Ma, and Shleifer (2017) providing a detailed introduction to the characteristics of the Chinese real estate market and Zhang (2017) for an explanation of the rising value of Chinese real estate properties.

[2] Through the acquisition of shares in mutual funds, households can hold shares in many companies, whereas it is uncommon for them to hold multiple shares in houses. When comparing a stock market index to a housing market index, it is possible that the former is more volatile than the latter. However, most households do not hold an index of houses

An important feature of the model is the presence of uninsurable idiosyncratic wealth risks. The model is related to the literature that shows the importance of entrepreneurial risks (Quadrini (2000) and Cagetti and DeNardi (2006)) or, more generally, investment risks (Benhabib, Bisin, and Luo (2017)) for generating very concentrated distributions of wealth, especially at the very top of the distribution. Our model focuses on investment risk associated with housing due to the importance of housing wealth for households in China. The presence of uninsurable risks and the interaction between housing and inequality in China is also a feature of the model studied in Zhang (2016).

After calibrating the model to the Chinese economy, we conduct two experiments. In the first one, we relax the financial constraints faced by households and find that higher debt generates a significant increase in wealth inequality. It also increases mobility, but only for those with greater financial participation. In terms of macroeconomic effects, greater accessibility to credit increases capital accumulation and aggregate production.

In the second experiment, we consider a policy in which the government privatizes state-owned enterprises. This increases the shares of corporate companies that can be held by households. Effectively, this increases the size of the stock market and allows for greater households' diversification (with respect to the idiosyncratic risk). We find that this policy reduces both wealth inequality and mobility. The effects on capital accumulation and production, however, are negative. The two experiments conducted in the paper point out a trade-off between economic equality and macroeconomic performance. Greater accessibility to credit and higher public ownership of productive capital have positive macroeconomic effects. However, they are also associated with greater wealth inequality.

1 Empirical analysis

The main goal of the analysis conducted in this section is to outline some of the factors that could be important in affecting the growth rate of individual wealth. To do so we first derive an accounting expression that decomposes the growth rate of wealth at the individual level in few components (Subsection

but only one or a few houses. Some households also choose not to diversify their stock market investments, but this decision is more a matter of choice rather than a consequence of limited diversification options.

1.1). We will then use the data to explore how these components relate to some economically relevant factors (Subsection 1.2).

1.1 Accounting framework

Denote by W_t the net worth of an individual household at time t. We will also refer to W_t simply as 'wealth'. The growth rate of wealth between t and $t+1$, denoted by $g_t^W = W_{t+1}/W_t - 1$, can be decomposed as follows:

$$\begin{aligned} g_t^W &= \frac{g_t W_t}{W_t} + \frac{Y_t - C_t}{Y_t} \times \frac{Y_t}{W_t} \\ &= g_t + s_t \times r_t^W. \end{aligned} \tag{1}$$

The variable g_t is the capital gain earned on one unit of wealth, $s_t = \frac{Y_t - C_t}{Y_t}$ is the saving rate (with Y_t and C_t denoting, respectively, income and consumption), and $r_t^W = \frac{Y_t}{W_t}$ can be thought as a broad measure of the return on wealth, excluding capital gains. It is a broad measure because the return includes income from labor as if earnings were also generated by wealth.

We can further decompose the broad return on wealth—net of capital gains—into capital income return and labor income return, that is, $r_t^W = \frac{Y_t^K + Y_t^L}{W_t} = r_t^K + r_t^L$, where Y_t^K and Y_t^L are, respectively, capital and labor incomes earned by an individual household. We can then rewrite the decomposition of wealth growth as

$$g_t^W = g_t + s_t(r_t^K + r_t^L) \tag{2}$$

The empirical analysis will be based on Equations (1) and (2).

1.2 Data source

The main source of data is China Household Finance Survey (CHFS). The survey has been conducted bi-annually starting in 2011 and there are five waves available: 2011, 2013, 2015, 2017, 2019. However, since there are some consistency issues when comparing consumption in the 2019 survey to previous surveys, we do not include in our analysis the 2019 survey. A feature of the CHFS is that it samples the same households over time, which allows us to track individual wealth over time. These dynamic features are studied by linking the 2011-2013, 2013-2015, and 2015-2017 waves. Although the cross-sectional dimension of the data has been used in other studies (see, for example, Cooper and Zhu (2017) and Piketty, Yang, and Zucman (2019)), the use of the longitudinal dimension for the study of wealth mobility is fairly new. Further details about the survey are provided in Appendix A.

There are some issues related to the timing in which income and wealth are measured in the survey. Wealth and its components (assets and liabilities) are observed in the middle of the survey years, that is, 2011, 2013, 2015 and 2017. Income and consumption data, instead, are for the prior year, that is, 2010, 2012, 2014 and 2016. To circumvent this problem, we proxy income and consumption for the missing years with the average of two adjacent years. Specifically, the proxy for 2011 is the average of 2010 and 2012; for 2013 we use the average of 2012 and 2014; for 2015 we use the average of 2014 and 2016.

The statistics reported in the paper are based on the urban sample which is thought to be more accurate and contaminated by smaller measurement errors. This is especially important for the value of housing wealth. Nevertheless, the main results are similar if we include the rural sample. We do not report these extended results in the paper but are available upon request.

Another data issue is the measurement of the value of assets, especially housing. Since the value of houses is self-reported in the survey, there is no guarantee that the reported values are accurate. To mitigate this concern, we impute the value of each house by using the community-level median house price in the sample.[3] More specifically, we compute the median per-square-meter price reported in the survey for houses located in a particular community. The imputed value of a house, then, is calculated by multiplying the square meter size of the house by the median price of the community in which the house is located. The logic behind this procedure is that in China the price of a house or apartment per square meter within a community is quite similar. However, the price variation across communities can be quite large. As we will see, the analysis based on the imputed price of houses reduces somewhat the magnitude of mobility but the overall message does not changed.

[3] A residential community refers to a social living collective composed of people who reside within a certain geographical area and does not coincide with any administrative unit. Communities are managed residents' committees, which are legally mandated grassroots self-governing organizations for urban residents. The size of a community varies but typically contains thousand of households. For example, there are about 5,000 communities across 16 districts in the city of Beijing.

1.3 Wealth mobility

The analysis of mobility is typically done by constructing transition matrices. A transition matrix shows the next period distribution of households that are located today in a particular wealth class (for example, those located today in the first quintile). But ultimately, in order to move from one wealth class to the other, a household needs to experience growth in wealth. Therefore, in this study, we complement the analysis based on wealth transition matrices with the growth rate analysis based on equations (1) and (2).

Table 1 reports the decomposition numbers based on equations (1) and (2). We first sort households into 5 quintiles based on the growth rate of individual wealth (net worth). Then, for each quintile, we calculate group-level aggregate variables and use them to compute the statistics of interest. For example, for each quintile, we first calculate the group income $Y_{d,t} = \sum_{i \in d} \omega_{i,t} Y_{it}$ and group consumption $C_{d,t} = \sum_{i \in d} \omega_{i,t} C_{it}$, where $\omega_{i,t}$ is the survey weight assigned to household i. We then compute the group-level saving rate as $s_{d,t} = \frac{Y_{d,t} - C_{d,t}}{Y_{d,t}}$. A more detailed description is provided at the bottom of the table.

The table shows that there are significant differences in wealth growth among households. For example, focusing on the 2015-2017 linked waves, we see that the top quintile has a growth rate of 184.9% while the bottom quintile has a growth rate of -80.1%. A similar variation among the five groups is observed for capital gains. This already indicates that the major source of wealth growth variation comes from capital gains. In the appendix we also report the results by sorting households with their initial wealth and the average wealth of two survey years, see Tables 24 and 25.

As already mentioned, since wealth data is self-reported by households, there could be significant measurement errors, especially for the value of houses. In order to alleviate this problem, we replace the reported value of houses with the imputed value calculated as the median or mean of the reported values in the community in which the housing property is located.

The survey collects the square meter size of the property as well as the total value of the property. We can then compute the value per square meter by dividing the total value by the size of the property. This allows us to compute the median or the mean of all square meter prices reported by the survey for a particular community. We impute the value of a property by multiplying the size of the property by the median or mean value per square meter in the community where the property is located. The resulting

Table 1: Decomposition of household wealth growth. Sorting based on growth rates of net worth.

(a) 2011-2013

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	3,705	19.4%	14.4%	26.8%	18.9%	11.9%	7.0%
Quintile 1	708	-59.3%	-61.3%	18.2%	11.0%	7.3%	3.7%
Quintile 2	733	-9.1%	-12.9%	25.0%	15.3%	9.9%	5.4%
Quintile 3	745	23.2%	17.3%	32.6%	18.2%	11.5%	6.7%
Quintile 4	735	63.8%	57.3%	27.4%	23.8%	15.0%	8.8%
Quintile 5	784	216.3%	204.5%	28.3%	41.5%	24.4%	17.1%

(b) 2013-2015

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	12,851	11.8%	7.2%	24.8%	18.5%	11.5%	7.1%
Quintile 1	2,749	-65.6%	-67.4%	15.0%	12.0%	8.6%	3.4%
Quintile 2	2,520	-17.0%	-20.7%	24.5%	15.3%	9.7%	5.6%
Quintile 3	2,459	10.2%	5.6%	27.1%	17.3%	10.8%	6.5%
Quintile 4	2,542	48.8%	42.8%	27.5%	21.9%	13.2%	8.8%
Quintile 5	2,581	187.9%	177.8%	27.0%	37.1%	20.4%	16.7%

(c) 2015-2017

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	15,742	13.8%	8.0%	30.0%	19.1%	11.8%	7.3%
Quintile 1	3,111	-80.1%	-82.0%	15.9%	12.3%	8.0%	4.3%
Quintile 2	3,068	-30.7%	-34.9%	27.2%	15.7%	10.2%	5.5%
Quintile 3	2,969	8.3%	2.2%	32.0%	18.9%	11.6%	7.3%
Quintile 4	3,300	50.1%	43.1%	33.9%	20.5%	12.6%	7.8%
Quintile 5	3,294	184.9%	173.3%	34.3%	33.8%	19.5%	14.3%

Notes: All calculations on this table, including house prices, are based on self-reported values. Households are sorted into 5 quintile groups based on the growth rate of individual net worth. Group-level aggregate variables are calculated by aggregating households included in each quintile using the weight $\omega_{i,t}$ assigned by the survey to each household i. For each quintile, we first calculate the group income $Y_{d,t} = \sum_{i \in d} \omega_{i,t} Y_{i,t}$ and group consumption $C_{d,t} = \sum_{i \in d} \omega_{i,t} C_{i,t}$, and then compute the saving rate as $s_{d,t} = \frac{Y_{d,t} - C_{d,t}}{Y_{d,t}}$. Returns are calculated as $r_{d,t}^K = \sum_{i \in d} \omega_{i,t} Y_{i,t}^K / \sum_{i \in d} \omega_{i,t} W_{i,t}$ and $r_{d,t}^L = \sum_{i \in d} \omega_{i,t} Y_{i,t}^L / \sum_{i \in d} \omega_{i,t} W_{i,t}$, where $Y_{i,t}^K$ and $Y_{i,t}^L$ are, respectively, capital and labor incomes. The wealth growth rate for each deciles is calculated as $\sum_{i \in d} \omega_{i,t} W_{i,t+1} / \sum_{i \in d} \omega_{i,t} W_{i,t}$-1.

statistics for the linked waves 2015-2017 are shown in Table 2. Notice that the sample size is smaller compared to the previous Table 1. This is because we only know the location of the first property reported in the survey. Therefore,

Table 2: Decomposition of household wealth growth when house prices are imputed. Sorting based on growth rate of net worth.

(a) 2015-2017 (House prices based on median prices)

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	8,627	20.1%	14.9%	24.0%	21.6%	12.7%	8.9%
Quintile 1	1,605	-51.7%	-54.6%	17.4%	16.5%	11.6%	4.9%
Quintile 2	1,635	-9.8%	-15.0%	22.9%	22.5%	12.8%	9.7%
Quintile 3	1,705	13.0%	7.6%	25.3%	21.4%	12.0%	9.5%
Quintile 4	1,942	43.5%	38.3%	26.4%	19.7%	11.1%	8.5%
Quintile 5	1,740	116.4%	108.5%	26.1%	30.2%	17.2%	12.9%

(b) 2015-2017 (House prices based on mean prices)

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	7,934	22.7%	17.2%	24.6%	22.1%	13.1%	9.0%
Quintile 1	1,461	-46.2%	-49.6%	18.1%	18.7%	13.2%	5.5%
Quintile 2	1,498	-8.3%	-13.9%	24.3%	23.2%	13.1%	10.1%
Quintile 3	1,573	14.1%	8.6%	25.4%	21.6%	12.1%	9.5%
Quintile 4	1,808	43.5%	38.3%	26.7%	19.7%	10.9%	8.7%
Quintile 5	1,594	108.7%	101.0%	27.0%	28.7%	16.8%	11.9%

Notes: House values are imputed based on median or mean of square meter values reported in the community where properties are located. The sample is restricted to households who own only one real estate property since we only know the location of the first property. See footnote in the previous table for the description of how the variables are computed.

to construct this table, we restrict the sample to households that own only one property.

With the imputed prices of housing, the quintile variation in growth rate is somewhat smaller: about -50% for the bottom quintile and about 110% for the top quintile, compared to -80% and 185% when housing prices are the self-reported. Still, there is a large variation and capital gains is the largest source of variation.

1.4 Variance decomposition of wealth growth

To characterize the driving forces for individual wealth growth, we conduct a variance decomposition based on equation (1). For convenience we rewrite the equation here as

$$g_t^W = \text{capital gain}_t + \text{saving}_t. \tag{3}$$

This allows us to compute the importance of two factors for the dispersion of wealth growth: capital gains and savings. The results for each of the linked surveys are reported in the top section of Table 3. Most of the variation in wealth growth can be attributed to capital gains as they account for more than 80% of the variance.

Table 3: Variance decomposition of wealth growth.

	Full sample			
	Std	Gain	Save	Cov
2011-2013	1.18	81.63%	2.69%	15.68%
2013-2015	1.10	83.09%	0.95%	15.96%
2015-2017	1.07	83.92%	3.24%	12.83%

	Without housing debt				**With housing debt**			
	Std	Gain	Save	Cov	Std	Gain	Save	Cov
2011-2013	1.17	79.28%	2.66%	18.06%	1.21	93.84%	3.07%	3.08%
2013-2015	1.10	82.45%	0.74%	16.80%	1.08	86.89%	2.62%	10.49%
2015-2017	1.05	82.51%	2.63%	14.86%	1.12	90.09%	5.87%	4.04%

	One-house owner				**Multiple-house owner**			
	Std	Gain	Save	Cov	Std	Gain	Save	Cov
2011-2013	1.12	74.34%	1.95%	23.71%	0.84	94.60%	0.08%	5.32%
2013-2015	0.99	77.45%	0.53%	22.01%	0.73	94.50%	0.29%	5.21%
2015-2017	1.01	79.54%	1.85%	18.61%	0.78	94.72%	4.23%	1.06%

Table 3 also conducts a variance decomposition for different sub-samples. We first separate households with and without housing debt. We then separate households that own one house from households that own multiple houses. The table shows that capital gains are relatively more important (they account for a larger share of variance) for households that have housing debt and own multiple houses. This suggests that borrowing against the owned house and owning multiple houses are important for wealth mobility.

Next we conduct a variance decomposition after sorting households in quintiles based on their initial wealth. The results are reported in Table 4. Low wealth households have higher variance. This was to be expected since the wealth quintiles are calculated based on the initial wealth (for example, for the 2015-2017 matched samples, households are sorted according to their wealth in 2015), and it is normal for households with low wealth to experience larger growth rates of wealth, both positive and negative. But the key message is that capital gains are the most important force for the volatility of growth for all wealth class and their contribution increases with the initial value of wealth.

Table 4: Variance decomposition of growth for different quintiles of wealth based on initial wealth. Linked surveys 2015-2017.

	Std	Gain	Save	Cov
Quintile 1	1.42	68.22%	3.64%	28.15%
Quintile 2	1.05	92.71%	2.64%	4.66%
Quintile 3	0.93	95.16%	3.22%	1.62%
Quintile 4	0.92	97.18%	4.04%	-1.22%
Quintile 5	0.79	98.78%	4.23%	-3.00%

Another way to look at the role of housing assets and housing debt in generating wealth mobility is by constructing wealth mobility matrices. Table 5 reports the wealth mobility matrices for the full samples and the sub-samples of households with housing debt and multiple houses. The thresholds used to calculate the transition probabilities in the sub-samples remain the same as those used for the whole sample. For economy of space we report here only the transition matrices for the last linked waves 2015-2017. The transition matrices constructed with the previous surveys are provided in the appendix.

Comparing the transition matrices for the whole sample and the two sub-samples we find that households with housing debt and multiple houses are more likely to move upward and less likely to move downward. As we will see, this property is consistent with the regression analysis we will present later in the empirical section of the paper.

1.5 The role of housing capital gains

We further decompose the capital gains into gains from housing assets and gains from other assets. To do so, we rewrite the wealth growth equation as

$$\Delta W = H\Delta p + A\Delta q + S, \tag{4}$$

where ΔW is the change in wealth (net worth). The variable H denotes the size of housing assets, p the price of houses, A the size of other assets and q the price of other assets. Thus, $H\Delta p$ represents the capital gains realized

Table 5: Wealth mobility matrices for whole sample and sub-samples with housing debt and multiple houses. Linked surveys 2015-2017.

Full sample (2015-2017)			
	Bottom	Middle	Top
Bottom	64.0%	30.9%	5.0%
Midde	18.5%	54.4%	27.1%
Top	7.1%	16.2%	76.7%

With housing debt (2015-2017)			
	Bottom	Middle	Top
Bottom	45.2%	44.2%	10.5%
Middle	11.1%	50.4%	38.5%
Top	5.5%	14.3%	80.2%

With multiple houses (2015-2017)			
	Bottom	Middle	Top
Bottom	60.3%	34.9%	4.8%
Middle	17.0%	49.5%	33.5%
Top	4.9%	15.3%	79.8%

from the ownership of houses, $A\Delta q$ represents the capital gains from other assets, and S is saving.

In the data we do not observe whether households have changed their houses during the two years of linked surveys. Therefore, we cannot compute the capital gains on houses precisely. As a proxy we then use the total change in the value of houses, $\Delta(Hp)$, instead of $H\Delta p$. This would be the right measure of housing capital gains if households did not change the size and location of the owned houses over the two survey years. If they changed the size and/or location of their houses, our measure is just a proxy for the capital gains on houses. Once we have the proxy for housing capital gains, $H\Delta p$, and the measure of saving, S, we can compute the capital gains on other assets as a residual from equation (4), that is, $A\Delta q = \Delta W - H\Delta p + S$.

Table 6 reports the variance for each component of wealth (housing, other assets and savings) as a percentage of total variance. The top section of the table reports the statistics computed on the whole sample, while the statistics reported in the bottom section are computed on the restricted sample of households who do not change houses over the two-year period (which we assume to be the case if they report the same housing size in the two survey years). In this case, $H\Delta p$ truly captures capital gains on housing. As can be seen from the table, capital gains on housing is the predominant source of variation for wealth growth.

To support the finding that changes in house prices could be quite het-

Table 6: Variance decomposition of wealth growth.

(a) Full sample

	Housing assets $\frac{\text{Var}(H\Delta p)}{\text{Var}(\Delta W)}$	Other assets $\frac{\text{Var}(A\Delta q)}{\text{Var}(\Delta W)}$	Savings $\frac{\text{Var}(S)}{\text{Var}(\Delta W)}$	Covariances $\frac{\text{Cov}}{\text{Var}(\Delta W)}$
2011-2013	53.42%	26.73%	2.69%	17.16%
2013-2015	67.37%	15.29%	0.95%	16.39%
2015-2017	70.05%	11.11%	3.24%	15.60%

(b) With unchanged housing only

	Housing assets $\frac{\text{Var}(H\Delta p)}{\text{Var}(\Delta W)}$	Other assets $\frac{\text{Var}(A\Delta q)}{\text{Var}(\Delta W)}$	Savings $\frac{\text{Var}(S)}{\text{Var}(\Delta W)}$	Covariances $\frac{\text{Cov}}{\text{Var}(\Delta W)}$
2011-2013	47.06%	34.10%	1.78%	17.06%
2013-2015	56.78%	23.76%	0.52%	18.94%
2015-2017	64.00%	17.12%	1.97%	16.91%

Notes: The decomposition of capital gains is based on the accounting equation $\Delta W = H\Delta p + A\Delta q + S$, where ΔW is the change in wealth, H the size of housing assets in first year, p the price of houses, A the size of other assets in first year and q the price of the other assets. Since in the data we do not know whether a household has changed the house over the two year period, we cannot compute the capital gains $H\Delta p$ exactly. Then, in order to proxy for these capital gains, we use the total change in the value of houses, $\Delta(Hp)$ instead of $H\Delta p$. Once we have the proxy for the capital gains in housing, $H\Delta p$, and the measure of saving, S, we can compute the capital gains in other assets as $A\Delta q = \Delta W - H\Delta p + S$.

erogeneous among households, Figure 2 in the appendix shows the median growth rate of house prices in Beijing during 2013-2015. Data is from Lianjia which is similar to Zillow for the United States. Since the data is based on actual house transactions, it should be quite accurate. As can be seen, the two-year change in prices has been quite heterogeneous. In certain areas of Beijing the price even declined while in others the growth rate was more than 80%. This shows that even within a single metropolitan area, the change in housing prices could be quite heterogeneous. A price index for the whole metropolitan area of Beijing would hide the fact that housing values grow at very different rates in different parts of the city. But predicting which areas will experience faster growth is not always possible.[4]

We can also show that there is significant heterogeneity in the growth

[4] If the market could predict which area will experience faster price growth, that prediction should be immediately reflected in current price. So, the fact that in 2013 the prices in areas that experienced fast growth were much lower than in 2015 is an indication that the market cannot predict the change in price.

of housing prices within other major cities in China using the CHFS data. Figure 3 in the appendix plots the median price growth between 2015 and 2017 for different communities within few major cities. The median price per square meter is the median of all prices reported by households living in the particular community of the city. As can be seen, there is significant price variation within each city.

Finally, we show that there is also heterogeneity in housing price growth across districts/counties. Table 26 in the appendix reports the two-year district/county level house price growth. In China, district or county is an administrative division under a city. For example, Beijing has 16 districts. We collect the district level house price data from Lianjia. For each year we calculate the house price growth rate in each district. We then sort districts into 5 quintile groups based on the growth rate of house prices and calculate the average growth rate for each quintile group. As can be seen, there is significant price variation across districts/counties.

One of the reasons capital gains on housing play the predominant role in generating volatility in wealth growth is because housing is the largest component of individual portfolios. As shown in Table 7, housing assets account for 70 percent of total households' assets. Financial assets which include the stock market account for only 13 percent which, as observed earlier is quite low. Business assets is the value of private businesses and account for about 11 percent ot total assets. We may have expected to find a larger share of business assets. However, even if private businesses might generate significant income for their owners, their sale value may be relatively low due to the central role played by the owner of the business.

1.6 Regression analysis

So far we have shown that capital gains, especially on housing, are the main determinant of cross-sectional variation in wealth growth. In this section we provide additional evidence on the determinant of wealth growth using regression analysis. For simplicity we report only the results based on the most recent surveys. The regression results based on earlier surveys are similar.

We consider five dependent variables:

1. g^w: growth rate of wealth.
2. $P(g^W_{High})$: probability of being in the top 33% of wealth growth.

Table 7: The share of housing assets in total assets. Quintile sorting based on growth rate of wealth.

(a) Asset composition in 2015

	Obs	$\frac{\text{hs-asset}}{\text{asset}}$	$\frac{\text{fin-asset}}{\text{asset}}$	$\frac{\text{bus-asset}}{\text{asset}}$	$\frac{\text{oth-asset}}{\text{asset}}$	$\frac{\text{t-debt}}{\text{asset}}$	$\frac{\text{hs-debt}}{\text{asset}}$
Full sample	15,742	69.14%	13.23%	10.75%	6.88%	4.99%	2.97%
Quintile 1	3,111	59.16%	9.32%	23.62%	7.90%	4.47%	1.74%
Quintile 2	3,068	65.24%	14.65%	12.12%	7.98%	4.31%	1.95%
Quintile 3	2,969	71.05%	15.93%	6.90%	6.12%	3.68%	2.24%
Quintile 4	3,300	77.05%	13.39%	3.94%	5.62%	5.11%	3.66%
Quintile 5	3,294	77.03%	12.77%	3.62%	6.59%	9.23%	7.05%

(b) Asset composition in 2017

	Obs	$\frac{\text{hs-asset}}{\text{asset}}$	$\frac{\text{fin-asset}}{\text{asset}}$	$\frac{\text{bus-asset}}{\text{asset}}$	$\frac{\text{oth-asset}}{\text{asset}}$	$\frac{\text{t-debt}}{\text{asset}}$	$\frac{\text{hs-debt}}{\text{asset}}$
Full sample	15,742	75.48%	12.46%	5.62%	6.44%	6.24%	3.85%
Quintile 1	3,111	61.55%	16.83%	7.80%	13.82%	34.84%	11.57%
Quintile 2	3,068	72.87%	13.56%	5.60%	7.97%	6.17%	3.79%
Quintile 3	2,969	74.18%	12.98%	6.22%	6.62%	5.13%	3.57%
Quintile 4	3,300	78.58%	12.25%	3.80%	5.37%	4.10%	3.08%
Quintile 5	3,294	76.55%	11.29%	6.48%	5.68%	5.15%	3.72%

Notes: hs=housing; fin=financial; bus=business; oth=other; t=total.

3. $P(g^W_{Low})$: probability of being in the bottom 33% of wealth growth.
4. $P(up)$: probability of moving out of the bottom 33% of wealth growth.
5. $P(down)$: probability of moving out of the top 33% of wealth growth.

We regress these variables on several indicators and the results are reported in Table 8. The first column of the table shows the results when the dependent variable is wealth growth. This variable is negatively correlated with initial wealth and positively correlated with savings. The estimates have an intuitive interpretation: the impact of initial wealth is negative because of 'reversal-to-the-mean' while savings raise next period wealth by definition.

Importantly, we find that wealth growth is positively associated with housing. Wealth increases more if households has multiple houses, purchased new houses during the sample period or had houses with increasing housing prices. We use the term 'newly purchased houses' for households who purchased a house during the sample period and 'housing appreciation' for households who owned at least one house whose price increased more than

Table 8: Housing debt and wealth growth. Linked surveys 2015-2017.

	(1) g^W	(2) $P(g^W_{High})$	(3) $P(g^W_{Low})$	(4) $P(\text{up})$	(5) $P(\text{down})$
Lag wealth	-0.35***	-0.12***	0.06***	-0.10***	0.07***
	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)
Saving	0.55***	0.17***	-0.19***	0.11***	-0.02***
	(0.05)	(0.02)	(0.02)	(0.02)	(0.00)
Lag housing debt	**1.81*** **	**0.98*** **	**-0.58*** **	**0.48*** **	**-0.07*** **
	(0.19)	(0.09)	(0.08)	(0.08)	(0.03)
New purchased house	0.79***	0.31***	-0.18***	0.09***	-0.04***
	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
Housing appreciation	1.07***	0.52***	-0.33***	0.15***	-0.06***
	(0.03)	(0.01)	(0.01)	(0.01)	(0.00)
Multiple house owner	0.21***	0.04***	-0.13***	-0.04***	-0.06***
	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
Business owner	0.28***	0.08***	0.00	0.01	0.01
	(0.04)	(0.02)	(0.02)	(0.01)	(0.01)
Age	0.02***	0.01***	-0.01***	0.00	-0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Age^2	-0.00***	-0.00***	0.00***	-0.00***	0.00***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
College	0.14***	0.05***	-0.08***	-0.00	-0.04***
	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)
Family size	-0.02*	-0.01***	0.01***	-0.00	0.00
	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)
Constant	4.10***	1.56***	-0.14*	1.37***	-0.73***
	(0.20)	(0.08)	(0.09)	(0.07)	(0.04)
Observations	15,551	15,551	15,551	15,551	15,551
R-squared	0.38	0.36	0.17	0.17	0.11

50% during the sample period. We also find that wealth growth is positively correlated with education and it has an inverse U-shape relation with age. Finally, owning a business is important for increasing the growth rate of wealth. A result that is consistent with studies based on US data (see Quadrini (1999)).

The housing variables and business ownership have positive effects on the probability of experiencing high growth of wealth (second column of Table 8) and moving to the upper group (fourth column of Table 8). The housing variables are also significant in explaining the probability of experiencing low growth but with the opposite sign. These results show that housing is an important factor for wealth mobility in China.

Table 8 also shows that housing debt is important for mobility: households with higher debt experience higher growth, face higher probability of moving up, and lower probability of moving down the distribution of wealth.

This shows that borrowing against housing assets could be an important way to enhance the likelihood of moving up in the distribution of wealth. The question, however, is whether many Chinese households have access to credit or choose to borrow.

Table 9 shows that in 2017 only 16.4 percent of households in the sample had housing debt. Conditional on having housing debt, the average debt to asset ratio was 13.48%, and the average housing debt to income ratio was 193.22%. Similar statistics are found in other survey years. Therefore, even if debt could be important for mobility, only a small fraction of Chinese households borrow against housing assets. Furthermore, the value of debt for those who borrow is relatively small. This points out that the financial structure of China is still in a development stage. Limited borrowing may be considered an impediment to enhance mobility. From a macro perspective, however, less debt may provide greater financial and macroeconomic stability.

Table 9: Statistics for housing debt. Average 2015-2017 surveys.

(a) Full sample

	Obs	Mean	Std	Min	Max
HouseDebt/Income	15,523	25.23%	65.64%	0.00%	247.39%
TotalDebt/Income	15,523	50.65%	109.10%	0.00%	399.93%
HouseDebt/Asset	15,742	2.40%	6.08%	0.00%	22.04%
TotalDebt/Asset	15,742	5.86%	12.16%	0.00%	43.80%

(b) With positive housing debt

	Obs	Mean	Std	Min	Max
HouseDebt/Income	2,539	141.67%	87.75%	0.00%	247.39%
TotalDebt/Income	2,539	193.22%	137.94%	0.00%	399.93%
HouseDebt/Asset	2,584	13.48%	7.62%	0.00%	22.04%
TotalDebt/Asset	2,584	19.93%	14.18%	0.00%	43.80%

1.7 Comparison with other surveys

We conclude the empirical section of the paper by showing that some of the statistics reported in the previous subsections are comparable to those obtained from two other surveys. The first survey is the China Family Panel Studies (CFPS). This is a nationally representative, annual longitudinal survey of Chinese communities, families, and individuals launched in 2010 by the Institute of Social Science Survey of Peking University. The second survey is the Survey of Consumer Finance (SCF) for the United States. The

SCF is a triennial survey of the balance sheet, pension, income and other demographic characteristics of families in the United States, sponsored by the United States Federal Reserve Board in cooperation with the U.S. Treasury Department.

Table 10 shows the same statistics reported in the previous Table 1 but computed from the 2015-2017 CFPS waves. The comparison of the statistics reported in the bottom section of Table 1 with those reported in Tables 10 reveals that the CHFS data displays similar properties as the CFPS data.

Table 10: Household wealth growth from 2015-2917 CFPS data

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	9104	16.95%	11.09%	16.34%	35.86%	19.65%	16.22%
Quintile 1	1846	-81.61%	-84.10%	8.98%	27.73%	15.48%	12.24%
Quintile 2	1810	-35.34%	-39.43%	13.45%	30.37%	16.66%	13.72%
Quintile 3	1770	4.25%	-0.49%	14.44%	32.84%	18.75%	14.09%
Quintile 4	1865	56.10%	50.02%	18.55%	32.80%	18.36%	14.45%
Quintile 5	1813	248.46%	230.87%	23.40%	75.19%	38.18%	37.01%

Notes: Sorted by growth rate of wealth.

The comparison with the Survey of Consumer Finances is more difficult because the SCF does not have a longitudinal structure as the CHFS and CFPS, which is necessary to compute individual growth rates in wealth. This is why studies that use the SCF characterize the cross-sectional properties of the data. Examples are Quadrini and Ríos-Rull (2015) and Kuhn, Ríos-Rull, et al. (2016). However, in 2009, the Federal Reserve Board (FRB) implemented a follow-up survey that sampled the same families interviewed in the 2007 survey. Therefore, for the years 2007-2009 we can compute the growth rate of wealth for each household. The longitudinal dimension is no longer available in subsequent waves, which explains why we focus only on these two years. The growth rate of wealth for the five quintiles, with sorting based on wealth growth, are shown in Table 11.

The growth rate of net worth in the SCF data is also quite volatile as for the CHFS and the CFPS. The average growth for the full sample is negative in the SCF. This is because the two surveys were conducted in the years before and after the financial crisis during which the prices of many assets declined. Yet, the top quintile experienced an average growth rate of 116.8%.

We would like to point out that, the reason we did not report the decom-

Table 11: Household wealth growth from 2007-2009 SCF data.

Full sample	Quintile 1	Quintile 2	Quintile 3	Quintile 4	Quintile 5
-25.64%	-81.11%	-48.75%	-22.20%	4.61%	116.8%

Notes: Sorted by growth rate of wealth.

position of wealth growth in capital gains and savings for the United States is because the SCF does not allows us to do so. While the CHFS (and the CFPS) includes data for income and consumption expenditures, which in turn allows us to compute savings, the SCF provides only data on income, not consumption expenditures. Without data on savings we cannot compute capital gains by subtracting them from the change in wealth.

To conclude, the high dispersion in wealth growth we showed in the previous subsections is not just an artifact of the Chinese economy. We observe a similar degree of dispersion in the United States.

2 The model

The empirical analysis highlights several features of the Chinese economy that could be relevant for understanding some of the forces behind wealth distribution and mobility. In this section we construct a general equilibrium model that incorporates the empirical features outlined in the previous section. We will then use the model to conduct counterfactual exercises to explore the distributional impact of certain changes such as financial development and policy reforms.

The economy is populated by a mass 1 of agents, each surviving with probability $1 - \omega$. Exiting agents are replaced by a mass ω of newborn agents so that population remains constant over time. Agents maximize the lifetime expected utility

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \ln(c_t).$$

where $c_t = \hat{c}_t + \chi h_t$ is total consumption resulting from the sum of non-housing consumption, $\hat{c}_t$, and housing services χh_t (h_t is the stock of houses and χ is a constant parameter). The housing services that enter the utility function are in addition to actual income earned on housing as specified below. The discount factor $\beta = \hat{\beta}(1-\omega)$ is the product of the inter-temporal discount factor $\hat{\beta}$ and the survival rate, $1-\omega$. Newborn agents are endowed with the average states of surviving agents as specified below.

At any point in time agents are heterogeneous in human capital denoted by l_t. Human capital evolves endogenously according to

$$l_{t+1} = \eta_t l_t + e_t.$$

The variable e_t is human capital investment and η_t is an idiosyncratic shock that increases or decreases the existing stock l_t. Human capital earns the competitive wage rate w_t.

We think of individual labor earnings as broadly defined. They include not only the typical wage income but also profits from small undiversified businesses. Therefore, human capital also includes investments in undiversified businesses.

The shock η_t is iid with mean value normalized to $\mathbb{E}\eta_t = 1$. Even though the shock is iid, earnings are very persistent because the shock affects the stock of human capital. Modelling labor earnings as an endogenous process that depends on human capital investment is analytically convenient since agents' decisions are linear in wealth which allows for linear aggregation.

Agents can hold three types of real and financial assets: housing, h_t, stock market, k_t, and bonds, b_t. Houses are in limited supply and they are traded at an average price P_t. Housing assets are subject to both aggregate and idiosyncratic shocks as we will describe shortly. The stock market represents the diversified ownership of business capital and it is subject to aggregate shocks only. We assume that business capital is reproducible without adjustment cost so that its price is always 1. Bonds do not carry any risk and pay the gross return R_t. Bond holdings can be negative in which case the agent borrows. Borrowing, however, is limited by the collateral constraint

$$-b_{t+1} \leq \xi\Big(P_t h_{t+1} + \lambda k_{t+1} + l_{t+1}\Big). \tag{5}$$

The collateral constraint depends on the value of owned houses, stock market and human capital. The latter is a proxy for labor income which is also taken into account by lenders when they screen loan applications.[5] We

[5]The dependence on human capital could be multiplied by the expected value of the next period wage, that is, $w_{t+1}l_{t+1}$. We decided ignore this term because it introduces some computational complications when we solve for the general equilibrium. However, since the wage rate is an aggregate variable, it should not play an important role in affecting the heterogeneity of households' decisions, which is the focus of our paper.

also allow stock market assets (business capital) to be used as a collateral. However, the collateral value of the stock market is likely to be lower than housing and labor income. In the calibration this will be captured by setting $\lambda < 1$.

The stock market generates the cash flow $r_t^k k_t$, while houses generate the cash flow $r_t^h h_t$. Houses, however, are subject to idiosyncratic stochastic appreciation/depreciation ψ_t: h_t units of houses purchased in the previous period become $\psi_t h_t$ effective units this period. The stochastic variable ψ_t is iid with its mean normalized to $\mathbb{E}\psi_t = 1$. We think of ψ_t as capturing idiosyncratic local factors that increase or decrease the value of a housing unit relatively to the aggregate price P_t. The aggregate price P_t is determined in the general equilibrium.

While the stock market is a diversified investment that is not subject to idiosyncratic shocks (although the return depends on aggregate shocks), housing and human capital are undiversified investments which depend on idiosyncratic shocks, in addition to aggregate shocks.

To capture heterogeneous participation in assets markets, we assume that agents face a heterogeneous cost to invest in housing, stock market and human capital. The cost is $\tau_t(h_{t+1}P_t + k_{t+1} + l_{t+1})$, where τ_t is an idiosyncratic stochastic variable that follows a discrete first-order Markov process. Since the cost is idiosyncratic and, therefore, differs among agents, those with a lower cost invest a larger fraction of their wealth in housing, stock market and human capital. This also implies that in equilibrium low-cost agents borrow from high-cost agents. We interpret this cost as reflecting a variety of factors such as financial literacy and, more generally, financial development. One of the goals of this paper is to explore how participation in high return markets—an indicator of financial development—impacts wealth distribution and mobility.

The budget constraint for an agent with investment cost τ_t is

$$c_t + (1+\tau_t)\Big[P_t h_{t+1} + k_{t+1} + l_{t+1}\Big] + b_{t+1} = R_t^h h_t + R_t^k k_t + R_t^l l_t + R_t b_t, \quad (6)$$

where R_t^h, R_t^k, R_t^l, R_t are the gross returns earned, respectively, on houses, stock market, human capital and bonds (or interest paid if b_t is negative). Since $c_t = \hat{c}_t + \chi h_t$ includes housing services that enter directly the utility function, the return on houses R_t^h also includes these services as we will see more explicitly below.

We now have all the elements to write down the optimization problem solved by an individual household. Given $\mathbf{x}_t = (\tau_t, \psi_t, \eta_t)$ the vector of idiosyncratic shocks, the household's problem can be written as

$$V_t(\mathbf{x}_t; h_t, k_t, b_t) \quad = \quad \max_{\substack{c_t, h_{t+1}, \\ k_{t+1}, b_{t+1}}} \left\{ \ln(c_t) + \beta \mathbb{E}_t V_t(\mathbf{x}_{t+1}; h_{t+1}, k_{t+1}, b_{t+1}) \right\}, \quad (7)$$

subject to the borrowing constraint—equation (5)—and the budget constraint—equation (6). The subscript t in the value function captures the dependence on aggregate states. The housing price P_t and the returns R_t^h, R_t^k, R_t^l, R_t are all determined in general equilibrium.

Production technology. There is a continuum of competitive firms with production function

$$Y_t = z_t H_t^{\theta_H} K_t^{\theta_K} L_t^{\theta_L},$$

where H_t is the input of houses, K_t is the input of capital, L_t is the input of labor and z_t is aggregate productivity. The share parameters satisfy $\theta_H + \theta_K + \theta_L = 1$ (constant return to scale).

Capital is held in part by the government, denoted by $K_{g,t}$, and in part by the private sector, denoted by $K_{p,t}$. Therefore, $K_t = K_{g,t} + K_{p,t}$. Through the choice of $K_{g,t}$, the government affects the stock market assets held by the private sector. Changes in government ownership of productive capital is one of the policies we will explore later in the quantitative section of the paper.

Notice that housing services enter the economy through two channels: as direct inputs of the production function and in the utility function. By having housing services as an input of production we can easily determine their contribution to aggregate output (GDP). However, with the sole contribution to output, it would not be possible to generate the high valuation of housing observed in the Chinese economy. In other words, the model would not be able to generate a share of housing in household's wealth similar to what we observe in the data (around 70%). The utility value of housing services, then, is pinned down by the share of housing in households' portfolios.

The optimality conditions for the representative firm are

$$\begin{aligned} r_t^h &= \theta_H z_t H_t^{\theta_H - 1} K_t^{\theta_K} L_t^{\theta_L}, \\ r_t^k &= \theta_K z_t H_t^{\theta_H} K_t^{\theta_K - 1} L_t^{\theta_L}, \\ w_t &= \theta_L z_t H_t^{\theta_H} K_t^{\theta_K} L_t^{\theta_L - 1}, \end{aligned}$$

and the gross returns on housing, stock market and human capital are

$$\begin{aligned} R_t^h &= r_t^h + \chi + \psi_t P_t, \\ R_t^k &= r_t^k + 1 - \delta, \\ R_t^l &= w_t + \eta_t. \end{aligned}$$

While r_t^h, r_t^k and w_t are subject to the aggregate shock z_t, the gross returns R_t^h and R_t^l depend also, respectively, on the idiosyncratic shocks ψ_t and η_t. This captures the fact that investments in housing and human capital are less diversified than stock market investments. Small businesses, of course, are also very undiversified. This is the reason we think of the income generated by small businesses as part of the process that determines earnings $w_t l_t$. Consistent with this interpretation, investments in human capital e_t also include investments in small businesses.[6]

The income earned by the government through the ownership of business capital is used to fund public consumption, G_t, and public investment, $K_{g,t+1} - (1-\delta)K_{g,t}$. The budget constraint for the government is

$$G_t + K_{g,t+1} = R_t^k K_{g,t}. \tag{8}$$

We assume that public consumption G_t affects agents' utility additively to the utility from private consumption c_t. Thus, G_t does not affect the marginal utility from private consumption.

2.1 First order conditions and portfolio choices

The linearity of the return from all types of investments, including human capital, is a convenient property that allows us to aggregate individual decisions for all agents that have the same access to financial markets, that is, the same value of τ_t.

[6]It may be claimed that stock market investments are also undiversified as in practice some households do not hold a well diversified portfolio of shares. For these households, however, this is a deliberate choice rather than the consequence of being unable to diversify stock market investments. Something that is much more difficult to do with housing and human capital. In some cases, the lack of diversification may be related to the particular position of the household vis-a-vis the company it invests. For example, a member of the household could have a managerial position in the firm it invests for incentive purposes. Effectively, this case is more a reflection of the lack of human capital diversification rather than the lack of financial diversification. In the model this is captured by the idiosyncratic risk on human capital.

Define $a_t = R^h_t h_t + R^k_t k_t + R^l_t l_t + R_t b_t$ the household's net worth at the end of the period. This is an 'extended' measure of net worth because it includes the household's human capital as well as the housing services that enter the utility function χh_t. Using the variable a_t and taking into account that the idiosyncratic shocks ψ_t and η_t are iid, we can rewrite the agent's problem as

$$V_t(\tau_t; a_t) = \max_{c_t, h_{t+1}, k_{t+1}, l_{t+1}, b_{t+1}} \Big\{ \ln(c_t) + \beta \mathbb{E}_t V_t(\tau_{t+1}; a_{t+1}) \Big\}, \tag{9}$$

subject to:

$$\begin{aligned} c_t &= a_t - (1+\tau_t)\Big[P_t h_{t+1} + k_{t+1} + l_{t+1}\Big] - b_{t+1} \\ a_{t+1} &= R^h_{t+1} h_{t+1} + R^k_{t+1} k_{t+1} + R^l_{t+1} l_{t+1} + R_{t+1} b_{t+1} \\ -b_{t+1} &\leq \xi\Big(P_t h_{t+1} + \lambda k_{t+1} + l_{t+1}\Big). \end{aligned}$$

The iid properties of ψ_{t+1} and η_{t+1} allow us to replace the vector of state variables $\mathbf{x}_t = (\tau_t, \psi_t, \eta_t)$ with only τ_t. Current realizations of the housing and human capital shocks, ψ_t and η_t, affect the agent's decision only through their impact on the end-of-period net worth a_t.

We can now normalize the household's problem by a_t and rewrite it as

$$\tilde{V}_t(\tau_t) = \max_{\substack{\tilde{c}_t, \tilde{h}_{t+1}, \tilde{k}_{t+1}, \\ \tilde{l}_{t+1}, \tilde{b}_{t+1}}} \Big\{ \log(\tilde{c}_t) + \beta \mathbb{E}_t \tilde{V}_{t+1}(\tau_{t+1}) + \frac{\beta}{1-\beta}\mathbb{E}_t \log(g_{t+1}) \Big\}, \tag{10}$$

subject to:

$$\begin{aligned} \tilde{c}_t &= 1 - (1+\tau_t)\Big[P_t \tilde{h}_{t+1} + \tilde{k}_{t+1} + \tilde{l}_{t+1}\Big] - \tilde{b}_{t+1} \\ g_{t+1} &= R^h_{t+1} \tilde{h}_{t+1} + R^k_{t+1} \tilde{k}_{t+1} + R^l_{t+1} \tilde{l}_{t+1} + R_{t+1} \tilde{b}_{t+1} \\ -\tilde{b}_{t+1} &\leq \xi\Big(P_t \tilde{h}_{t+1} + \lambda \tilde{k}_{t+1} + \tilde{l}_{t+1}\Big), \end{aligned}$$

where

$$\tilde{V}_t(\tau_t) = V_t(\tau_t; a_t) - \frac{\log(a_t)}{1-\beta}, \qquad g_{t+1} = \frac{a_{t+1}}{a_t}.$$

A tilde sign denotes variables divided (normalized) by a_t. For example, $\tilde{c}_t = c_t/a_t$ and $\tilde{h}_{t+1} = h_{t+1}/a_t$. The variable $g_{t+1} = a_{t+1}/a_t$ is the gross growth rate of net worth.

The normalized problem has only one individual state variable, that is, the current value of the investment cost τ_t. Still, agents are heterogeneous in wealth a_t. However, as long as they have the same τ_t, their decisions are linear multiple of their wealth a_t.

The first order conditions are

$$\tilde{h}_{t+1} \quad : \quad \frac{(1+\tau_t)P_t}{\tilde{c}_t} = \frac{\beta}{1-\beta}\mathbb{E}_t\left(\frac{R^h_{t+1}}{g_{t+1}}\right) + \mu_t \xi P_t, \tag{11}$$

$$\tilde{k}_{t+1} \quad : \quad \frac{1+\tau_t}{\tilde{c}_t} = \frac{\beta}{1-\beta}\mathbb{E}_t\left(\frac{R^k_{t+1}}{g_{t+1}}\right) + \mu_t \xi \lambda, \tag{12}$$

$$\tilde{l}_{t+1} \quad : \quad \frac{1+\tau_t}{\tilde{c}_t} = \frac{\beta}{1-\beta}\mathbb{E}_t\left(\frac{R^l_{t+1}}{g_{t+1}}\right) + \mu_t \xi, \tag{13}$$

$$\tilde{b}_{t+1} \quad : \quad \frac{1}{\tilde{c}_t} = \frac{\beta}{1-\beta}\mathbb{E}_t\left(\frac{R_{t+1}}{g_{t+1}}\right) + \mu_t, \tag{14}$$

where μ_t is the Lagrange multiplier associated with the collateral constraint.

The normalized first order conditions are the same for all agents with the same investment cost τ_t. Therefore, all households with the same τ_t make the same portfolio decisions. Households with different values of τ_t, however, face different (expected) returns on housing, stock market and human capital and, therefore, they choose different composition of portfolio, that is, they choose different values of $\tilde{h}_{t+1}$, $\tilde{k}_{t+1}$, $\tilde{l}_{t+1}$ and $\tilde{b}_{t+1}$. In particular, we will see that agents with higher investment cost τ_t choose positive values of $\tilde{b}_{t+1}$ and become lenders while agents with lower investment cost τ_t chose negative values of $\tilde{b}_{t+1}$ and become borrowers. For borrowing agents the collateral constraint could be binding or not binding. In the first case $\mu_t > 0$. In the second case $\mu_t = 0$. Differences in portfolio choices imply that agents experience different stochastic properties of wealth growth and mobility.

If we multiply the first order conditions (11)-(14) by $\tilde{h}_{t+1}$, $\tilde{k}_{t+1}$, $\tilde{l}_{t+1}$, $\tilde{b}_{t+1}$, respectively, and we add them together, we obtain

$$\frac{1-\tilde{c}_t}{\tilde{c}_t} = \frac{\beta}{1-\beta}.$$

This implies that $\tilde{c}_t = 1-\beta$ or, equivalently, $c_t = (1-\beta)a_t$. Thus, consumption is a constant fraction of net worth, which is a well-known property of models with log utility.

We now define $\tilde{\mu}_t = (1-\beta)\mu_t$ and substitute $\tilde{c}_t = 1-\beta$ in the first order conditions to obtain the following five equations,

$$
\begin{aligned}
(1+\tau_t)P_t &= \beta\mathbb{E}\left(\frac{R^h_{t+1}}{g_{t+1}}\right) + \tilde{\mu}_t\xi P_t, &(15)\\
1+\tau_t &= \beta\mathbb{E}\left(\frac{R^k_{t+1}}{g_{t+1}}\right) + \tilde{\mu}_t\xi\lambda, &(16)\\
1+\tau_t &= \beta\mathbb{E}\left(\frac{R^l_{t+1}}{g_{t+1}}\right) + \tilde{\mu}_t\xi, &(17)\\
1 &= \beta\mathbb{E}\left(\frac{R_{t+1}}{g_{t+1}}\right) + \tilde{\mu}_t, &(18)\\
\tilde{\mu}_t &= 0, \qquad \text{if} \qquad -\tilde{b}_{t+1} < \xi_t(P_t\tilde{h}_{t+1} + \lambda\tilde{k}_{t+1} + \tilde{l}_{t+1}), &(19)
\end{aligned}
$$

where the gross growth rate of net worth is given $g_{t+1} = R^h_{t+1}\tilde{h}_{t+1} + R^k_{t+1}\tilde{k}_{t+1} + R^l_{t+1}\tilde{l}_{t+1} + R_{t+1}\tilde{b}_{t+1}$.

The five conditions above represent a system of five equations in five individual unknowns: $\tilde{h}_{t+1}$, $\tilde{k}_{t+1}$, $\tilde{l}_{t+1}$, $\tilde{b}_{t+1}$, and $\tilde{\mu}_t$. We can solve for these five unknowns numerically using a nonlinear solver. Because of the normalization, the solution depends on the individual state τ_t but not on the extended net worth a_t.

2.2 General equilibrium and numerical solution

A key object that we need to find in order to solve for the general equilibrium is the housing price P_t. This is a function of the aggregate states denoted by $\mathbf{s}_t$. We can then express the housing price as $P_t = \mathcal{P}(\mathbf{s}_t)$.

Since the idiosyncratic state τ_t is a finite state Markov process, it can take I values. Therefore, at any point in time there are I groups or types of agents, each characterized by a particular realization of τ_t. Agents in each group differ in the endogenous states. However, to characterize the general equilibrium, we only need the aggregation of the endogenous states for each group $i = 1, ..., I$. We denote the group-aggregate states by H^i_t, K^i_t, L^i_t, B^i_t. Even if an agent is in a group i today, it may be in a different group in the next period because τ_t changes stochastically. Thus, the sufficient set of aggregate states we need to keep track in order to solve for the equilibrium is $\mathbf{s}_t = \left\{z_t, \{H^i_t, K^i_t, L^i_t, B^i_t\}^I_{i=1}\right\}$.

We can reduce further the set of sufficient aggregate states by defining the variable

$$N_t^i = (r_t^h + \chi)H_t^i + R_t^k K_t^i + \bar{R}_t^l L_t^i + R_t B_t^i,$$

where $\bar{R}_t^l$ is the gross return on human capital averaged over the idiosyncratic shock η_t, that is, $\bar{R}_t^l = \int_\eta R_t^l f(\eta) d\eta$. The variable N_t^i is the aggregate net worth of type i agents, including human capital, but without housing wealth $P_t H_t^i$. More specifically, given A_t^i the aggregation of the extended net worth of type i agents, the variable N_t^i is equal to $A_t^i - P_t H_t^i$. Using the variable N_t^i, the sufficient set of aggregate states are $\mathbf{s}_t = \left\{z_t, \{H_t^i, N_t^i\}_{i=1}^I\right\}$.

If we knew the price function $\mathcal{P}(\mathbf{s}_t)$, we could predict the next period price $P_{t+1} = \mathcal{P}(\mathbf{s}_{t+1})$ for each value of next period states $\mathbf{s}_{t+1}$. This would allow us to solve for the general equilibrium at any period t. However, since we do not know the housing price function $\mathcal{P}(.)$, we also need to solve for it as part of our computational procedure.

To make the numerical procedure operational, we approximate $\mathcal{P}(.)$ with some functional form. In particular, we use the following approximation

$$P_{t+1} = \sum_{j}^{I^z} \alpha_z^j D_{t+1}^j + \sum_{i=1}^{I-1} \alpha_H^i H_{t+1}^i + \sum_{i=1}^{I} \alpha_N^i N_{t+1}^i \tag{20}$$

where D_{t+1}^j is the dummy variable for the j realization of aggregate productivity z_{t+1} (which can take I^z values). Finding the function $\mathcal{P}(\mathbf{s}_t)$ would then require finding the values of the coefficients α_z^j, α_H^i, α_N^i. Note that the summation for housing contains only $I-1$ terms because aggregate housing is constant in the model. The detailed numerical procedure is described in Appendix B

3 Quantitative analysis

The goal of quantitative analysis is to use the calibrated model to conduct counterfactual exercises in order to address specific questions. In particular, we inquire about how financial development, such as greater access to credit and/or higher financial participation, affects wealth distribution and mobility. Additionally, we will utilize the model to investigate how government ownership of productive capital impacts wealth concentration and mobility. We begin with a description of the calibration.

3.1 Calibration

Since the China Household Finance Survey (CHFS) is conducted every two years and some of the statistics used to calibrate the model require merging two consecutive surveys (for example, to compute the growth rate of individual wealth), the period in the model is set to two years.

The production function takes the form $Y_t = z_t H_t^{\theta_H} K_t^{\theta_K} L_t^{\theta_L}$ with share parameters $\theta_H = 0.15$, $\theta_K = 0.51$ and $\theta_L = 0.34$. The calibrated income share of housing, 15 percent, is higher than the number reported in the official Chinese statistics for rental income. However, the general view is that the official number underestimates the actual income generated by housing, which motivates the higher calibration number. After setting $\theta_H = 0.15$, the remaining income compensates capital and labor. The data suggests that, abstracting from housing, capital income accounts for about 60% and labor income for 40%, consistent with the estimates in Bai and Qian (2010). Therefore, we set $\theta^K = 0.85 \times 0.6 = 0.51$ and $\theta_L = 0.85 \times 0.4 = 0.34$.

For the counterfactual exercises we consider the version of the model without aggregate shocks and set z_t to 0.5. Different values of productivity would simply re-scale the model and are irrelevant for the results of the paper. The depreciation rate of capital over the two-year period is set to $\delta = 0.15$.

The model features three idiosyncratic shocks: ψ_t (housing), η_t (human capital) and τ_t (investment cost). The first shock (housing) is iid and takes five values with equal probability. To pin down these five values we use the cross-sectional distribution of housing values from the 2013, 2015 and 2017 surveys. We first compute the individual growth rate in housing price between 2013 and 2015. We then order households according to their individual growth rate and arrange them in quintiles. The five values of ψ are set using the deviation of the aggregate growth rate of each quintile from the sample mean. More specifically, we set $\psi^i = 1 - (g^i - \sum_{i=1}^{5} g^i/5)$, where g^i is the aggregate growth rate for decile $i = 1, ...5$. We do the same using the 2015-2017 surveys and then we average the values of ψ^i computed from the 2013-2015 surveys and the 2015-2017 surveys.

To calibrate the shock to human capital, η_t, we use the same procedure. We first construct quintiles for the growth rate of labor and business incomes combined, separately for the linked 2013-2015 surveys and the 2015-2017 surveys. After calculating the deviation of growth from the sample mean of each quintile, we average over 2013-2015 and 2015-2017. The resulting numbers are reported in Table 12. It is important to point out that, even if the idiosyncratic shocks ψ_t and η_t are iid, their impact is highly persistent since they affect the stocks of housing and human capital.

Table 12: Distribution of housing price growth and earning growth

	Quintile				
	1st	*2nd*	*3rd*	*4th*	*5th*
Housing price growth (ψ)	0.568	0.875	0.963	1.076	1.516
Labor earning growth (η)	0.454	0.853	1.011	1.170	1.510

The remaining idiosyncratic shock is the investment cost τ_t. We assume that τ_t follows a two-state first order Markov process and normalize the lower value $\underline{\tau}$ to zero. Thus, a fraction of households can access high return investments without incurring any cost. The actual identity of this group of households changes over time since a new τ_t is drawn in every period. With the normalization $\underline{\tau} = 0$ we only need to calibrate the high value $\bar{\tau}$, and the transition probabilities.

To calibrate the transition probabilities we use the equilibrium property for which households with $\tau_t = \underline{\tau}$ borrow while households with $\tau_t = \bar{\tau}$ do not borrow. We then use the empirical two-year transition probabilities from borrowers (positive housing debt) to non-borrowers (zero housing debt) as the probability of switching from $\underline{\tau}$ to $\bar{\tau}$. Similarly, we use the empirical two-year transition probability from non-borrowers (zero housing debt) to borrowers (positive housing debt) as the probability of switching from $\bar{\tau}$ to $\underline{\tau}$. The empirical two-year transition matrix, averaged over 2013-2015 and 2015-2017, is reported in Table 13. The calibrated transition probabilities imply that the steady state fraction of households with low investment cost is about 15 percent.

Table 13: Two-year transition probability matrix $\Gamma(\tau_t, \tau_{t+1})$

	Borrowers	Non-borrowers
Borrowers	0.52	0.48
Non-borrowers	0.08	0.92

To calibrate the last parameter pertaining to the investment cost, $\bar{\tau}$, we use conditions (12) and (14). Agents with $\tau_t = \bar{\tau}$ do not borrow. Therefore, $\tilde{\mu}_t = 0$. Since we are considering the steady state without aggregate shocks,

R^k_{t+1} is not stochastic. Conditions (12) and (14) then imply that

$$1 + \bar{\tau} = \frac{R^k_{t+1}}{R_{t+1}}.$$

Thus, $\bar{\tau}$ is directly related to the spread between the average return on the stock market, R^k_{t+1}, and the interest rate, R_{t+1}, which we set to 14% for the bi-annual period. This is consistent with estimates of two-year average stock market returns for China.

At this point we are left with five parameters: the utility from owning houses, χ, the collateral parameters, ξ and λ, the discount factor $\hat{\beta}$ and the death probability ω (remember that $\beta = \hat{\beta}(1-\omega)$). In addition, we need to fix the stock of physical capital held by the government.

For the parameter λ, we assume that the collateral value of the stock market is 50% lower than the collateral value of houses. In China, the maximum leverage ratio for stocks is 50% but most of stocks cannot be used as a collateral. Therefore, the average collateral rate for stocks is much lower than 50%. The collateral rate of physical capital is around 30%-50% for firms in China (see for example, Song, Storesletten, and Zilibotti (2011)), while the collateral rate for housing is 70%-80%. All considered, the assumption that stocks have a collateral value that is half the collateral value of houses is a plausible approximation.

After fixing the value of $\lambda = 0.5$, we calibrate the remaining four parameters and the capital held by the government jointly to match the following targets: (i) 70% share of housing in households' portfolio; (ii) 15% aggregate debt over (two-year) output; (iii) 6% bi-annual interest rate ($R_{t+1} = 1.06$); (iv) 50% share of productive capital held by the government, which is consistent with the estimates in Piketty et al. (2019); (v) 0.7 Gini index for wealth.[7] These calibration targets are taken from the data (see empirical Section 1). The full set of parameter values are reported in the top section of Table 14 while the bottom section reports some steady state statistics.

[7] Although the four parameters contribute jointly to the calibrated targets, individually they are more important for certain targets. In particular, χ is especially important for targeting the share of housing wealth; $\hat{\beta}$ for targeting the interest rate (given the equity spread captured by the parameter $\bar{\tau}$); ξ for the level of debt; ω for the Gini index (lower mortality increases wealth concentration).

Table 14: Calibration and steady state statistics

(a) Calibration values

Parameter description	*Values*
Discount factor	$\beta = 0.8985$
Death probability	$\omega = 0.0116$
Utility from housing	$\chi = 0.028$
Aggregate productivity	$\bar{z} = 0.5$
Income shares	$\theta_H = 0.15$, $\theta_K = 0.51$, $\theta_L = 0.34$
Capital depreciation	$\delta = 0.15$
Collateral parameter	$\xi = 0.176$
Collateral on k	$\lambda = 0.5$
Investment cost	$\tau_t \in \left\{0, 0.1321\right\}$, $\Gamma(\tau_t, \tau_{t+1}) = \begin{bmatrix} 0.92 & 0.08 \\ 0.48 & 0.52 \end{bmatrix}$
Housing shocks	$\psi_t \in \left\{0.5686, 0.8756, 0.9636, 1.0761, 1.5161\right\}$
Labor earning shocks	$\eta_t \in \left\{0.4544, 0.8534, 1.0114, 1.1704, 1.5104\right\}$

(b) Steady state statistics

Variable description	*Values*
House price	0.172
Output	0.082
Debt-Output ratio	0.151
Privately owned capital	0.060
Publicly owned capital	0.060
Housing share in wealth	0.695
Return on bonds	0.060 (3% annually)
Return on stock market	0.199 (10% annually)
Wealth Gini	0.699
Top percentiles of wealth	0.458 (top 5%), 0.277 (top 1%), 0.135 (top 0.1%)

3.2 Steady state statistics

Most of the statistics reported at the bottom section of Table 14 are calibration targets. For example, we impose that the model generates a wealth Gini of 0.7. The other distributional statistics such as the top percentiles, though, are not targeted in the calibration.

Figure 1 plots the Lorenz curve for wealth, measured by net worth, in the data (2015 CHFS) and in the model. As can be seen, the two curves are very similar. The model generates a slightly more concentrated distribution of wealth than in the data as the artificial Lorenz curve (dashed line) lies somewhat below the empirical Lorenz curve (continuous line). But the difference is not big. This can also be seen in the last row of Table 14 which reports the shares of wealth held by the top 1 percent of households. In the model this share is 27.7%, which is close to the number reported in Piketty et al. (2019)—about 29% of national wealth—but it is bigger than the share computed from the CHFS. We would like to point out, though, that the CHFS survey misses the super wealthy despite the over-sampling of wealthier households. Accounting for the super rich may increase the concentration statistics at the very top of the distribution.

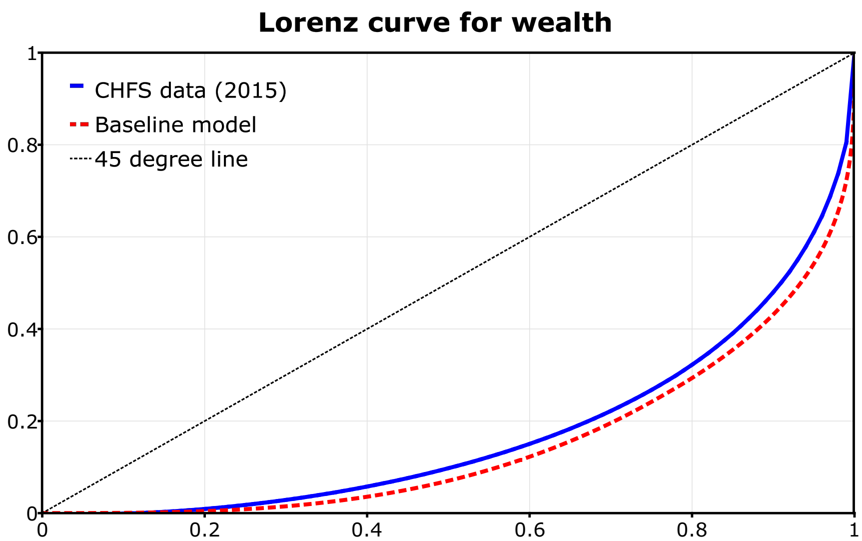

Figure 1: Lorenz curve for net worth constructed using data from the 2015 CHFS and from the steady state of the calibrated model.

One dimension of interest is the participation in investment markets which is determined by the cost τ_t. At any point in time about 85% of households face the high cost $\bar{\tau}$ while the remaining 15% face the low cost $\underline{\tau} = 0$. The cost determines the portfolio choices made by households, which in turn affect their wealth mobility. The portfolio choices and mobility statistics are shown in Table 15.

Agents with low investment cost allocate a larger fraction of their wealth in housing, stock market and human capital. Their bond ownership is negative, meaning that they borrow from agents with high investment costs. As a result of allocating a larger share of wealth in high return assets (housing, stock market and human capital), the average growth rate of wealth of low cost households is much higher than for high cost households. In fact, for high cost households, the expected growth rate of wealth is slightly negative.

Table 15: Portfolio composition and wealth growth properties

	Portfolio composition				Wealth Stats	
	Housing	*Stock market*	*Human capital*	*Bonds*	*Mean growth*	*Std growth*
High cost, $\bar{\tau}$	0.462	0.170	0.324	0.044	-0.004	0.348
Low cost, $\underline{\tau}$	0.600	0.181	0.419	-0.200	0.200	0.517

At the same time, because low-cost households allocate a larger faction of their wealth in high volatile assets (housing and human capital), they experience higher standard deviation of growth. Thus, low-cost households are characterized by higher upward mobility (on average they experience a higher rate of wealth growth) and higher overall volatility of growth (up and down).

3.3 Structural changes

In this section, we conduct counterfactual exercises to assess the impact of several changes that could emerge as a result of financial development or policy reforms. The first change considers an increase in access to credit. The second change allows for greater accessability or participation to investment markets. The third change is a reduction in government ownership of productive capital, represented by K_g, through privatization.

3.3.1 Higher access to credit

Higher access to credit is obtained by increasing the collateral parameter ξ. We increase ξ so that debt over output doubles in the new steady state—from 15% to 30%. This can be interpreted as replicating the rise in household debt that increased from 30% of GDP to more than 60% during the 2010-2020 period. We then compare the steady state properties of the model with the new value ξ to the steady state of the baseline model. The results are reported in Table 16

Higher access to credit leads to an increase in aggregate production due to higher investment made by low-cost households (households with $\tau = \underline{\tau}$). These households can now use more debt to fund investments in housing, stock market and human capital. As a result, more savings are allocated in reproducible factors (physical and human capital), which has a positive impact on aggregate production. Low-cost households also invest more in housing but this impacts only the market price of houses since the aggregate

Table 16: Counterfactual exercise with higher access to credit.

(a) Steady state properties of macro variables

	Baseline model	*Model with higher ξ*
House price	0.172	0.180
Output	0.082	0.088
Debt-Output ratio	0.151	0.302
Private capital	0.060	0.070
Public capital	0.060	0.060
Housing share	0.695	0.635
Return on bonds	0.060	0.057
Return on stocks	0.199	0.195

(b) Portfolio and distributional properties

		Baseline model		*Model with higher ξ*	
		$\bar{\tau}$	$\underline{\tau}$	$\bar{\tau}$	$\underline{\tau}$
Portfolio Composition	Housing	0.462,	0.600	0.422,	0.677
	Stocks	0.170,	0.181	0.176,	0.225
	Human	0.324,	0.419	0.312,	0.486
	Bonds	0.044,	−0.200	0.091,	−0.388
Wealth Mobility	Ave of growth	−0.004,	0.200	−0.024,	0.331
	Std of growth%	0.348,	0.517	0.319,	0.641
Wealth distribution	Gini	0.699		0.716	
	Top 5%	0.458		0.482	
	Top 1%	0.277		0.301	
	Top 0.1%	0.135		0.154	

supply is fixed. Since physical capital rises more than the price of houses, the share of housing in the portfolio of households declines.

We look now at the composition of portfolio for low and high cost households. Remember that about 85% of households incur the high investment cost $\bar{\tau}$ while the remaining 15% face the low cost $\underline{\tau} = 0$. The portfolio of low-cost households contains a larger share of high return assets (housing, stock market and human capital), in part funded with debt acquired by high-cost households. The portfolio compositions of the two types of households is important for the overall distribution of wealth and mobility. Higher access to credit, induced by the higher value of ξ, makes the differences in portfolio composition between low-cost and high-cost households even bigger. As a result of this change, the distribution of wealth becomes more concentrated. For example, the share held by the top 1% increases from 0.277 to 0.301. This follows from the fact that low-cost households now hold a more leveraged portfolio which allows them to experience higher mean growth as well as higher volatility of growth. On the other hand, high-cost households hold a larger share of safer assets (the debt issued by low-cost households) and they experience lower volatility of growth.

In summary, greater access to credit has a positive macroeconomic impact but it makes the overall distribution of wealth more unequal. It also increases the differences in mobility: higher mobility for the low-cost households and lower mobility for high cost households.

3.3.2 Financial participation

Higher participation in high return markets can be generated in the model in two ways. The first is with a reduction in the investment cost $\bar{\tau}$. The second is with a reduction in the fraction of households that face the high investment cost $\bar{\tau}$. The first change induces more participation through the *intensive* margin, that is, high-cost households allocate a larger share of their wealth in high return assets. The second change induces more participation through the *extensive* margin, that is, more households hold portfolios with a larger shares of high return assets. Both of these changes can be seen as the result of financial development, part of which induced by regulatory reforms.

We begin by lowering the high investment cost $\bar{\tau}$. We would like to reiterate that the investment cost τ should be interpreted broadly. Besides capturing actual transaction costs in financial markets, it could also reflect lack of information (low financial literacy) or aversion to more complex investment operations (also due to low financial literacy). In the exercise conducted here we reduce $\bar{\tau}$ by half. The results are reported in Table 17.

The lower value of $\bar{\tau}$ increases the effective return from investing in physical and human capital for high-cost households, thus raising the overall return from savings. This, in turn, leads to a significant increase in savings and capital accumulation. In general equilibrium, the returns from physical and human capital decline due to the higher inputs of physical and human capital, which lower their marginal products. However, since physical and human capital combined account for 85 percent of the production inputs ($\theta_K + \theta_L = 0.85$), the marginal products are not very sensitive to their supply. Consequently, even moderate reductions in marginal products result in large increases in the supplies of physical and human capital, which lead to a substantial increase in aggregate production.

Since the supply of houses is fixed, the large increases in physical and human capital raise in the marginal product of houses, which in turn generates

Table 17: Counterfactual with lower investment cost $\bar{\tau}$.

(a) Steady state properties of macro variables

	Baseline model	*Model with lower $\bar{\tau}$*
House price	0.172	0.355
Output	0.082	0.226
Debt-Output ratio	0.151	0.141
Private capital	0.060	0.325
Public capital	0.060	0.060
Housing share	0.695	0.496
Return on bonds	0.060	0.078
Return on stocks	0.199	0.150

(b) Portfolio and distributional properties

		Baseline model		*Model with lower $\bar{\tau}$*	
		$\bar{\tau}$	$\underline{\tau}$	$\bar{\tau}$	$\underline{\tau}$
Portfolio Composition	Housing	0.462,	0.600	0.315,	0.380
	Stocks	0.170,	0.181	0.282,	0.390
	Human	0.324,	0.419	0.367,	0.414
	Bonds	0.044,	−0.200	0.035,	−0.184
Wealth Mobility	Ave of growth	−0.004,	0.200	−0.001,	0.096
	Std of growth%	0.348,	0.517	0.266,	0.333
Wealth distribution	Gini	0.699		0.613	
	Top 5%	0.458		0.366	
	Top 1%	0.277		0.196	
	Top 0.1%	0.135		0.080	

a large increase in price. The increase in production and housing price would be smaller if the production share of reproducible factors was lower. For example, if human capital was fixed in aggregate, the share of reproducible factors would be $\theta_K = 0.51$ instead of $\theta_K + \theta_L = 0.85$.[8] But what is important for our paper are the distributional and mobility consequences resulting from lower investment costs. We can see that the distribution of wealth becomes less unequal. For example, the share of wealth held by the top 1% declines from 0.277 to 0.196. This is because the composition of portfolios for low-cost and high-cost households becomes more similar. Consequently, we observe a decrease in mobility, as measured by the volatility of growth,

[8]We could modify the model and assume that individual human capital h_t represents the *share* of a fixed aggregate human capital. If a household invests more than other households, it will increase its relative position within the distribution, but it will not affect the aggregate stock of human capital. This alternative model would have similar distributional properties, but the overall impact of the structural changes would be smaller.

for both low and high-cost households.

To summarize, greater participation in capital markets induced by lower investment cost (intensive margin) has a positive effect on the aggregate economy and leads to a more equal distribution of wealth.

We now consider the counterfactual in which we increase the number of households that incur the low investment cost (extensive participation margin). We do this by changing the structure of the transition probability matrix for the investment cost τ_t.

The baseline calibration of the transition matrix implies that 85% of households face the high investment cost $\bar{\tau}$, and 15% the low investment cost $\underline{\tau} = 0$. In the new calibration we impose that the transition probability matrix is symmetric so that in the steady state 50% of households face the high cost $\bar{\tau}$ and 50% no cost. More specifically, we change $\Gamma(\bar{\tau}, \bar{\tau})$ from 0.92 to 0.52, which is also the number for $\Gamma(\underline{\tau}, \underline{\tau})$. The results are reported in Table 18

The impact of having a larger number of households facing a low investment cost is similar to lowering $\bar{\tau}$. The macroeconomic impact is positive and large for the same reasons a lower value of $\bar{\tau}$ induces a macroeconomic boom (described above). The distribution of wealth becomes less concentrated as a result of the fact that the average growth of wealth experienced by the two groups of households is more similar and the standard deviation of growth decreases for both types of households. Therefore, we also have that a more extensive participation is beneficial for the aggregate economy and reduces wealth inequality.

3.3.3 Privatization

A large share of Chinese businesses is under the corporate control of the government through State-Owned Enterprises (SOEs). We would like to use the model to investigate whether privatization of the SOEs could affect wealth distribution and mobility. We do this by comparing the steady-state equilibrium in the baseline model, where half of the physical capital is held by the public sector, with the steady-state equilibrium in which physical capital is entirely private. The results are reported in Table 19.

From a macro perspective, privatization has negative consequences as it reduces aggregate production. This result has an intuitive explanation: to incentivize the private sector to hold more capital, its return must increase. Since the return from capital is determined by its marginal product, the

Table 18: Counterfactual with a higher fraction of low-cost households (higher participation).

(a) Steady state properties of macro variables

	Baseline model	*Model with higher part.*
House price	0.172	0.383
Output	0.082	0.240
Debt-Output ratio	0.151	0.481
Private capital	0.060	0.348
Public capital	0.060	0.060
Housing share	0.695	0.438
Return on bonds	0.060	0.014
Return on stocks	0.199	0.150

(b) Portfolio and distributional properties

		Baseline model $\bar{\tau}$	*Baseline model* $\underline{\tau}$	*Model with higher part.* $\bar{\tau}$	*Model with higher part.* $\underline{\tau}$
Portfolio Composition	Housing	0.462,	0.600	0.256,	0.377
	Stocks	0.170,	0.181	0.195,	0.386
	Human	0.324,	0.419	0.333,	0.422
	Bonds	0.044,	−0.200	0.216,	−0.185
Wealth Mobility	Ave of growth	−0.004,	0.200	−0.092,	0.124
	Std of growth%	0.348,	0.517	0.226,	0.324
Wealth distribution	Gini	0.699		0.650	
	Top 5%	0.458		0.402	
	Top 1%	0.277		0.225	
	Top 0.1%	0.135		0.098	

aggregate stock of capital must decrease. This, in turn, reduces the marginal product of housing and human capital, leading to a decline in both the price of houses and investment in human capital. However, it is important to note that this prediction does not consider the potential impact of privatization on corporate governance, which could also affect output. This aspect is not analyzed in our paper as it falls beyond its scope.

While the consequences for the aggregate economy are negative, privatization leads to a more equal distribution of wealth. The primary reason is that privatization changes the composition of households' portfolios: in equilibrium, households hold a smaller share of housing assets and a larger share of the stock market. The share of housing in total household wealth (houses plus stock market) declines from 70 percent before privatization to 57 percent.

Essentially, when the government owns a large share of physical capital,

Table 19: Counterfactual with privatization.

(a) Steady state properties of macro variables

	Baseline model	*Model with $K^g = 0$*
House price	0.172	0.165
Output	0.082	0.075
Debt-Output ratio	0.151	0.172
Private capital	0.060	0.107
Public capital	0.060	0.000
Housing share	0.695	0.571
Return on bonds	0.060	0.067
Return on stocks	0.199	0.207

(b) Portfolio and distributional properties

		Baseline model $\bar{\tau}$	*Baseline model* $\underline{\tau}$	*Model with $K^g = 0$* $\bar{\tau}$	*Model with $K^g = 0$* $\underline{\tau}$
Portfolio Composition	Housing	0.462,	0.600	0.406,	0.545
	Stocks	0.170,	0.181	0.284,	0.283
	Human	0.324,	0.419	0.268,	0.364
	Bonds	0.044,	−0.200	0.043,	−0.192
Wealth Mobility	Ave of growth	−0.004,	0.200	−0.010,	0.181
	Std of growth%	0.348,	0.517	0.290,	0.437
Wealth distribution	Gini	0.699		0.655	
	Top 5%	0.458		0.414	
	Top 1%	0.277		0.238	
	Top 0.1%	0.135		0.108	

it creates a shortage of alternative saving instruments that can be held by households for the allocation savings. Consequently, households hold more houses relative to other assets in equilibrium. Since houses are subject to more idiosyncratic risks than the stock market, the growth rate of individual wealth is also more volatile, which leads to more wealth concentration.[9]

This finding may seem at odds with the view that housing wealth is a safer form of investment compared to the stock market. According to this view, privatization would generate more inequality because households would invest a larger share of their savings in the stock market. However, this conclusion is based on the comparison of housing price indices and stock

[9] Our counterfactual exercise highlights one mechanism that, in certain countries, could contribute to the shortage of saving assets discussed in Caballero, Farhi, and Gourinchas (2017): the government ownership of corporations. Although the definition of safe assets often discussed in the literature does not include the stock market, a broader definition would include it for reasons discussed in the main text.

market indices, both of which do not reflect the idiosyncratic volatility of their components. But while diversified investments in the stock market are easily achievable, diversified investments in housing are rare.

To summarize, privatization may have a negative impact on aggregate economic activity because it would be associated with lower aggregate savings. However, it allows for more diversified portfolios which lead to lower volatility in individual wealth growth. This, in turn, results in a more equal distribution of wealth.

4 Conclusion

We have explored the properties of individual wealth growth and mobility in China using the China Household Finance Survey (CHFS) and outlined three main findings. The first finding is that savings play a relatively minor role in explaining individual wealth mobility. While households with higher savings experience higher growth in wealth, the heterogeneity in savings across households explains only a small portion of the heterogeneity in wealth growth. Instead, the most significant contribution to the dispersion of individual wealth growth is the cross-sectional heterogeneity in capital gains from assets. This indicates that individual wealth in China is highly undiversified.

The second finding is that housing wealth plays an important role in generating wealth mobility. This derives from two features of the Chinese economy. First, housing represents the largest component of households' wealth. Second, there is significant cross-sectional dispersion in capital gains on houses. These two facts further indicate that individual wealth is very undiversified, and households' portfolios are exposed to large idiosyncratic risks.

The third finding is that households' debt increases wealth mobility. Households that hold more debt (higher leverage) tend to experience greater volatility of wealth growth.

These findings raise several questions. If housing ownership is so risky, why do Chinese households allocate such a large fraction of their portfolio to housing assets? If housing debt enhances mobility, should borrowing be encouraged through policies?

To address these questions, we built a general equilibrium model with heterogeneous agents where households choose three types of assets: housing, stock market investment, and bonds (or debt when negative). An important form of heterogeneity is the ability to participate in investment markets. After calibrating the model to the Chinese economy, we conduct several experiments. We first relax the financial constraints faced by households. We then allow for greater participation in investment markets. Finally, we consider the privatization of State-Owned Enterprises.

The quantitative results show that higher access to credit and higher financial participation have positive effects on aggregate production. However, while the expansion of credit makes the distribution of wealth more concentrated, it reduces mobility for households with lower access to capital markets, higher participation leads to a more equal distribution of wealth. Finally, we find that privatization could lead to a more equal distribution of wealth, although it could have negative aggregate effects by reducing national savings.

While some of the changes considered in these experiments could be the natural consequence of financial development—as financial markets evolve, credit and investment markets become more accessible to the wider society—they could also be encouraged through policies. This is certainly the case for privatization. The fact that certain changes have different effects on aggregate outcomes and wealth distribution implies that some changes may be more desirable than others.

References

Bai, C.-E., & Qian, Z. (2010). The factor income distribution in China: 1978–2007. *China Economic Review*, *21*(4), 650–670.

Benhabib, J., Bisin, A., & Luo, M. (2017). Earnings inequality and other determinants of wealth inequality. *American Economic Review*, *107*(5), 593–597.

Caballero, R. J., Farhi, E., & Gourinchas, P.-O. (2017). The safe assets shortage conundrum. *Journal of economic perspectives*, *31*(3), 29–46.

Cagetti, M., & DeNardi, M. (2006). Entrepreneurship, Frictions and Wealth. *Journal of Political Economy*, *114*(5), 835–70.

Chen, K., & Wen, Y. (2017). The great housing boom of China. *American Economic Journal: Macroeconomics*, *9*(2), 73–114.

Cooper, R., & Zhu, G. (2017). Household Finance in China. NBER Working Paper No. 23741.

Gan, L., Yin, Z., Jia, N., Xu, S., Ma, S., Zheng, L., et al. (2014). Data you need to know about China. *Springer Berlin Heidelberg. https://doi*, *10*(978), 3.

Glaeser, E., Huang, W., Ma, Y., & Shleifer, A. (2017). A Real Estate Boom with Chinese Characteristics. *Journal of Economic Perspectives*, *31*(1), 93–116.

Han, B., Han, L., & Zhu, G. (2018). Housing price and fundamentals in a transition economy: The case of the Beijing market. *International Economic Review*, *59*(3), 1653–1677.

Jiang, S., Miao, J., & Zhang, Y. (2022). China's Housing Bubble, Infrastructure Investment and Economic Growth. *International Economic Review*, *63*(3), 1189–1237.

Kuhn, M., Ríos-Rull, J.-V., et al. (2016). 2013 Update on the US earnings, income, and wealth distributional facts: A View from Macroeconomics. *Federal Reserve Bank of Minneapolis Quarterly Review*, *37*(1), 2–73.

Piketty, T., Yang, L., & Zucman, G. (2019). Capital accumulation, private property, and rising inequality in China, 1978–2015. *American Economic Review*, *109*(7), 2469–2496.

Quadrini, V. (1999). The Importance of Entrepreneurship for Wealth Concentration and Mobility. *Review of Income and Wealth*, *45*(1), 1–19.

Quadrini, V. (2000). Entrepreneurship, Saving and Social Mobility. *Review of Economic Dynamics*, *3*(1), 1–40.

Quadrini, V., & Ríos-Rull, J.-V. (2015). Inequality in macroeconomics. In *Handbook of income distribution*, Vol. 2, pp. 1229–1302. Elsevier.

Song, Z., Storesletten, K., & Zilibotti, F. (2011). Growing like china. *American economic review*, *101*(1), 196–233.

Zeng, T., & Zhu, S. (2022). The mobility of top earnings, income, and wealth in China: Facts from the 2011–2017 China household finance survey. *Journal of Asian Economics*, *80*, 101461.

Zhang, F. (2016). Inequality and House Prices. Working Paper, PBC School of Finance, Tsinghua University.

Zhang, Y. (2017). Liquidity constraints, transition dynamics, and the Chinese housing return premium. Working Paper, Peking University.

Appendix

A China Household Finance Survey (CHFS)

The CHFS employs a stratified three-stage probability proportion to size (PPS) random sample design. The primary sampling units (PSU) include 2,585 counties (including county level cities and districts) from all Chinese provinces and municipalities except Tibet, Xinjiang, Inner Mongolia, Hong Kong, Macau, and Taiwan. The second stage of sampling selects residential communities from the counties/cities selected in the first sampling stage. The third stage selects households from the residential communities chosen in the previous stage. Every sampling stage is performed with the PPS method and weighted by its population size. See Gan, Yin, Jia, Xu, Ma, Zheng, et al. (2014) for more details about the sampling scheme.

A residential community is a social living collective composed of people who reside within a certain geographical area and does not coincide with any administrative unit. Communities are the management scope of residents' committees, which are legally mandated grassroots self-governing organizations for urban residents. The size of a community varies but typically contains thousand of households. For example, there are about 5,000 communities across 16 districts in the city of Beijing.

In the urban areas, the number of households selected varies according to the average housing price in the selected community. Communities are sorted into quartiles based on the average housing price of each neighborhood: 50 households are drawn from each residential community in the top quartile and 25 from each residential community in the bottom quartile. This implies that the survey oversamples wealthy households.

B Numerical procedure

The numerical procedure consists of three steps:

1. Guess the coefficients of the approximated price function $\mathcal{P}(\mathbf{s}_t)$: α_z^j, α_H^i, α_N^i. See equation (20).

2. Solve for the general equilibrium at time $t = 1, ..., T$ with the following steps:

    (a) Given the states $\mathbf{s}_t = (z_t, H_t^i, N_t^i)$, for $i = 1, ..., I$, we guess the equilibrium prices P_t, R_{t+1} and the normalized individual decisions $\tilde{h}_{t+1}^i$,

$\tilde{k}^i_{t+1}$, $\tilde{l}^i_{t+1}$, $\tilde{b}^i_{t+1}$ for each group i. Since the individual decisions are normalized by net worth a^i_t, they are the same for all agents of type i (that is, agents with the same τ_t).

(b) Using the states H^i_t, N^i_t and the guessed price P_t, we compute the net worth for each group i,

$$A^i_t = P_t H^i_t + N^i_t.$$

This allows us to compute the next period aggregate variables

$$\begin{aligned} H^j_{t+1} &= \sum_i \left(\tilde{h}^i_{t+1} A^i_t\right) \Gamma_{ij} \\ K^j_{t+1} &= \sum_i \left(\tilde{k}^i_{t+1} A^i_t\right) \Gamma_{ij} \\ L^j_{t+1} &= \sum_i \left(\tilde{l}^i_{t+1} A^i_t\right) \Gamma_{ij} \\ B^j_{t+1} &= \sum_i \left(\tilde{b}^i_{t+1} A^i_t\right) \Gamma_{ij}. \end{aligned}$$

The term Γ_{ij} is the transition probability for the investment cost τ_t.

(c) We now compute the aggregate values of the production inputs in the next period,

$$\begin{aligned} H_{t+1} &= \sum_j H^j_{t+1} \\ K_{t+1} &= \sum_j K^j_{t+1} \\ L_{t+1} &= \sum_j L^j_{t+1}, \end{aligned}$$

which in turn allows us to compute the next period returns for each realization of the aggregate shock z_{t+1},

$$\begin{aligned} r^h_{t+1} &= \theta_H z_{t+1} \bar{H}^{\theta_H - 1} K^{\theta_K}_{t+1} L^{\theta_L}_{t+1}, \\ R^k_{t+1} &= \theta_K z_{t+1} H^{\theta_H}_{t+1} K^{\theta_K - 1}_{t+1} L^{\theta_L}_{t+1} + 1 - \delta, \\ \bar{R}^l_{t+1} &= \theta_L z_{t+1} H^{\theta_H}_{t+1} K^{\theta_K}_{t+1} L^{\theta_L - 1}_{t+1} + \bar{\eta}. \end{aligned}$$

(d) At this point we have all the ingredients needed to compute the next period aggregate state N^j_{t+1} for each type j,

$$N^j_{t+1} = r^h_{t+1} H^j_{t+1} + R^k_{t+1} K^j_{t+1} + \bar{R}^l_{t+1} L^j_{t+1}.$$

We use N^j_{t+1} with the guessed price function for housing to compute the next period price for each realization of z_{t+1}, that is,

$$P_{t+1} = \sum_{j}^{I^z} \alpha_z^j D^j_{t+1} + \sum_{i=1}^{I-1} \alpha_H^i H^i_{t+1} + \sum_{i=1}^{I} \alpha_N^i N^i_{t+1}.$$

(e) We now check the accuracy of the initial guesses for the individual decisions and the prices P_t and R_t we made in step 2a. To verify the guesses we check the following conditions:

- First order conditions for individual decisions (15)-(19) for each $i = 1, ..., I$. In particular, using $g_{t+1} = R^h_{t+1}\tilde{h}_{t+1} + R^k_{t+1}\tilde{k}_{t+1} + R^l_{t+1}\tilde{l}_{t+1} + R_{t+1}\tilde{b}_{t+1}$, equations (15)-(19) can be rewritten as

$$\begin{aligned}
(1+\tau_t)P_t &= \beta\mathbb{E}\left(\frac{R^h_{t+1}}{R^h_{t+1}\tilde{h}_{t+1} + R^k_{t+1}\tilde{k}_{t+1} + R^l_{t+1}\tilde{l}_{t+1} + R_{t+1}\tilde{b}_{t+1}}\right) + \tilde{\mu}_t\xi P_t, \\
1+\tau_t &= \beta\mathbb{E}\left(\frac{R^k_{t+1}}{R^h_{t+1}\tilde{h}_{t+1} + R^k_{t+1}\tilde{k}_{t+1} + R^l_{t+1}\tilde{l}_{t+1} + R_{t+1}\tilde{b}_{t+1}}\right) + \tilde{\mu}_t\xi\lambda, \\
1+\tau_t &= \beta\mathbb{E}\left(\frac{R^l_{t+1}}{R^h_{t+1}\tilde{h}_{t+1} + R^k_{t+1}\tilde{k}_{t+1} + R^l_{t+1}\tilde{l}_{t+1} + R_{t+1}\tilde{b}_{t+1}}\right) + \tilde{\mu}_t\xi, \\
1 &= \beta\mathbb{E}\left(\frac{R_{t+1}}{R^h_{t+1}\tilde{h}_{t+1} + R^k_{t+1}\tilde{k}_{t+1} + R^l_{t+1}\tilde{l}_{t+1} + R_{t+1}\tilde{b}_{t+1}}\right) + \tilde{\mu}_t, \\
\tilde{\mu}_t &= 0, \qquad \text{if} \qquad -\tilde{b}_{t+1} < \xi_t(P_t\tilde{h}_{t+1} + \lambda\tilde{k}_{t+1} + \tilde{l}_{t+1}).
\end{aligned}$$

  Given the aggregate states, individual state τ_t, and the next period returns, the above conditions form a system of five equations in five unknowns: $\tilde{h}_{t+1}$, $\tilde{k}_{t+1}$, $\tilde{l}_{t+1}$, $\tilde{b}_{t+1}$, $\tilde{\mu}_t$. We solve for the unknown variables using a nonlinear solver.

- The clearing conditions in the market for housing, $\sum_i^I H^i_{t+1} = 1$, and the market for bonds, $\sum_i B^i_{t+1} = 0$.

These conditions are used to update the guesses made in step 2a after which we restart the procedure from step 2b until the approximation error is sufficiently small. In the actual code these steps are performed jointly using a nonlinear solver that finds the numerical solution of the nonlinear system.

3. Using the solutions for $t = 1, ..., T$, we determine the coefficients of the approximated price function by estimating equation (20) using the data generated by the model solution for T periods. The estimated coefficients are

then used to update the parameters of the price function and the procedure is restarted from step 2 until convergence.

C Additional tables and figures

Table 20: Wealth mobility matrices for whole sample and sub-samples with housing debt and multiple houses. Linked surveys 2011-2013 and 2013-2015.

Full sample (2011-2013)			
	Bottom	Middle	Top
Bottom	74.3%	24.4%	1.3%
Middel	20.8%	55.8%	23.4%
Top	5.1%	11.4%	83.6%

With housing debt (2011-2013)			
	Bottom	Middle	Top
Bottom	61.9%	37.7%	0.4%
Middle	21.7%	45.5%	32.8%
Top	0.8%	10.9%	88.3%

With multiple houses (2011-2013)			
	Bottom	Middle	Top
Bottom	65.5%	34.5%	0.0%
Middle	17.7%	56.6%	25.7%
Top	3.0%	11.2%	85.8%

Full sample (2013-2015)			
	Bottom	Middle	Top
Bottom	78.7%	18.5%	2.8%
Middel	29.1%	53.8%	17.2%
Top	6.1%	19.0%	75.0%

With housing debt (2013-2015)			
	Bottom	Middle	Top
Bottom	75.7%	21.2%	3.1%
Middle	29.4%	50.8%	19.8%
Top	5.2%	17.8%	77.0%

With multiple houses, (2013-2015)			
	Bottom	Middle	Top
Bottom	77.3%	18.0%	4.7%
Middle	27.4%	56.8%	15.8%
Top	3.9%	16.1%	80.0%

Table 21: Distributional statistics

	Wealth		Income		Consumption	
	2015	2017	2015	2017	2015	2017
Gini index	0.65	0.66	0.59	0.56	0.45	0.43
Share top 20%	0.68	0.68	0.61	0.58	0.51	0.49
Share top 10%	0.52	0.50	0.46	0.42	0.36	0.33
Share top 5%	0.39	0.35	0.35	0.31	0.25	0.22
Share top 1%	0.19	0.12	0.18	0.14	0.10	0.09

Notes: The statistics are calculated on the full cross-sectional sample of 2015 and 2017, instead of the 2015-2017 panel sample.

Table 22: Additional statistics

	Obs	Age	2-home owner	1-home owner	Entrepr.	College	Tier-1 cities	House price rider	House buyer
Full sample	15,742	50.57	21.06%	71.82%	9.16%	12.89%	8.99%	17.30%	15.57%
Quintile 1	3,111	53.06	7.82%	70.40%	8.51%	6.26%	7.94%	2.54%	6.84%
Quintile 2	3,068	50.68	16.35%	79.92%	9.92%	11.05%	5.27%	3.94%	8.84%
Quintile 3	2,969	50.80	21.81%	75.60%	8.96%	15.31%	6.84%	7.21%	10.79%
Quintile 4	3,300	49.65	26.86%	69.90%	8.20%	16.08%	12.19%	27.17%	17.22%
Quintile 5	3,294	48.64	32.47%	63.31%	10.21%	15.73%	12.70%	45.62%	34.13%

	Wealth		Income		Consumption		Debt	
	2015	2017	2015	2017	2015	2017	2015	2017
Full sample	890,997	1,013,665	73,100	89,225	58,144	60,188	36,686	49,268
Quintile 1	915,890	182,634	59,240	55,337	52,850	45,499	33,683	45,398
Quintile 2	1,010,162	700,397	76,971	79,975	59,772	56,967	30,134	37,999
Quintile 3	974,514	1,054,948	81,521	95,537	60,233	63,399	31,987	39,323
Quintile 4	963,200	1,445,491	79,184	105,007	61,387	66,399	40,103	54,123
Quintile 5	591,226	1,684,440	68,585	110,261	56,479	68,675	47,514	69,487

Notes: Quintile sorting by growth rate of wealth. Bottom values are 2015-2017 averages.

Table 23: Wealth growth, sorting by marital status, education and age.

(a) 2013-2015

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Marriage status							
Single	2,234	8.2%	4.8%	19.4%	17.1%	8.2%	8.9%
Married	10,617	12.6%	7.3%	27.9%	19.0%	9.2%	9.8%
Education level							
Secondary and below	6,005	4.6%	2.2%	13.7%	17.6%	7.1%	10.5%
High school and equivalent	4,846	12.4%	7.2%	27.8%	18.7%	9.0%	9.7%
Bachelor and above	1,965	19.6%	11.7%	39.1%	20.2%	11.5%	8.8%
Age group							
Below 25	473	10.1%	7.4%	15.7%	17.4%	10.3%	7.0%
25-34	2,361	17.3%	10.2%	31.3%	22.6%	13.0%	9.5%
35-44	3,072	15.2%	9.8%	26.1%	20.8%	11.8%	9.1%
45-54	2,916	12.5%	8.0%	25.2%	18.0%	9.9%	8.1%
55-64	2,006	5.1%	1.3%	25.6%	14.7%	4.9%	9.9%
65 and above	2,023	5.7%	1.3%	27.6%	15.6%	0.9%	14.7%

(b) 2015-2017

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Marriage status							
Single	2,862	3.9%	-1.1%	29.4%	17.2%	8.1%	9.1%
Married	12,880	13.4%	7.2%	31.7%	19.6%	9.4%	10.1%
Education level							
Secondary and below	7,984	4.2%	0.6%	19.4%	18.1%	7.6%	10.6%
High school and equivalent	5,567	11.7%	5.5%	32.9%	18.8%	8.8%	10.0%
Bachelor and above	2,185	26.1%	16.0%	45.9%	21.9%	12.9%	9.0%
Age group							
Below 25	330	-10.8%	-18.3%	37.7%	19.9%	11.9%	8.0%
25-34	2,356	17.7%	8.3%	39.1%	24.1%	14.6%	9.5%
35-44	3,249	19.8%	13.0%	30.8%	22.1%	12.2%	10.0%
45-54	3,782	12.5%	6.7%	30.3%	18.9%	10.8%	8.1%
55-64	2,798	8.2%	3.7%	28.0%	16.2%	6.0%	10.1%
65 and above	3,227	2.4%	-1.5%	26.9%	14.5%	1.0%	13.5%

Table 24: Wealth growth across households sorted by initial wealth.

(a) 2011-2013

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	3,705	19.4%	14.4%	26.8%	18.9%	11.9%	7.0%
Quintile 1	682	95.1%	80.3%	11.9%	124.2%	81.2%	43.0%
Quintile 2	701	74.8%	64.3%	19.9%	53.0%	35.0%	17.9%
Quintile 3	755	52.1%	44.8%	21.9%	33.7%	22.3%	11.4%
Quintile 4	827	42.7%	35.7%	28.7%	24.3%	14.0%	10.3%
Quintile 5	740	2.2%	-1.3%	34.7%	10.2%	6.4%	3.8%

(b) 2013-2015

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	12,851	11.8%	7.2%	24.8%	18.5%	11.5%	7.1%
Quintile 1	2,811	123.7%	116.5%	8.2%	88.0%	52.4%	35.6%
Quintile 2	2,425	52.3%	44.6%	19.1%	40.7%	25.2%	15.5%
Quintile 3	2,328	26.4%	20.0%	22.2%	29.1%	17.7%	11.3%
Quintile 4	2,393	20.3%	14.8%	26.9%	20.7%	12.6%	8.1%
Quintile 5	2,894	-2.8%	-6.2%	32.2%	10.6%	6.8%	3.8%

(c) 2015-2017

	Obs	g_t^W	g_t	s_t	r_t^W	r_t^L	r_t^K
Full sample	15,742	13.8%	8.0%	30.0%	19.1%	11.8%	7.3%
Quintile 1	3,134	83.8%	77.2%	7.6%	86.5%	51.9%	34.6%
Quintile 2	2,836	42.4%	32.3%	23.2%	43.1%	26.0%	17.1%
Quintile 3	2,837	23.3%	15.0%	27.3%	30.4%	18.9%	11.5%
Quintile 4	3,106	22.5%	15.6%	32.0%	21.5%	13.2%	8.3%
Quintile 5	3,829	3.5%	-0.9%	38.7%	11.3%	7.1%	4.2%

Table 25: The growth rate of each asset component

(a) 2013-2015

	obs	asset	hs-asset	fin-asset	bus-asset	oth-asset	debt	hs-debt
Full sample	12,851	12.47%	12.78%	40.00%	48.07%	-25.80%	19.82%	16.05%
Quintile 1	2,749	-61.36%	-58.24%	-45.49%	-49.32%	-65.30%	39.18%	30.15%
Quintile 2	2,520	-15.85%	-12.90%	-6.74%	1.38%	-42.50%	10.31%	12.02%
Quintile 3	2,459	10.51%	10.20%	35.32%	19.61%	-21.59%	3.87%	1.06%
Quintile 4	2,542	46.79%	39.81%	98.12%	139.99%	1.80%	2.93%	-4.59%
Quintile 5	2,581	176.06%	173.29%	165.68%	305.80%	32.60%	47.55%	56.79%

(a) 2015-2017

	obs	asset	hs-asset	fin-asset	bus-asset	oth-asset	debt	hs-debt
Full sample	15,742	14.83%	18.83%	17.38%	-31.07%	9.91%	34.30%	39.65%
Quintile 1	3,111	-75.00%	-77.54%	-51.10%	-77.97%	-49.81%	34.78%	31.53%
Quintile 2	3,068	-28.58%	-25.40%	-28.89%	-55.81%	-21.19%	26.10%	37.20%
Quintile 3	2,969	8.58%	10.62%	9.68%	-25.78%	12.68%	22.93%	28.77%
Quintile 4	3,300	49.47%	51.09%	49.63%	1.87%	43.43%	34.96%	35.90%
Quintile 5	3,294	173.50%	173.20%	127.19%	91.84%	97.66%	46.25%	54.17%

Notes: Sorted by the growth rate of wealth. Variables: total assets, house assets, financial assets, business assets, other assets, total debt, house debt.

Table 26: Two-year house price growth rate (district level)

	2011-2013		2013-2015		2015-2017	
	Obs	Growth	Obs	Growth	Obs	Growth
Full sample	301	5.6%	338	4.4%	915	32.7%
Quintile 1	61	-11.4%	68	-20.6%	183	-7.3%
Quintile 2	60	-1.3%	68	-5.3%	183	8.7%
Quintile 3	60	2.6%	67	0.6%	183	20.7%
Quintile 4	60	8.8%	68	9.9%	183	38.5%
Quintile 5	60	29.6%	67	37.9%	183	103.1%

Notes: This table reports the two-year district/county level house price growth rate. In China, district or county is an administrative division under a city. For example, Beijing has 16 districts. We collect the district level house price data from Lianjia. For each year, we calculate the house price growth rate for each district, and sort districts into 5 quintile groups based on growth rate. Then, we calculate the average growth rate for each quintile group.

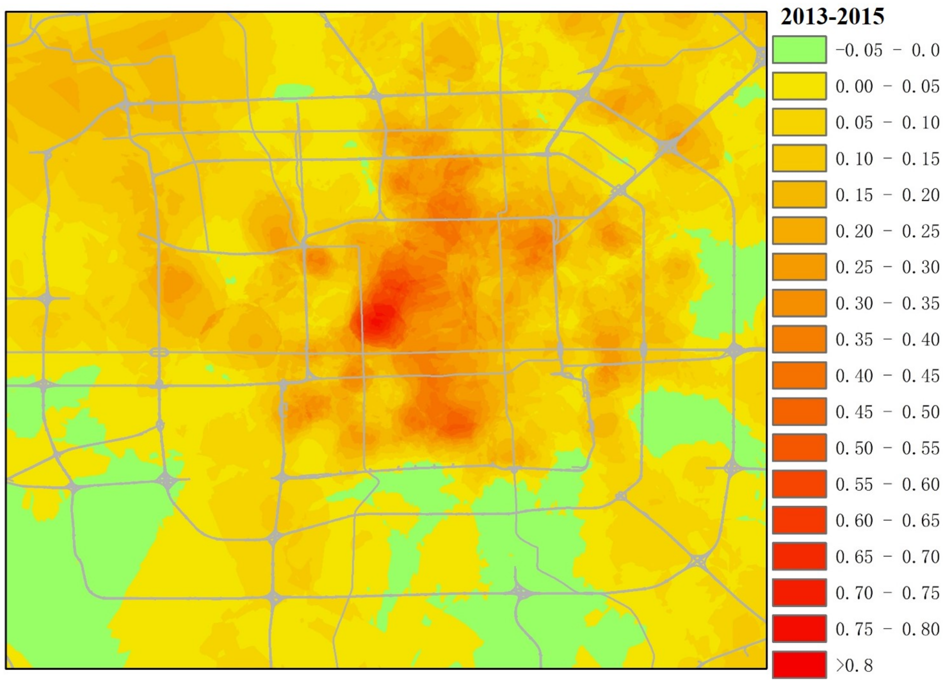

Figure 2: Community level house price growth rate in Beijing (from transaction data of Lianjia).

Notes: The figure presents the two-year housing price growth rate during the period 2013-2015 across communities in the city of Beijing. Warm colors indicate positive growth rates, while green color represents negative growth. The grey circles represent the highways of Beijing. The data is from the largest real-estate brokerage company Lianjia (similar to Zillow in the US). In 2015, there were around 70,000 housing transactions across 3,000 communities in Beijing. Based on the transaction data, we first calculate the average housing price for each community and then compute the two-year growth rate.

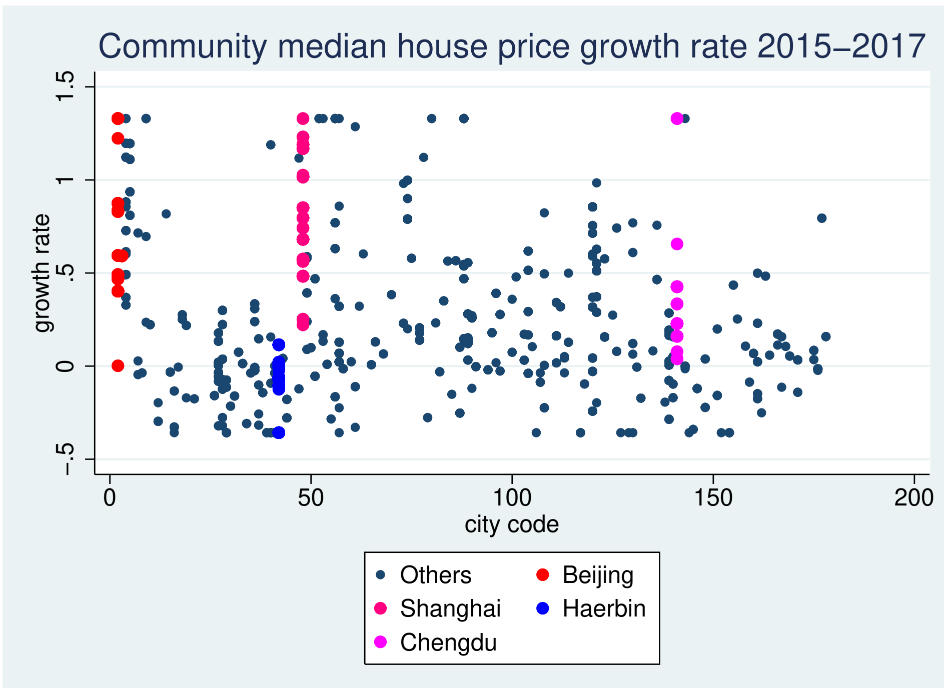

Figure 3: Community level median house price growth rate (calculated from our data sample)

Notes: This figure presents the two-year housing price growth rate during the period 2015-2017 across communities in our CHFS data. The y-axis represents the two-year growth rate of housing price, and the x-axis represents the city code. Each city code encompasses multiple communities arranged horizontally along the y-axis. Each dot represents a community. We also mark the communities in four selected cities with different colors. For example, we used the red color to highlight Beijing, and the blue color for Haerbin, a city in northeast part of China. We first calculate the median housing price for each community and then compute the two-year growth rate.