Migration, Housing Constraints, and Inequality: A Quantitative Analysis of China [*]

Min Fang
University of Rochester [†]

Zibin Huang
University of Rochester [‡]

This Version: February 7, 2021
First Version: April 30, 2020
click here for most recent version

Abstract

We investigate the role of migration and housing constraints in determining income inequality within and across Chinese cities. Combining microdata and a spatial equilibrium model, we quantify the impact of the massive spatial reallocation of workers and the rapid growth of housing costs on the national income distribution. We first show several stylized facts detailing the strong positive correlation between migration flows, housing costs, and imputed income inequality among Chinese cities. We then build a spatial equilibrium model featuring workers with heterogeneous skills, housing constraints, and heterogeneous returns from housing ownership to explain these facts. Our quantitative results indicate that reductions in migration costs and the divergent growth in productivity across cities and skills result in the observed massive migration to developed areas. Combined with tight land supply policies in big cities, the expansion of the housing demand causes the rapid growth of housing costs and increases the inequality between local housing owners and migrants. The counterfactual analysis shows that a migration-based land supply reform with regional transfers or a US-level property tax both lower within-city income inequality, by 34% and 21%, respectively. Meanwhile, both reforms lower national income inequality by 20%. However, only the land supply reform encourages more workers to migrate to higher productivity cities.

Keywords: Income Inequality, Migration, Cities, Housing Constraints, China;
JEL Classification Numbers: E24, J61, R23, R31;

---

[*]We want to send special thanks to Junsen Zhang for generously providing us some of the data, and to the River Campus Libraries at University of Rochester for generously providing us a Dataset Grant. We thank George Alessandria, Yan Bai, Mark Bils, Jingting Fan (discussant), Kaixin Liu, Dan Lu, Ming Lu, Ronni Pavan, Lisa Kahn, Narayana Kocherlakota, and Tianchen Song for their comments and suggestions. We also thank other audiences at the University of Rochester, ASSA/AEA 2021 Annual Meeting, Urban Economics Association 2020 Annual Meeting, Chinese Economist Society 2020 Annual Meeting, Seminars in Economic Geography, Junior Migration Webinar, and 7th International Conference on the Chinese Economy: Past, Present and Future for their valuable comments. All errors are ours.

[†]Email: min.fang.ur@gmail.com. Department of Economics, Harkness Hall, Rochester, NY 14620.

[‡]Email: zibin.huang@rochester.edu. Department of Economics, Harkness Hall, Rochester, NY 14620.

1 Introduction

As documented by Piketty, Yang, and Zucman (2019), Chinese income and asset inequality rose from a level similar to that of the Scandinavian countries to approaching that of the United States. More importantly, much of this rise was driven by uneven ownership of the dramatically appreciating housing assets in developed cities. This housing boom in developed cities has been accompanied by massive inflows of migrant workers as well as tightening housing constraints.[1] Could this massive migration inflow and tight housing constraints in these developed cities explain rapidly rising inequality in China? If this is true, is there any policy we could implement to alleviate this rising inequality?

In this paper, we take two approaches to answer both questions. First, using administrative data, we document that housing costs and income inequality are significantly positively correlated with the number of migrant workers in different cities. Second, we construct a spatial equilibrium model incorporating both heterogeneous wage income and heterogeneous housing asset income to quantify the effects of migration and housing constraints on the observed income inequality. In the counterfactual analysis, we find that easing the housing constraints in developed cities can reduce income inequality in China.

In the first step, we show three main stylized facts from the data. First, migration in China is highly concentrated into developed areas. The concentration is accelerating across time because of improvements in the transportation system and the relaxation of the Hukou system.[2] Second, housing costs have increased drastically over time, especially for cities with large numbers of migrants. There is a positive correlation between housing costs and the net stock of migrant workers across cities. Third, income inequality within cities is positively correlated with the net stock of migrant workers. We show two additional facts as supplement. First, wage inequality within cities is not correlated with the net stock of migrant workers. This implies that the positive correlation between the within-city income inequality and the number of migrant workers is due to housing ownership rather than wages. Second, cities with more migrant workers contribute more to national income inequality. These five stylized facts give us a preliminary picture of the whole story. As the economy grows, more and more migrants concentrate in large and developed cities in coastal areas. The massive increase in housing demand, together with the highly regulated land supply, push up real estate rents in big cities and benefit all local housing owners. Increasing housing rents then translate into increasing income inequality between local property

[1]China has experienced impressive economic growth over the last four decades after the start of economic reforms and opening-up in 1978. This triggered a massive wave of migrant workers moving from under-developed areas to developed areas. There was also a huge housing boom, especially in large cities such as Beijing and Shanghai. Housing prices increased by 660% from 2003 to 2013 in Beijing (Fang et al., 2016), which can partly be attributed to the tight land supply regulatory policy.

[2]The Hukou system is a unique household registration system. In China, each household has to register in the place where they are initially from, and it is hard to change the registration locale during one's lifetime. The Hukou system is closely related to access to public services. For instance, a family migrating from Henan to Shanghai may not be able to send their children to public schools in Shanghai. For more details, please refer to Song (2014).

owners and migrant renters.

In the second step, we construct a spatial equilibrium model to quantify all the facts, explain the mechanism, and conduct counterfactual analysis. The model comprises heterogeneous workers making migration choices, a representative firm, and a state-regulated housing sector in each city. The key mechanism is that with the universal drop in migration costs and uneven productivities, workers migrate from under-developed cities to developed cities with higher wages. Since housing supply is heavily regulated and inelastic in these developed cities, housing costs increase dramatically which drives up local property owners' housing income. However, due to various frictions, a migrant worker could not participate in local property market and sharing in any increases in home values. They could only earn housing asset returns in his under-developed home city. As more and more migrants move into these developed cities, the housing ownership gap between locals and migrants results in the rapidly rising income inequality.

Using administrative data from 2005 and 2010, we solve the model quantitatively. We find that from 2005 to 2010, the average migration costs decreased by 35% for low-skill workers and 21% for high-skill workers. Meanwhile, productivity growth is faster in absolute terms in large cities which attracts large numbers of migrant workers. These large developed cities also have slower growth growth in land supply. Construction land supply increased by only 10% in the largest cities which together attracted more than twenty million workers, while at the same time, average land supply growth was 40% for cities which lost almost half of their working population. This inefficient land supply policy causes severe housing constraints in developed cities and increases income inequality.[3]

Finally, we conduct counterfactual policy reforms to ease housing constraints in developed cities and reduce income inequality. The main counterfactual we impose is a migration-based land supply reform. The idea is straightforward and intuitive: to allocate new land quotas by migration flows. We reallocate the increment of the total land supply quota from 2005 to 2010. Instead of giving more land quota to under-developed areas, we allocate land quota proportionally to the change in the migration inflows to different cities, while keeping the national total land supply constant. That is, cities attracting more migrants are given more quota. Meanwhile, all revenues from additional lands in developed cities are collected and transferred to under-developed cities who lost quota as compensation. This policy mimics a "land quota transfer market" (Lu, 2016) where developed cities can buy land quota from under-developed cities and compensate them with direct transfers. Thus, we can achieve balanced development between regions and simultaneously avoid policy distortion. With the counterfactual land supply policy, housing cost increases in big cities are significantly attenuated. Compared with the real world, this policy can reduce housing costs in 2010 by 30% in first-tier cities and by 25% in second-tier cities and also incentivizes more workers to migrate to these developed cities. Simultaneously, within-city income inequality falls

[3] There is a quota of land for construction usage in each city. The quota is determined by the central government and utilized as a tool to balance development across different regions. Thus, under-developed western regions get much more construction land than they need while land supply is severely suppressed in developed eastern regions. This potential policy distortion creates a substantial spatial misallocation, as suggested by Hsieh and Moretti (2019).

by 34% and national level inequality by 20%. In another counterfactual policy, we show that a US-level property tax and redistribution policy could also help to reduce income inequality.

Literature Review. Our study extends the current literature in three dimensions. First, we investigate a new mechanism for income inequality and extend knowledge about increasing inequality in China. There are many studies on income and wage inequality. Different papers investigate many causes of inequality, including skill-biased technological changes and the increase in the return to human capital (Berman, Bound, and Machin, 1998; Card and DiNardo, 2002; Moore and Ranjan, 2005), education inequality (Gregorio and Lee, 2002; Sylwester, 2002), trade liberalization (Goldberg and Pavcnik, 2004; Han, Liu, and Zhang, 2012; Verhoogen, 2008), and privatization (Chao, Hazari, and Yu, 2006; Cuadrado-Ballesteros and Peña-Miguel, 2018). The closest study to ours is Chen, Liu, and Lu (2018). They find that larger cities have higher income inequality and claim that it is because migration inflows into larger cities change the skill composition of the workers, yielding a higher skill premium. In this study, we investigate a new mechanism of migration interacting with housing constraints that can also increase income inequality.

Second, this paper contributes to the literature that studies the spatial distribution of labor supply using the EK-Migration framework. Since Ahlfeldt et al. (2015), the literature extends the canonical Eaton and Kortum (2002) international trade framework to introduce worker mobility to explicitly model worker location choices in the presence of migration costs and heterogeneous worker preferences regarding locations. Many of them investigate internal migration costs, such as Morten and Oliveira (2014), Bryan and Morten (2019), Tombe and Zhu (2019), and Fan (2019). The closet studies to us are Tombe and Zhu (2019) and Fan (2019). The former focuses on how trade and migration costs affect labor productivity in China without differentiating between worker skill types, and the latter focuses on understanding how international trade affects overall domestic wage inequality and the aggregate skill premium without considering the distribution of property ownership. Our paper aims to understand income inequality stemming from both human capital and wealth ownership differences. Guided by this target, our model is extended to introduce high/low-skill workers and heterogeneous housing ownership. Second, instead of inferring wages from the model, which is the most important ingredient for calculating inequality, we manually collect the wages by industry for as many Chinese cities as we can from individual city statistical yearbooks. Combining this unique dataset with the population census, we construct a comprehensive and spatially decomposable inequality measure for China and investigate the most realistic policy reforms.

Third, this paper contributes to the literature that studies the housing and land market in China. The so-called *Great Housing Boom* of China is well documented in Garriga et al. (2020), Garriga et al. (2017), Fang et al. (2016), Chen and Wen (2017), and Glaeser et al. (2017). The housing boom is unevenly distributed spatially. As Fang et al. (2016) shows, the boom is not universal. More developed cities have seen disproportional gains in housing prices while less developed cities actually see their housing prices grow more slowly than GDP. Various theories attempt to explain this pattern: Garriga et al. (2017), Liang, Lu, and Zhang (2016), and Wu, Gyourko, and Deng (2016). We contribute to this literature by showing that the inefficient land supply policy and massive migration inflows into the larger cities jointly caused the *Great Housing Boom* of China. We also provide counterfactual policies for the housing sector, which could lower housing costs and reduce housing inequality.

Layout. This paper is organized as follows. Section 2 describes the data and variables. Section 3 documents five stylized facts of migration, housing, and inequality in China. Section 4 shows the spatial equilibrium model. Section 5 quantifies the model and shows model results. Section 6 shows counterfactual policy reforms. Section 7 concludes.

2 Data and Variables

2.1 Data Sources

In this study, we need a comprehensive dataset that records an individual's Hukou registration location, their current work location, wage earnings, occupation, housing ownership, and rent payment. Our interest in housing costs and spatial inequality implies that the data must be geographically representative. Moreover, since we want to estimate migration elasticities, the dataset must be large enough to record flows between all pairs of locations. Only the *Chinese Population Census* (*Census* for short) meets all these specifications. We also supplement the *Census* with the *City Statistic Yearbooks* and the *Urban Statistic Yearbooks* for city-level aggregate variables. We introduce these datasets sequentially below in depth.

We use the *Census* as the main dataset in this study. It is the most comprehensive household-level survey in China. It is conducted every ten years, and all residents in mainland China are surveyed. In the survey, 90% of the households report only basic demographic information, including their Hukou registration location and current living location. The other 10% of the households take a so-called "long-survey," which asks additional questions including items dealing with housing conditions, housing rents, and job details. Midway between two *Census* years, a *Mini-census* is conducted. The *Mini-census* randomly selects 1% of the population and asks a list of questions similar to the ones in the long survey of the decennial *Census*. In this study, we use the decennial *Census* in 2010 and the *Mini-census* in 2005 to calculate city-level migration flows and housing rents for individuals with different education levels.[4] In our sample, we have 2,585,481 observations in the year 2005, which covers 0.2% of the Chinese population. Additionally, we have 4,803,589 observations in the year 2010, which covers 0.36% of the population.

We supplement the *Census* data with the *City Statistic Yearbooks* and the *Urban Statistic Yearbook*. *City Statistic Yearbooks* are books including socioeconomic data for specific cities. Each city has its own yearbook, and the data is collected by the local branch of the National Bureau of Statistics. We derive industry level average wages in each city from these yearbooks. They will be used to impute the city-skill level wages as we will explain in the next section. The *Urban*

[4]From now on we call the decennial *Census* and *Mini-census* as simply the *Census* in general for conciseness.

Statistic Yearbook is a book with a summary of key economic indicators across all Chinese cities in a specific year. We derive the city-level GDP growth rate and the constructed land area data from it.

2.2 Impute City-skill Level Wages

In the model part of this study, we need average wages for different skills (education levels) in different cities in 2005 and 2010. However, there is no data directly showing average high-skill (college-educated) wages and average low-skill (not college-educated) wages in each city. Ideally, if we have wages for all individuals in the *Census* data, we can calculate city-skill level average wages as:

$$w_j^s = \frac{1}{N_j^s} \sum_s w_{ij}^s \tag{1}$$

where w_j^s is the average wage of workers with skill s in city j. N_j^s is the number of workers in city j with skill s. w_{ij}^s is the wage of individual worker i with skill s, working in city j. However, the *Census* data contains wage information only for the year 2005 but not 2010. Fortunately, in the *City Statistic Yearbooks* of each city, they have average wages in different industries in this city. In addition, in the *Census* data, there is information about an individual's education and industry. Thus, we can first impute an individual's wage by using the average wage in the industry-city the individual is working in. Then we use equation (1) to calculate the city-skill level average wages. In essence, what we do is to calculate city-skill level wages using average city-industry wages, weighted by the number of workers with different education levels in each industry. Since the *City Statistic Yearbooks* are published separately by different local governments, we have to manually collect over 600 books for 2005 and 2010. There are some cities for which we cannot find for the exact years of 2005 and 2010. We replace these missing years by the closest year we could find and impute the wages using city-level GDP growth rates.[5] The replacement is less than 5% of the observations.

There is another concern that the wages from *City Statistic Yearbooks* may not be representative since the National Bureau of Statistics usually does not include informal jobs when it collects the data. Thus, we try another imputation of the city-skill level average wages to check the robustness of our results. In this method, we directly use individual level wages in 2005 and calculate city-skill level wages using equation (1) in 2005. We then impute the city-skill level wages in 2010 by multiplying the city-skill level wages in 2005 with GDP growth from 2005 to 2010 in each city. We repeat all the analysis using this method and the results are robust. They are available upon request.

[5]For example, if we cannot find the *City Statistic Yearbook* of Beijing in 2005 and can only find the one in 2004, then we use city-industry level wages for Beijing in 2004 and multiply them by Beijing's GDP growth rate in 2005 to estimate city-industry level wages for Beijing in 2005.

3 Stylized Facts: Migration, Housing, and Inequality

From our data, we calculate the net stock of migrant worker, the skill share, and the housing costs in each city. We then calculate within-city inequality for each city and nationwide inequality. From these observations, we document three major and two additional supplementary stylized facts of migration, housing and inequality in China.

Fact 1: Migration is highly and increasingly concentrated in certain large cities.

To document Fact 1, we calculate the net stock of migrant workers and the share of the net stock of migrant workers across all Chinese cities in 2005 and 2010, respectively. The net stock (in numbers N) and share of the net stock (in percentage %) for city j are calculated as follows:

$$\text{Net Stock}_j(N) = \text{Current Workers}_j - \text{Hukou Workers}_j$$
$$\text{Net Stock}_j(\%) = \frac{\text{Current Workers}_j - \text{Hukou Workers}_j}{\text{Hukou Workers}_j}$$

where Current Workers$_j$ is the total number of workers who are currently working in city j, and Hukou Workers$_j$ is the total number of workers whose Hukou registration is located in city j. Therefore, the net stock reflects the net gain or loss of the working population of each city and the share of net stock reflects the net gain or loss proportional to the Hukou registered working population of each city.[6] The former measure avoids the potential outliers when measuring in percentages and the latter measure avoids the potential city size effect when measuring in absolute numbers.

Table 1: **Distribution of Net Stock of Migrant Workers**

Panel A: Net Stock (measured in numbers, Unit: million)											
Year	No.	(-4,-2)	(-2,-1)	(-1,-0.5)	(-0.5,0)	(0, 0.5)	(0.5,1)	(1,2)	(2,4)	(4,8)	(8+)
2005	287	1	1	23	188	59	4	4	4	2	1
2010	266	6	29	41	115	39	9	13	7	3	4
Panel B: Net Stock (measured in percentage, Unit: %)											
Year	No.	(-80, -45)	(-45,-30)	(-30,-15)	(-15,0)	(0, 15)	(15,30)	(30,45)	(45,60)	(60,75)	(75+)
2005	287	0	11	63	139	48	9	5	3	3	6
2010	266	12	47	61	71	19	17	14	6	4	15

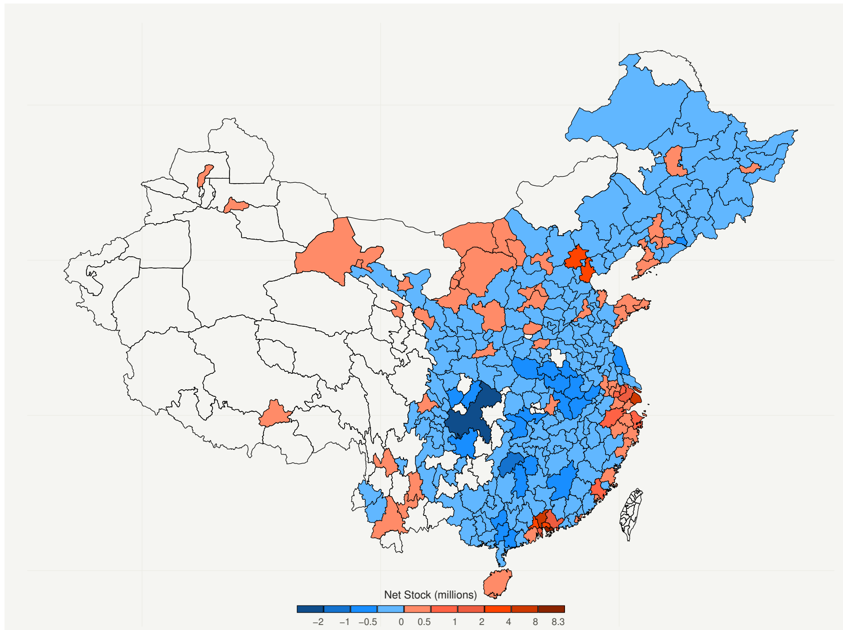
Notes: This table shows the distribution of the net stock of migrant workers across Chinese cities. There are 287 and 266 cities available in 2005 and 2010, respectively in our data. Panel A is the standard of coloring in the Figure 3 map. Panel B is the standard of coloring in the Figure 7 map.

First, migration is highly concentrated in certain large cities. Table 1 shows the distribution of the net stock of migrant workers across cities measured in both absolute numbers and percentages.

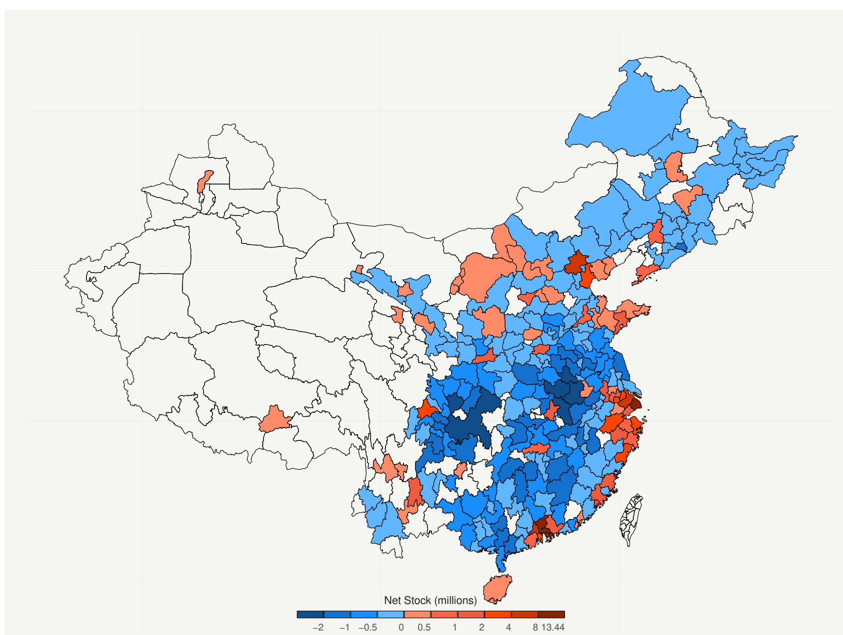
[6] We do not choose a percentage measure as $\{(\text{Current Workers}_j - \text{Hukou Workers}_j)/\text{Current Workers}_j\}$ which is strictly bounded between 0 and 1 because we do not just want to capture the relative share of migrant workers among working populations in each city. We emphasize each city's net gain or loss relative to its Hukou registrations.

Figure 1: Net stock of migrant workers by city in China

(a) Net stock of migrant workers in 2005



(b) Net stock of migrant workers in 2010



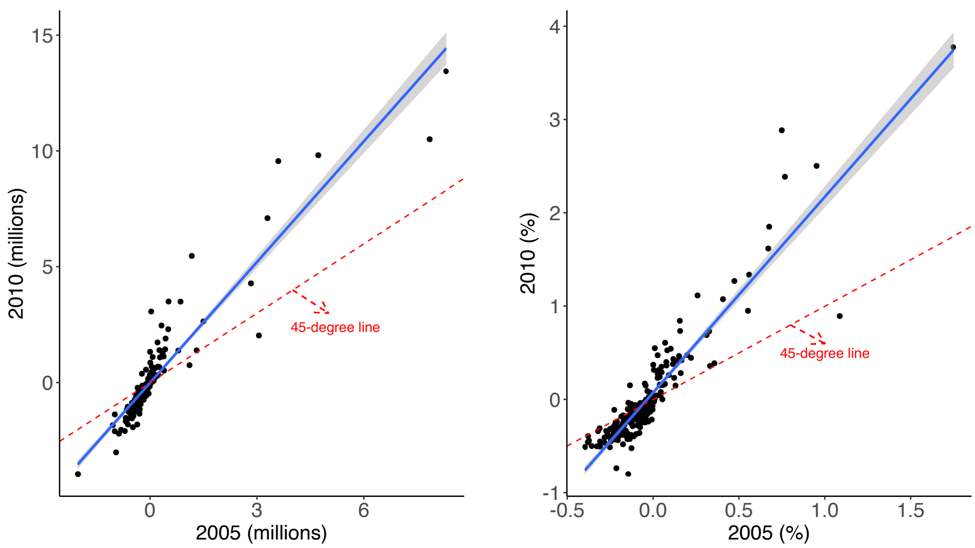
Notes: The sample only includes workers with wage income, which means that we exclude retired workers, persistently unemployed workers (zero wage income for the whole year), children, students, homemakers, and others. The net stock of workers in city i is calculated as current workers in city i minus Hukou workers in city i. Therefore, this measure reflects the net gain in the working population for each city. We only have data on 287 and 266 cities in 2005 and 2010 respectively. Though the blank parts are missing, our available data covers more than 95% of the Chinese population.

To visualize the migration patterns, we also geographically plot the *Net Stock(N)* by cities in both 2005 and 2010 in Figure 3.[7] Each interval in the table means the range of the net migration. For instance, in 2010, there are 34 cities with a net stock of more than 8 million migrants. Most cities lose workers, and only about one-fourth of cities have positive net stocks. From the map, it is obvious that workers are migrating from western and central regions to eastern regions, and from inner-land cities to coastal cities.[8]

Second, migration is increasingly concentrated in certain large cities. As the colors in Figure 3 indicate, the concentration of migration has grown during these five years. From 2005 to 2010, inland cities lose more workers and big eastern cities gain more. To provide more intuitive result, we also plot the correlation between the net stock of migrants in 2005 and in 2010 in Figure 2. The red dashed line is the 45-degree line. The fitted line has a slope much larger than one, and the big cities with a net gain of workers in 2010 are all above the 45-degree line, which means that the concentration is rapidly increasing over these five years.

Figure 2: Correlation of Net Stock of Migrants in 2005 and 2010

(a) Net Stock (Numbers N) (b) Net Stock (Percentage %)



Notes: In plot (b), 1 means 100%. The percentage plot (b) excludes two outlier new cities Shenzhen and Dongguan which do not fit the scale of the plot but still fit the pattern. Both cities were established in the 1980s. Because of low initial stocks of Hukou population and high appeal to migrants, both cities have *Net Stock (%)* measures larger than 500% in 2005, growing to larger than 1000% in 2010.

Fact 2: Housing cost increases drastically with net stock of migrants and across time.

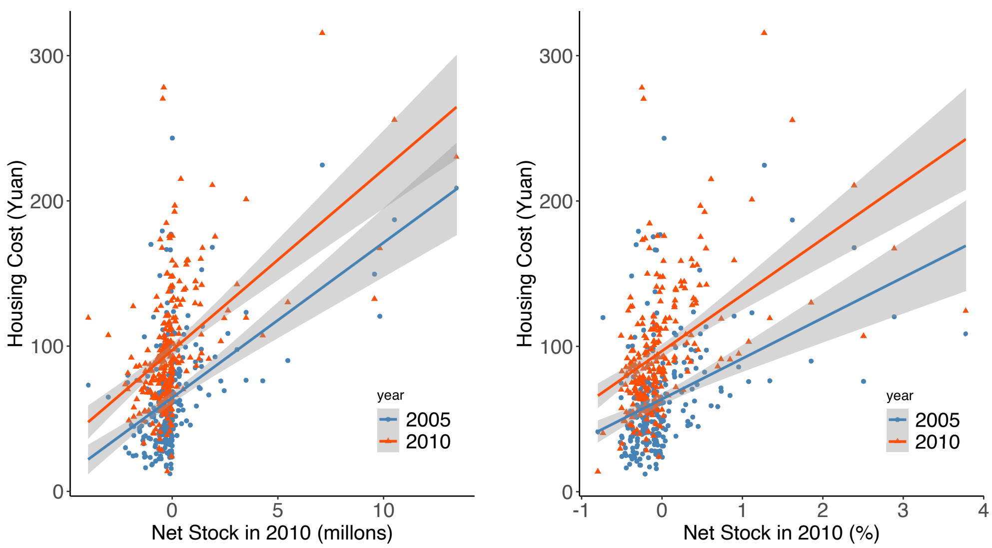
[7] For the sake of space, we show the map of *Net Stock (%)* in the appendix A.1.

[8] Most of the big industrialized cities are located along the eastern coastline. There are four main economic zones where cities with huge number of migrant workers concentrate: (1) the Bohai Economic Rim, led by Beijing and Tianjin; (2) the Yangtze River Delta Zone, led by Shanghai, Suzhou, and Hangzhou; (3) the Western Taiwan Straits Zone, led by Xiamen; (4) the Pearl River Delta Zone, led by Guangzhou (Canton), Shenzhen, and Hong Kong.

To document Fact 2, we calculate housing costs for each city in both 2005 and 2010 using micro data from the *Chinese Population Census*. We first annualize individual housing rent per square meter, then take the average for each city. One concern here is that cities may systematically differ in their housing qualities since we are using raw rents. We addressed this concern in appendix A.2 using quality-adjusted housing rents and migration inflows. The results are robust.

Figure 3: Net Stock of Migrants and Housing Cost

(a) Net Stock (Numbers N) (b) Net Stock (Percentage %)



Notes: Housing cost is measured as rent per square meter using the micro data from the *Chinese Population Census*. In plot (b), 1 means 100%. The percentage plot (b) excludes two outlier new cities Shenzhen and Dongguan which do not fit the scale of the plot but still fit the pattern. Both cities are established at the end of 1980s. Because of low initial stocks of Hukou population and high attractions to migrants, both cities have their *Net Stock (%)* measures larger than 1000% in 2010 and almost the highest housing costs among all Chinese cities.

Figure 5 plots housing costs against the net stock of migrant workers for both 2005 and 2010 in both absolute number and percentage measures, respectively. Red dots are values in 2010 and blue dots are values in 2005. We fix the x-axis using the 2010 value for each city so we can easily compare changes in housing costs across cities over the five year period. For instance, the highest dot (>300 Yuan) in sub-figure (a) is Beijing's average housing cost in 2010; we can then easily identify Beijing's average housing cost in 2005 as roughly 220 Yuan right below the highest dot. We keep this plotting format for all the plots in the rest of this paper.

Figure 5 delivers two messages. First, housing costs increase drastically with the net stock of migrants. It is clear that the net stock of migrants is positively correlated with housing rent costs. The fitted lines for both net stock measures and both years are significantly upward sloping. A simple regression suggests that increasing the migrant inflow by one million workers is associated with a 17.9 RMB (about 2.6 USD) increase in the annual average housing rent per square meter in 2005. The corresponding number in 2010 is 13.4 RMB (about 1.9 USD).

Second, housing costs increased drastically over time. The average living space for tenants in 2005 and 2010 is 43.3 and 43.9 square meters. Thus, according to sub-figure (a), on average, one million new migrants is associated with a 775.1 RMB (about 110.7 USD) or a 588.3 RMB (84.0 USD) increase in annual housing rent in 2005 and 2010, respectively. In addition, the national average housing rent increased sharply from 288.6 RMB per square meter to 460.5 RMB per square meter, which corresponds to a 60% increase. The pattern in the plot and these numbers indicate that housing costs increase drastically with the net stock of migrants and across time.

Fact 3: Income inequality increases drastically with the net stock of migrants and across time.

To document Fact 3, we calculate the Theil Index[9] for total income at city-level for all Chinese cities in 2005 and 2010, respectively. Each worker's income consists of wage income and capital income. For wage income, we directly take the imputed city-skill wages as each individual workers' wage income, which is explained in section 2.2.[10] For capital income, however, there is no available data for each city. Therefore, we adopt a lower bound imputation through two compromises using the *Census* data.

Our first compromise is to assume housing assets are the only assets. Housing accounts for the absolute majority (74.2%) of total assets in Chinese families[11] and families with more housing assets usually own more financial assets. This compromise potentially underestimate income inequality due to the exclusion of financial assets. Our second compromise is to assume that housing returns are totally captured by the flow of rent income (homologous to the dividend of an always fairly priced stock). Since housing prices in larger and more developed Chinese cities have increased much faster than rents, this compromise also potentially underestimates the housing income of property owners in these larger and more developed cities. In general, our inequality measure is a lower bound and focuses on housing assets.

Conditional on these two compromises, we calculate housing income for local workers owning houses by multiplying the size of their houses with city-level average rent divided by family sizes. Then we take the average of this housing income for each city and attribute it to the local residents who own houses. Thus, an individual worker's total income is the sum of the city-skill wage and

[9] We also try the traditional Gini Index, and the results are robust. We use the Theil Index because it can be easily decomposed into small groups. Specifically, a national level Theil Index can be decomposed into two terms. The first term is a weighted average of Theil Index scores for city-level mean (inequality across cities). The second term is a weighted average of the Theil Index of individuals within different cities (inequality within cities). Therefore, it is natural to calculate the contribution of each city to national level inequality (Novotnỳ, 2007).

[10] One concern is that when we use city-skill level imputed wages, we may erase a large portion of heterogeneity. Thus, we also check the results when using real individual level wages in 2005. The results are robust, consistently finding that cities with more migrant workers have higher inequality. We stick with imputed city-skill level wages in the main context for two reasons. First, there is no real individual wage data available in the 2010 *Census*. Second, we want to present data that is the most consistent with the model part.

[11] This is according to a report by the Central Bank of China. Please refer to http://paper.people.com.cn/zgjjzk/html/2020-05/15/content_1987791.htm (in Chinese). An average urban family owns 1.5 unit of houses/apartments and only 43% of Chinese families carry mortgage.

the imputed housing asset income (including both self-consumed housing and actual rent income from migrant renters). This construction is more consistent with our model.[12]

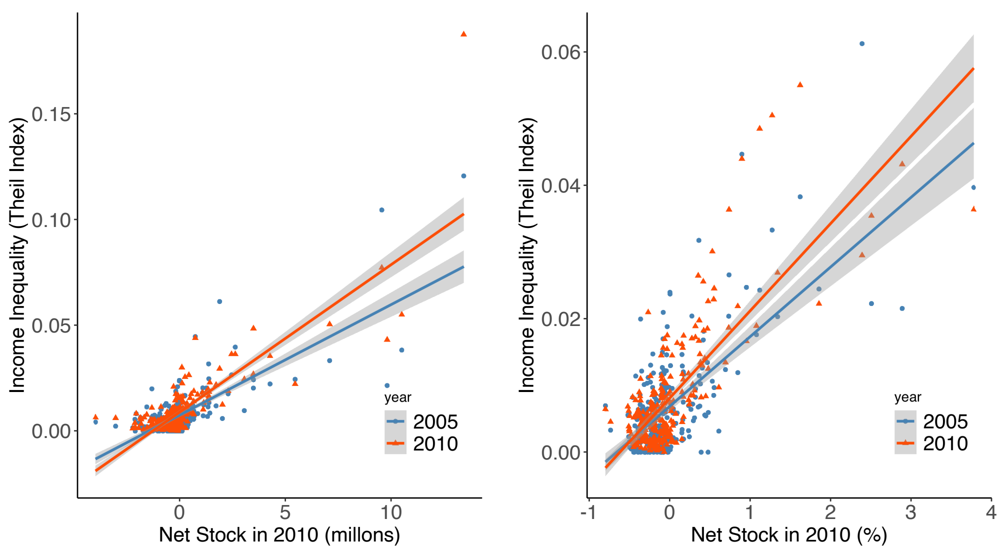
The city-level Income Theil Index for city j is then:

$$T_j^{\text{Inc}} = \frac{1}{N_j} \sum_{n=1}^{N_j^s} \sum_{s=1}^{S} \frac{i_{jn}^s}{\bar{i}_j} ln \frac{i_{jn}^s}{\bar{i}_j}, \quad i_{jn}^s = w_{jn}^s + \text{housing asset income}_{jn} \tag{2}$$

where j, n, s indicate city, worker, and skill, respectively. N_j is the total number of current workers in city j, N_j^s is the total number of current workers with skill s in j, $\bar{i}_j$ is the average income in city j, i_{jn}^s is the income of each individual n with skill s in city j, and w_{jn}^s is the wage of each individual n with skill s in city j.

Figure 4: Net Stock of Migrants and Income Inequality

(a) Net Stock (Numbers N) (b) Net Stock (Percentage %)



Notes: Income Theil Index for each city is calculated by equation (2) using the micro data from the *Chinese Population Census*. In plot (b), 1 means 100%. The percentage plot (b) excludes two outlier new cities Shenzhen and Dongguan which do not fit the scale of the plot but still fit the pattern. Both cities are established at the end of 1980s. Because of low initial stocks of Hukou population and high attractions to migrants, both cities have their *Net Stock (%)* measures larger than 1000% in 2010 and the two highest income inequality among all Chinese cities.

Figure 6 shows that income inequality is positively correlated with the net inflow of migrants. When net stock is measured in levels as in sub-figure (a), bigger cities with more migrants are

[12] Alternatively, we investigate this correlation with two other definitions of housing asset income. First, we calculate housing asset income by using the actual square meters a worker owns times the per square meter rent in that city. For instance, a three-person household who owns a $90m^2$ apartment in Beijing, where the average rent is $300/m^2$, yields a household head's estimated housing asset income of $\frac{90}{3} \times 300 = 9000$RMB. Second, we calculate housing asset income for all housing owners rather than just local housing owners. The basic patterns in inequality are similar in these two different definitions. The results are available upon request.

much more unequal. When net stock is measured as a percentage as in sub-figure (b), cities with a larger proportional net gain of migrant workers are much more unequal even though we excluded the two most unequal cities in sub-figure (a). This indicates that cities with more migrants and higher housing costs also exhibit higher income inequality, due to the high housing asset income inequality between local Hukou residents and migrant workers.[13]

Two additional supplementary stylized facts on the decomposition of inequality.

To further document the potential mechanisms associated with the observed income inequality patterns in stylized fact 3, we show two additional supplementary stylized facts on the decomposition of inequality. First, we decompose income into wage income and housing asset income and then calculate the Wage Theil Index. This stylized fact helps illustrate the key source of the observed income inequality pattern. Second, we calculate national income inequality and then decompose it by calculating each city's contribution to national income inequality. This stylized fact shows how the national income inequality is distributed across cities.

Fact 4: ***Wage inequality is only weakly correlated with the net stock of migrants.***

Wage inequality is not the major source of the observed income inequality patterns in Fact 3. Figure 5 displays the correlation between wage inequality and the net stock of migrants in the city. The figure indicates that there is only a weak positive correlation between wage inequality within cities and the net migrant inflows. The slope coefficient of the fitted line is also not significant statistically. Thus, the observed income inequality patterns in Fact 3 are mainly due to housing asset income inequality.

Fact 5: ***National inequality drops but developed city's contribution remains high.***

To document Fact 5, we calculate a national Income Theil Index and then calculate each city's contribution to the national Income Theil Index as follows:

$$T = \sum_{j=1}^{J} s_j (T_j + ln\frac{\bar{i}_j}{\bar{i}}), \quad s_j = \frac{N_j}{N}\frac{\bar{i}_j}{\bar{i}}$$

$$Contri_j = s_j (T_j + ln\frac{\bar{i}_j}{\bar{i}})/T$$

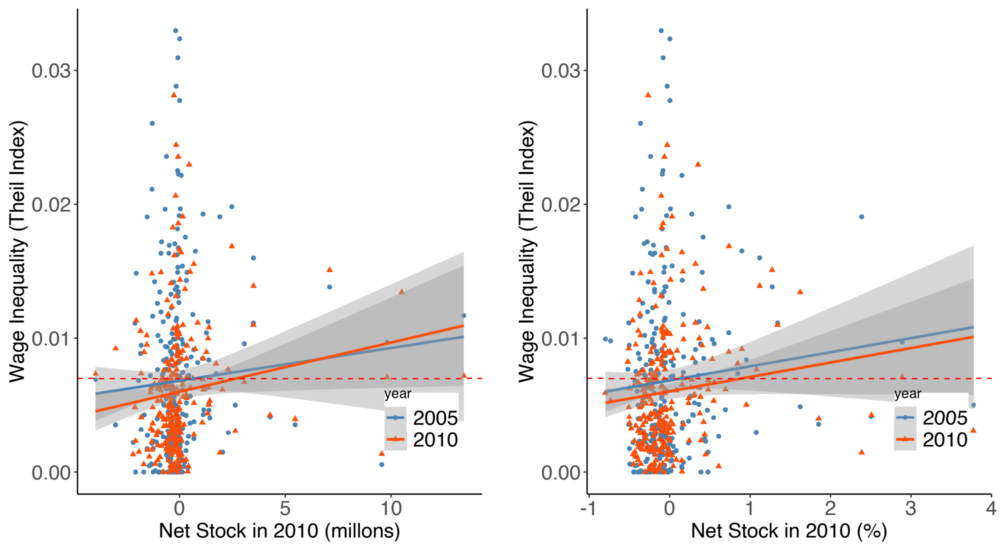
where j indicates city, T is the national Theil Index, T_j is the Theil Index of city j, N_j is the total number of current workers in city j, N is the national total number of workers, $\bar{i}_j$ is the average income in city j, and $\bar{i}$ is the average national income.

The calculated national Income Theil Index dropped by 16% from 0.19 in 2005 to 0.16 in 2010. However the developed city's contribution to national income inequality remains high. Figure 6

[13]Since the financial information of households in the *Census* dataset is limited, we use another dataset called the *Chinese Household Income Project* to investigate other details about inequality between local residents and migrants. Although it is a much smaller dataset covering only a few regions, it still adds value to our investigation. Please refer to Appendix A.3.

Figure 5: Net Stock of Migrants and Wage Inequality

(a) Net Stock (Numbers N) (b) Net Stock (Percentage %)



Notes: The Income Theil Index for each city is calculated by equation (2) using micro data from the *Chinese Population Census*. In plot (b), 1 means 100%. The percentage plot (b) excludes two outlier new cities Shenzhen and Dongguan which do not fit the scale of the plot but still fit the pattern. Both cities were established at the end of the 1980s. Because of low initial stocks of Hukou registrants and high appeal to migrants, both cities have *Net Stock (%)* measures larger than 1000% in 2010. However, them do not have much higher Wage Theil Indexes than other cities.

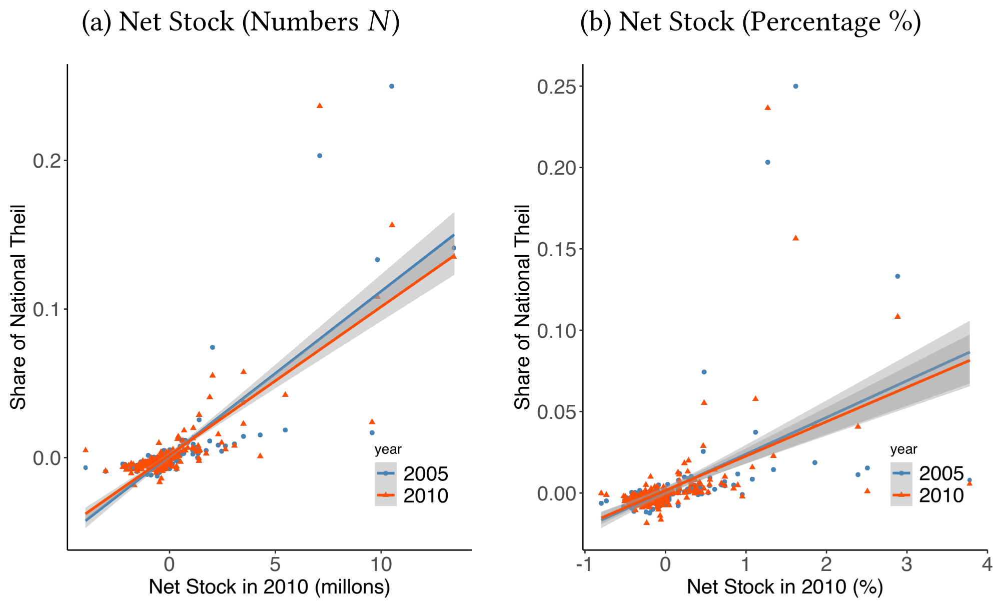
shows the correlation between cities' contributions and their net stock of migrants. The strong positive relationship indicates that larger developed cities with more migrants are contributing much more to national income inequality. This pattern is especially salient for the largest cities. For instance, at the corner of the figure, Shanghai, Shenzhen, and Beijing contribute almost 60%, 40%, and 50%, respectively, to national income inequality in 2010[14]. These numbers were much lower in 2005 (45%, 30%, and 37%, respectively). This indicates that certain Chinese cities with sizeable net stocks of migrant workers contribute much more to national income inequality than other cities.

Remarks of the Stylized Facts

We have shown five important patterns about migration, housing costs, and income inequality in China. They illustrate that as China continues to grow, more and more workers are migrating from under-developed inland areas to developed coastal areas. Because of the restrictive land supply regulations in China, the huge stock of working-age migrants lifts housing demand in big industrialized cities and results in a rapid increase in housing costs. Because most property-owners are local residents, incumbent locals benefit a lot from the rising rents at the expense of

[14]The majority of small cities contribute negatively to national income inequality. That is why the total contribution still sums up to 100% even though the collective contribution of larger cities is larger than 100%.

Figure 6: Net Stock of Migrants and Share of Contribution to National Inequality



Notes: Income Theil Index for each city is calculated by equation (2) using the micro data from the *Chinese Population Census*. In plot (b), 1 means 100%. The percentage plot (b) excludes two outlier new cities Shenzhen and Dongguan which do not fit the scale of the plot but still fit the pattern. Both cities were established at the end of 1980s. Because of low initial stocks of Hukou registrants and high attractions to migrants, both cities have *Net Stock (%)* measures larger than 1000% in 2010 and the two highest share of Income Theil Index among all Chinese cities.

the migrant tenants. This yields the observed positive relation between income inequality and the net stock of migrants, even without any correlation between wage inequality and net stock of migrants. One natural question then arises. How can we alleviate income inequality and motivate more migration flows from less developed/productive areas to more developed/productive areas? This is the main target of this study. To answer this, we construct a spatial equilibrium model with a housing sector and evaluate different policy counterfactuals.

4 The Model

This section describes how we construct the spatial equilibrium model, which will be used in the quantitative analysis and the policy counterfactual analysis.

4.1 Environment

The economy consists of a set of discrete locations, specifically in this paper, **cities**, which are indexed by $j = 1, ..., K$. The economy is populated by an exogenous measure of H workers, who are imperfectly mobile within the economy subject to migration costs. Each worker is either low skill $s = l$ or high skill $s = h$. The total labor in a city is the sum of the two skills, that is, $H_j = H_j^l + H_j^h$. Each location j has an effective supply of floor space S_j, which is produced by a fixed amount of land supply L_j.

Workers decide whether or not to move after observing an idiosyncratic utility shock for each possible destination location. Firms produce a single final good, which is costlessly traded within the city and across the country, and which we take as the numeraire. Locations differ in terms of their final goods productivity A_j^s and the supply of floor space S_j.

4.2 Worker Preferences

The utility of a worker o with skill s, originating from region i and migrating to region j is an aggregation of final good consumption (c_{ijo}), residential space consumption (s_{ijo}), migration costs (τ_{ij}^s), and an idiosyncratic shock (z_{ijo}) in a Cobb-Douglas form:

$$U_{ijo} = \frac{z_{ijo}}{\tau_{ij}^s} \left(\frac{c_{ijo}}{\beta} \right)^{\beta} \left(\frac{s_{ijo}}{1-\beta} \right)^{1-\beta} \tag{3}$$

We model the heterogeneity in the utility that workers derive from working in different parts of the economy following Eaton and Kortum (2002). For each worker o originating from city i and migrating to city j, the idiosyncratic component of utility (z_{ijo}) is drawn from an independent Fréchet distribution:

$$F(z_{ijo}) = e^{-z_{ijo}^{-\epsilon}},\ \epsilon > 1 \tag{4}$$

where the shape parameter $\epsilon > 1$ controls the dispersion of the idiosyncratic shock. We assume that the migration costs can be separated into two parts:

$$\tau_{ij}^s = \bar{\tau}_i^s d_{ij} \tag{5}$$

where d_{ij} captures the physical distance and institutional costs, due to the Hukou system and other potential frictions, in migrating from city i to city j, and $\bar{\tau}_i^s$ captures the difference in the cost across individuals with different skills. It may include differences in high/low skill workers' preferences for amenities such as education for children, entertainment, transportation, and many others.

After observing the realizations for idiosyncratic utility for each employment location, each worker chooses his location of employment to maximize his utility, taking as given residential amenities, goods prices, factor prices, and the location decisions of other workers and firms. Each worker is endowed with one unit of labor that is supplied inelastically with zero disutility. Combining our choice of the final good as numeraire with the first-order conditions for the consumer, we obtain the following demands for the final good and residential land for worker o with skill s

from location i who is migrating to location j:

$$c_{ijo} = \beta v_{ij}^{s} \tag{6}$$

$$s_{ijo} = (1-\beta)\frac{v_{ij}^{s}}{Q_j} \tag{7}$$

where v_{ij}^{s} is the total income, including wage income and return from owning floor space, received by workers in city j. Q_j is the unit rent of residential floor space in city j.

Floor space is not tradable and is commonly owned by all workers whose Hukou is registered in that city.[15] This assumption is broadly consistent with the institutional features of China and is the key component of the observed income inequality. Many migrants only own local properties in their Hukou cities and do not have access to the local housing market in their current city of employment due to financial frictions and policy regulations.[16] Therefore, the income v_{ij}^{s} is a combination of the wage income of skill s workers and the equally-divided rent income among local Hukou residents:

$$v_{ij}^{s} = w_j^{s} + \frac{Q_i S_i}{H_i^R} \tag{8}$$

where H_i^R is the number of Hukou residents registered in their origination city i and S_i is the residential floor space in city i.

Substituting equilibrium consumption of the final good and residential land use into the utility function, we obtain the following expression for the indirect utility function:

$$U_{ijo} = \frac{z_{ijo} v_{ij}^{s} Q_j^{\beta-1}}{\tau_{ij}^{s}} \tag{9}$$

4.3 Distribution of Utility and Migration Flow

Using the monotonic relationship between the utility and the idiosyncratic shock, the distribution of utility for a worker migrating from city i to city j is also Fréchet distributed:

$$G_{ij}^{s}(u) = Pr[U \leq u] = F\Big(\frac{u\tau_{ij}^{s} Q_j^{1-\beta}}{v_{ij}^{s}}\Big) \tag{10}$$

[15] Tombe and Zhu (2019) makes a stronger assumption such that migrant workers have no claim to any fixed factor income from land of either their current working city or their Hukou city. In their model, whenever a worker migrate, she losses all the fixed factor income from her previously owned local property in her Hukou city. We also solve a variation of our model using their assumption. Our mechanism of migration interacting with housing constraints increasing income inequality is further amplified with this assumption.

[16] Migrant workers are usually poorer and not able to pay the down payment in their destination cities. Even though some workers could pay the down payment, they still face a lot of regulations to purchase real estate because they do not own a Hukou registration in the city where they currently reside. As a result, only a very small fraction of migrants are able to participate in the local housing market as locals.

$$G_{ij}^s(u) = e^{-\Phi_{ij}^s u^{-\epsilon}},\ \Phi_{ij}^s = (\tau_{ij}^s Q_j^{1-\beta})^{-\epsilon}(v_{ij}^s)^{\epsilon} \tag{11}$$

Since the maximum of a sequence of Fréchet distributed random variables is itself Fréchet distributed, the distribution of utility across all possible destinations is:

$$1 - G_i^s(u) = 1 - \prod_{k=1}^{K} e^{-\Phi_{ik}^s u^{-\epsilon}} \tag{12}$$

we have

$$G_i^s(u) = e^{-\Phi_i^s u^{-\epsilon}},\ \Phi_i^s = \sum_{k=1}^{K} \Phi_{ik}^s \tag{13}$$

Let π_{ij}^s denote the share of workers with skill s registered in city i who migrate to city j. The proportion of workers who migrate to city j is:

$$\pi_{ij}^s = \frac{(\tau_{ij}^s Q_j^{1-\beta})^{-\epsilon}(v_{ij}^s)^{\epsilon}}{\sum_{k=1}^{K}(\tau_{ik}^s Q_k^{1-\beta})^{-\epsilon}(v_{ik}^s)^{\epsilon}} = \frac{\Phi_{ij}^s}{\Phi_i^s} \tag{14}$$

4.4 Production

We assume there is a single final good that is costlessly traded in the economy. It is produced with perfect competition and constant returns to scale with the following technology:

$$X_j = [(A_j^h H_j^h)^{\frac{\sigma-1}{\sigma}} + (A_j^l H_j^l)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \tag{15}$$

where X_j is a CES combination of high-skill labor H_j^h and low-skill labor H_j^l multiplied by their corresponding city-level efficiency A_j^h and A_j^l respectively.

Firms choose their inputs of workers with different skills to maximize profits, taking as given the final goods productivity ($\{A_j^h, A_j^l\}$), the distribution of idiosyncratic utility, factor prices, and the location decisions of other firms and workers. From the first-order conditions for profit maximization, we obtain:

$$w_j^l = A_j^{l\,\frac{\sigma-1}{\sigma}} X_j^{\frac{1}{\sigma}} H_j^{l\,-\frac{1}{\sigma}} \tag{16}$$

$$w_j^h = A_j^{h\,\frac{\sigma-1}{\sigma}} X_j^{\frac{1}{\sigma}} H_j^{h\,-\frac{1}{\sigma}} \tag{17}$$

This also gives us a measure of the skill premium ω in city j:

$$\omega_j = \frac{w_j^h}{w_j^l} = \left(\frac{A_j^h}{A_j^l}\right)^{\frac{\sigma-1}{\sigma}} \left(\frac{H_j^h}{H_j^l}\right)^{-\frac{1}{\sigma}} \tag{18}$$

The zero profit assumption gives us:

$$X_j = w_j^l H_j^l + w_j^h H_j^h \tag{19}$$

4.5 Floor Space Market Clearing

The standard approach in the urban literature is to assume that floor space S is supplied by a competitive construction sector that uses a Cobb-Douglas technology with geographic land L and construction intensity K as inputs. However, the Chinese land market is highly regulated. The central government restrictively determines both the construction intensity and land supply. Therefore, we assume the following floor space production function with regulated intensity ϕ_j and regulated land supply L_j in each city j:

$$S_j = \phi_j L_j \tag{20}$$

where ϕ_j represents the allowed density of development (the ratio of floor space to land.)

Residential land market clearing implies that the demand for residential floor space equals the supply of residential floor space in each location. Using utility maximization for each worker and taking expectations over the distribution for idiosyncratic utility, this residential land market clearing condition can be expressed as:

$$S_j = E[s_j]H_j = (1-\beta)\frac{E[v_j]H_j}{Q_j} \tag{21}$$

4.6 Definition of Spatial General Equilibrium

We define and characterize the properties of this spatial general equilibrium given the model's fixed parameters $\{\beta, \epsilon, \sigma, \eta\}$.

Definition 4.1. *A **Spatial General Equilibrium** for this economy is defined by a list of exogenous economic conditions $\{\tau_{ij}^s, A_j^s, \phi_j, L_j, H_i^s\}$, a list of endogenous prices $\{Q_j, w_j^s\}$, quantities $\{v_{ij}^s, y_j, H_j^s, S_j\}$, and proportions $\{\pi_{ij}^s\}$ that solve the firm's problem, the worker's problem, the floor space producer's problem, and market clearing such that:*

> *(i).**[Worker Optimization]** Taking the exogenous economic conditions $\{\tau_{ij}^s\}$ and the aggregate prices $\{Q_j, w_j^s\}$ as given, the optimal migration choices of workers pins down the equilibrium labor supply in each city H_j^s and the migration flow between each city pairs π_{ij}^s.*
>
> *(ii).**[Firm Optimization]** Taking the exogenous economic conditions $\{A_j^s\}$ and the aggregate prices $\{w_j^s\}$ as given, firms' optimal production choices pin down the equilibrium labor demand H_j^s.*

*(iv).**[Market Clearing]** For all cities, labor supply equals labor demand and floor space supply equals floor space demand. This pins down the equilibrium aggregate prices $\{Q_j, w_j^s\}$, the equilibrium floor space S_j, and the equilibrium output y_j.*

5 Quantitative Analysis

In this section, we quantify productivities, housing construction intensities, and migration costs for each of the Chinese cities in our sample (which is 233 cities for both years). We first parameterize the model and solve the model with the estimated parameters and the Census data we have in 2005 and 2010. We then show the model results and solve the unobserved variables. Specifically, we show how migration costs, productivity, and housing markets change during these five years.

5.1 Parameterization

Worker Preferences: We first match $(1-\beta)$ to the share of residential floor space cost in consumer expenditure to pin down the share parameters in the worker preferences (β). We use the average accommodation expenditure share of total consumption from UHS to match $(1-\beta)$. The survey is conducted by the National Bureau of Statistics of China who changed their measurement approach in 2012. We think the new approach is more realistic which gives us an average around 23% from 2013 to 2017.[17] Hence, we choose β to be equal to 0.77.

Elasticity of Substitution between Skills: The estimation results of the elasticity for the substitution between high and low-skill labor in China are mixed in previous studies (Dong, Wang, and Gao, 2013; Song, Wang, and Dong, 2010). Therefore, we choose to follow the canonical model of Katz and Murphy (1992) to calibrate the elasticity of substitution between high-skill and low-skill labor (σ) to be equal to 1.4. We also test the model results using alternative calibrations from 1.2 to 3 to ensure the results are robust to our parameter choice.

Migration Elasticity: We estimate the migration elasticity (ϵ) from the gravity equation of migration flow (14). We assume $\tau_{ij}^s = \tau_i^s d_{ij}$, where τ_i^s is the origination-skill fixed component and d_{ij} is the distance index between location i and j. Under these assumptions and given data on migration shares and real incomes, we estimate ϵ using the fixed effect regression:

$$ln(\pi_{ij}^s) = \epsilon ln(v_j^s) + \psi_{ij} + \gamma_{is} + \zeta_j + \phi_{ijs}, \text{ for } i \neq j$$

where $\psi_{ij} = -\epsilon ln(d_{ij})$ is the origination-destination pair fixed effect, $\gamma_{is} = -\epsilon ln(\tau_i^s) - ln(\Phi_i^s)$

[17] According to the old statistical standard, the average housing expenditure share ranges from 11.7% in 2012 to 14.3% in 2002 which is very low because they did not include the converted rent costs of self-owned houses and apartments. From 2013, the converted rent costs of self-owned houses and apartments are added to housing costs, which results in a range of 22.7% in 2017 to 23.3% in 2013. We find that the average expenditure share is very stable across time within both periods.

is the origination-skill fixed effect, $\zeta_j = -\epsilon(1-\beta)ln(Q_j)$ is the destination fixed effect, and ϕ_{ijs} is the measurement error term. We assume that the error term ϕ_{ijs} is not correlated with $ln(v_j^s)$ after controlling for all these fixed effects. Given our estimation, we choose ϵ to be equal to 1.90. The details of the estimation are in the appendix B.1.

Summary of Parameters: Table 2 below summarizes our parameters. The first parameter β is quite standard, as in the literature, such as Ahlfeldt et al. (2015). Chinese citizens have a slightly higher share of final consumption in utility. However, the number is generally similar. As for the migration elasticity, Tombe and Zhu (2019) estimates at the province-sector level and ends with a number of 1.5. Since cities are more substitutable than provinces, the migration elasiticity between cities is slightly larger as expected.

Table 2: **Parameterization**

Parameter	Description	Value
β	share of consumption in utility	0.77
σ	elasticity of substitution between skills	1.40
ϵ	migration elasticity	1.90

5.2 Solving the model

Based on our parameterization and the observed data variables $\{H_i^s, H_j^s, \pi_{ij}^s, w_j^s, Q_j, L_j\}$, we can now calculate all the unobserved variables in each city: productivity $\{A_j^l, A_j^h\}$, migration cost $\{\tau_{ij}^s\}$, floor spaces $\{S_j\}$, and construction density $\{\phi_j\}$ for both 2005 and 2010.

A.Productivity

From profit maximization and the zero profit conditions, we can infer productivity for each city from the data on employment and wages. First, we solve for productivity A_j^h as a function of A_j^l using first order conditions $A_j^h = A_j^l(H_j^h/H_j^l)^{1/(\sigma-1)}(w_j^h/w_j^l)^{\sigma/(\sigma-1)}$. Second, we plug A_j^h into the production function of X_j and apply the zero profit condition to yield:

$$X_j = A_j^l H_j^l \Big[\frac{w_j^h H_j^h + w_j^l H_j^l}{w_j^l H_j^l}\Big]^{\frac{\sigma}{\sigma-1}} = w_j^h H_j^h + w_j^l H_j^l$$

Defining $\Xi_j^l = \frac{w_j^l H_j^l}{w_j^h H_j^h + w_j^l H_j^l}$ as the share of labor income of low-skill workers, we can then calculate the productivities for both skill types as follows:

$$A_j^l = w_j^l(\Xi_j^l)^{\frac{1}{\sigma-1}}$$
$$A_j^h = w_j^h(1-\Xi_j^l)^{\frac{1}{\sigma-1}}$$

B.Construction Intensity

From the workers' first order conditions for floor space and the summation over all workers residing in each city j, we are able to calculate the total amount of floor space S_j:

$$
\begin{aligned}
S_j = E[s_j]H_j &= (1-\beta)\frac{E[v_j]H_j}{Q_j} \\
&= \frac{1-\beta}{Q_j}\left[w_j^l H_j^l + w_j^h H_j^h\right] + (1-\beta)S_j \\
&= \frac{1-\beta}{\beta}\cdot\frac{w_j^l H_j^l + w_j^h H_j^h}{Q_j}
\end{aligned}
$$

and then back out the construction intensity ϕ_j by dividing the land supply data:

$$\phi_j = S_j / L_j$$

C.Migration Costs

To compute migration costs, we need first to compute the city-level rent income which we assume to be equally divided among local residents $\frac{Q_i S_i}{H_i^R}$ from the floor space S_i we calculated above. Then, we can calculate individual worker's income $v_{ij}^s = w_j^s + \frac{Q_i S_i}{H_i^R}$. From the gravity equations, we can then calculate the migration costs between all city pairs. We assume that the iceberg migration cost for staying in the original city is one, that is $\tau_{ii}^s = 1$. With data on rent Q_i, income v_{ij}^s and migration flow π_{ij}^s, and the gravity equation, we have:

$$\Phi_i^s = \sum_{k=1}^{K} (\tau_{ik}^s Q_k^{1-\beta})^{-\epsilon}(v_{ik}^s)^{\epsilon} = \frac{(Q_j^{1-\beta})^{-\epsilon}(v_{ii}^s)^{\epsilon}}{\pi_{ii}^s}$$

Inserting Φ_i^s into the original gravity equation, we have the migration cost as follows:[18]

$$\tau_{ij}^s = \frac{v_{ij}^s}{Q_j^{1-\beta}(\pi_{ij}^s \Phi_i^s)^{1/\epsilon}}\text{ , for } i \neq j$$

5.3 What does the model tell us about the unobservables?

In this subsection, we show the unobserved fundamentals of the model and how they change over time, including migration costs, productivities, and housing construction intensities.

[18] For city pairs with zero migration flow, we assign a migration probability $\pi_{ij}^s \sim 0$, resulting in a huge migration cost approaching infinity, which we will not include while calculating the changes in migration costs.

A.Universal reduction in migration costs

Table 3 reports the share of migrants relative to the total working population, and the mean value of the migration cost τ_{ij}^{s}. On average, in 2005, migrants comprised 11% of total employment in China. As for workers by skill type, the statistics for low-skill workers are very similar to the overall statistics since they are the majority.

Table 3: **Average Migration Costs**

	Share of Emp.		Migration Costs			
	2005	2010	2005	2010	Relative	Changes
Overall	11%	22%	11.0	7.2	65%	-3.8
Low-skill	11%	23%	11.2	7.3	65%	-3.9
High-skill	9%	17%	8.9	7.0	79%	-1.9

Notes: This table displays migration-weighted harmonic means of migration costs in 2005 and 2010. Share of Employment among high-skill is high-skill migrants over high-skill population. Because τ_{ij}^{s} is proportional in the model, we show % changes.

In 2010, overall migration costs dropped dramatically by 35% relative to 2005. For low-skill workers, the changes were similar to the national average, while for high-skill workers, the drop on average was smaller (21%). With these huge drops in migration costs, we observe the share of migrants relative to the total working population doubling to 22%. More importantly, high-skill workers started to move more. These results indicate that the decreasing migration costs contribute a lot to the increasing migration flows.

As documented in Bryan and Morten (2019), the dramatic drop in migration costs is essential for the observed massive flow of migrant workers in developing countries. Tombe and Zhu (2019) also shows that province-sector level migration costs dropped a lot between 2000 and 2005. Our results indicate that the same pattern holds at the city-skill level as well. Though these changes are not the key we want to address in this paper, it is still important to capture them in the model so that the model will not overestimate the contribution of other elements.

B.Uneven productivities and uneven growth in productivities

Table 4 presents the average productivities A_j^s for both high-skill and low-skill workers, for all cities j grouped by net stock of migrant workers. On average, the overall productivity for all the cities grows by 87% for high-skill and by 94% for low-skill workers. To show the results in a more compact way, we group cities by their net stock of migrant workers. (6,13) refers to cities having a net stock of migrant workers between 6 million and 13 million. Similarly, (-4,-1) refers to cities having a net stock of migrant workers between -4 million and -1 million. We will use the these groupings through out the paper.

We find that, first, cities with a larger net stock of migrant workers have much higher productivities than cities with smaller or negative net stock for both high-skill and low-skill workers.

Table 4: **Average Productivity Growth**

Net Migrant Range(2010)	No. of Cities	High-skill 2005	2010	Relative	Changes	Low-skill 2005	2010	Relative	Changes
Average	233	6.4	14.0	219%	+7.6	9.4	17.1	182%	+7.7
(6,13)	5	19.2	45.7	240%	+26.5	12.6	21.2	168%	+8.6
(1,6)	19	3.9	12.0	308%	+8.1	12.2	19.5	160%	+7.3
(0, 1)	45	3.7	10.5	184%	+6.8	10.2	16.3	160%	+6.1
(-1,0)	134	0.9	2.3	256%	+1.4	8.2	16.3	199%	+8.1
(-4,-1)	30	0.4	1.6	400%	+1.2	7.8	15.2	195%	+7.4

Notes: This table displays population-weighted means in both 2005 and 2010 and their changes. The level of high-skill and low-skill productivity are not directly comparable. For readability, we normalize both numbers. The unit of high-skill productivity is $1e2$ and the unit of low-skill productivity is $1e3$. The net stock of migrant worker range groups are classified by net stock of migrant workers in 2010 (unit: millions). Each Net Migrant Range Group consists of the same cities in 2005 and 2010. There are 233 cities in the model.

For instance, Tier 1 cities, including Beijing, Shanghai, Shenzhen, Guangzhou, and Dongguan, had more than thirty million net stock of migrant workers in 2010. These cities had much higher productivity for both high-skill and low-skill workers in both 2005 and 2010. In 2005, their average high-skill productivity was 19.2, which was 200% higher than the national average, 290% higher than Tier 2 cities. However for low-skill, the differences between city groups are smaller. Tier 1 cities' average low-skill productivity is 12.6, which is 34% higher than the national average and only 3% higher than Tier 2 cities.

Second, productivities improved massively from 2005 to 2010 and especially for the high-skill productivities in developed cities with more migrants. The national average productivity improved by 119% for high-skill and 82% for low-skill. While smaller cities' productivity improved more in percentage terms because they had a smaller base in 2005, if we focus on the changes in absolute value, it is easy to spot that the improvement of high-skill productivity is much larger in cities with more migrant workers. The high-skill productivity increased by 26.5 in Tier 1 cities but only increased by 1.2 in Tier 5 cities.

All these results indicate that the reallocation of workers, especially high-skill workers, from these less productive cities to more productive cities, will significantly improve national productivity and therefore improve the national level of welfare.

C.Tightening housing constraints in developed cities

The land supply for each city in China is determined administratively by the central government. Table 5 shows the supply of construction land and floor space and how they change from 2005 to 2010. The national total land supply increased by 31%. However, the total land supply in Tier 1 cities only increased by 10% despite the massive migration. Tier 2 cities increased their total land supply the most (55%). Meanwhile, Tier 4 and 5 cities which are losing more than massive amount of workers gain 30% and 38% construction land, respectively. Given that most of the Tier

1 and 2 cities are located on plains, their construction land supply is essentially less than 10% of their administrative districts (except Shenzhen). This leaves substantial room for increasing or spatially reallocating the total land supply to larger cities to loosen housing constraints.

Table 5: **Construction Land Supply and Floor Space**

Net Migrant Range(2010)	No. of City	Total Land Supply Floor Space				Total Floor Space			
		2005	2010	Relative	Changes	2005	2010	Relative	Changes
Overall	233	24,277	31,705	131%	+7,428	2.19	3.30	150%	+1.11
(6,13)	5	5,135	5,648	110%	+513	5.92	7.84	132%	+1.92
(1,6)	19	3,801	5,912	155%	+2,111	1.79	4.10	229%	+2.31
(0, 1)	45	5,555	7,250	131%	+1,695	1.53	2.48	162%	+0.95
(-1,0)	134	7,950	10,363	130%	+2,413	1.48	2.17	147%	+0.69
(-4,-1)	30	1,836	2,532	138%	+696	2.55	3.12	122%	+0.57

Notes: This table displays total land supply within groups (unit: km^2) and total floor space (unit: $1e8$ m^2). The Net Migrant Range is classified by the net stock of migrant workers in 2010 (unit: millions). Each Net Migrant Range Group consists of the same cities in 2005 and 2010. There are 233 cities in the model.

5.4 Wage Inequality and Income Inequality

In this subsection, we show wage inequality and income inequality measured by the Theil Index in our model in both 2005 and 2010.

Table 6: **Within-city Theil Index**

Net Migrant Range(2010)	No. of City	Wage Theil Index			Income Theil Index		
		2005	2010	Relative	2005	2010	Relative
Average	233	0.0072	0.0070	97%	0.0100	0.0184	184%
(6,13)	5	0.0087	0.0097	111%	0.0442	0.0908	205%
(1,6)	19	0.0065	0.0079	122%	0.0092	0.0223	242%
(0, 1)	45	0.0075	0.0083	111%	0.0060	0.0092	153%
(-1,0)	134	0.0071	0.0058	82%	0.0049	0.0052	106%
(-4,-1)	30	0.0072	0.0058	80%	0.0054	0.0062	115%

Notes: This table displays population-weighted means in 2005 and 2010.

Table 6 shows the within-city Theil Index for both wages and income. The average Wage Theil Index is 0.0072 in 2005 and declined slightly to 0.0070 in 2010. Larger cities with more migrant workers have slightly higher wage inequality, and their wage inequality increased slightly from 2005 to 2010. On the other hand, during the same period, wage inequality decreased in smaller cities with negative net migrant workers. However, the differences and the changes in wage inequality across cities and across time are not comparable to these patterns of income inequality. The average within-city Income Theil Index was much higher than the average within-city Wage

Theil Index, and it doubled from 2005 to 2010. If we break down the statistics by city groups, we easily observe that this huge jump is attributable to cities with positive net migrants, especially Tier 1 and Tier 2 cities, with more than 100% increases.

Table 7 shows contribution shares to national Theil Indexes. The first row shows the national Wage Theil Index and Income Theil Index for both 2005 and 2010. At the national level, income inequality is still higher than wage inequality. Both measures dropped as more workers migrated from lower productivity areas to higher productivity areas.[19] Moreover, if we examine by city groups, we observe that larger cities with positive net migration contribute massively to both national Theil Index measures. For instance, for the Wage Theil of Tier 1 cities in 2005, +1.49 means that if we do not account for all workers in Tier 1 cities, the national Wage Theil would decrease by 149%. This pattern holds for both inequality measures and does not change much from 2005 to 2010.

Table 7: **Share of Contribution to National Theil Index**

Net Migrant	No. of	Share of Wage Theil			Share of Income Theil		
Range(2010)	City	2005	2010	Relative	2005	2010	Relative
National Theil	233	0.0985	0.0622	64%	0.1156	0.0921	80%
(6,13)	5	+1.49	+1.41	97%	+1.43	+1.27	89%
(1,6)	19	+0.58	+0.83	143%	+0.53	+0.70	132%
(0, 1)	45	+0.22	+0.26	118%	+0.19	+0.20	105%
(-1,0)	134	-0.92	-1.00	108%	-0.81	-0.78	96%
(-4,-1)	30	-0.37	-0.49	132%	-0.35	-0.39	111%

Notes: This table displays city groups' contribution to the national Theil Index in 2005 and 2010. The first row displays the national wage/income Theil Index in 2005 and 2010.

To further indicate how housing constraints play an essential role, we show the skill premium and the housing premium (measured as average annual housing return over the average annual wage) and their changes in Table 8. The national average skill premium and the city groups' skill premiums are very similar and do not change much over time. However, the average housing premium increased from 0.36 in 2005 to 0.49 in 2010, resulting in a 36% jump. For an "average" worker, housing asset income is almost 50% of his wage income. Furthermore, if we break down by city groups, we observe that in Tier 1 cities, the housing premium increased from 0.93 to 1.89, which is substantially above the average rate of growth. Given that houses in these large cities are almost all owned by locals and many more migrants are moving into these cities, it is not hard to understand the astonishing income inequality in Table 6.

[19] The trend is similar to the Gini Index published by the National Bureau of Statistics. The Gini Index in 2010 is 0.481 and the Gini Index in 2005 is 0.485.

Table 8: **Skill Premium and Housing Premium**

Net Migrant Range(2010)	No. of Cities	Skill Premium 2005	Skill Premium 2010	Skill Premium Relative	Housing Premium 2005	Housing Premium 2010	Housing Premium Relative
Average	233	1.47	1.40	95%	0.36	0.49	136%
(6,13)	5	1.35	1.39	103%	0.93	1.89	203%
(1,6)	19	1.40	1.40	100%	0.39	0.56	144%
(0, 1)	45	1.42	1.39	97%	0.31	0.35	113%
(-1,0)	134	1.50	1.40	93%	0.27	0.25	93%
(-4,-1)	30	1.58	1.45	92%	0.24	0.31	129%

Notes: This table displays population-weighted means in 2005 and 2010. Skill Premium is measured as annual high-skill wage over annual low-skill wage for each city, and Housing Premium is measured as average annual housing return over the average annual wage for each city.

5.5 Remarks on the Quantitative Analysis

In this section, we show that the universal reduction in migration costs, the uneven productivities, and the uneven growth in productivities are the major drivers of the massive migration flows in China. Furthermore, the restrictive housing constraints in cities with positive net stock of migrants are much tighter. These housing constraints increase income inequality in these larger cities and dissuade more migrants from entering these cities with higher productivities.

6 Counterfactual Analysis

In this section, we simulate some policies recommended in previous literature using our model. We try to recover how the policies could change the spatial distribution of workers with different skills in China. Most importantly, we investigate the effect of these policies on the housing market in different regions, and on national and within-city inequality. We employ an iteration algorithm to compute the counterfactuals. The details of the algorithm are in appendix B.2.

6.1 A Migration-based Land Supply Policy Reform

The most important reason that housing constraints are very tight in larger cities is because China has had a very restrictive construction land supply policy since the 1950s. The central government decides the total amount and the distribution of the total land supply for all Chinese cities year by year. The local governments follow these instructions to change their city-level land supply to match their city quotas. These quotas cannot be traded between cities. Therefore, land deficient cities and land abundant cities co-exist at the same time. In this section, we propose a policy to allocate more land to large cities with more migrants.

A.Current Land Supply Policy in China

Since 2003, the central government changed the principles of its land supply policy. The purpose is to balance regional development using land quotas as a regional income redistribution device. This is documented by a large urban literature (Han and Lu, 2017, 2018; Liang, Lu, and Zhang, 2016). There are two general guidelines. First, redistributing extra land supply away from the coastal areas (more developed) to favoring the inland areas (less developed). The inland share of the national land supply quota rose from 30% in 2003 to 60% in 2014. Second, redistributing land supply from favoring large cities (more developed) to favoring smaller cities (less developed). The small cities' share of the national land supply increased from 49% in 2003 to 64% in 2014. This trend has persisted since the beginning of 2003 until today.

However, from the stylized facts of migration flows and housing costs, we think the current land supply policy is inefficient. It is increasing land supply in cities which are less productive and losing workers while restricting land supply in cities which are much more productive and gaining workers. Even though workers in less developed cities do receive additional land income just due to having more land, this policy is economically poor in terms of both productivity and equality. Therefore, we propose an alternative land supply policy that favors high productivity cities with an endogenous cross-city transfer based on migration flows.

B.Migration-Based Land Supply Policy Reform

We propose a counterfactual policy of redistributing the total land supply increment from 2005 to 2010 according to the changes in the net stock of migrant workers. More specifically, the rule for land supply redistribution is as follows. Call the total land supply increment from 2005 to 2010 as ΔL, and the increase in the net stock of migrant workers in each city as $\Delta^+ H_j$ which sums up to total worker population growth $\Delta^+ H$. Then city $j's$ counterfactual land supply increment is: $\Delta^+ L_j = \Delta L \times (\Delta^+ H_j / \Delta^+ H)$. Since it is very costly to revoke current land supply, for cities with negative migrant changes, we assign $\Delta^0 L_j = 0$. The counterfactual land policy changes are summarized in Table 9.

Table 9: **Counterfactual Construction Land Supply**

Net Migrant Range(2010)	No. of Cities	Land Supply (Data) 2005	2010	Relative	Changes	Counterfactual $\widehat{2010}$	$\widehat{Relative}$	$\widehat{Changes}$
National	233	24,277	31,705	131%	+7,428	31,705	131%	+7,428
(6,13)	5	5,135	5,648	110%	+513	7,762	151%	+2,627
(1,6)	19	3,801	5,912	155%	+2,111	7,131	188%	+3,330
(0, 1)	45	5,555	7,250	131%	+1,695	6,829	123%	+1,274
(-1,0)	134	7,950	10,363	130%	+2,413	7,988	100.5%	+38
(-4,-1)	30	1,836	2,532	138%	+696	1,836	100%	+0

Notes: This table displays total land supply data by group in 2005 and 2010, as well as the counterfactual migration-based land supply in 2010 (unit: km^2). The Range is classified by net stock of migrant workers in 2010 as in the data (unit: millions). Each Net Migrant Range Group consists of the same cities in 2005 and 2010.

This counterfactual is feasible to implement and still fulfills the central government's goal of balancing regional development. We subtract land income from the additional land allocated to land-gaining cities and compensate land-losing cities for their losses to achieve the redistribution motive. This mechanism mimics a policy called the "land quota market", which has been recommended by previous literature (Lu, 2016). The basic idea is that central government can balance the development of different regions by transferring revenues from developed cities to under-developed cities, rather than allocating the land supply directly. Since the land income in land-gaining cities is higher than the land income in land-losing cities and the total amount of land supply is unchanged, this redistribution is feasible and central government can even generate profit from it.

C.Land Supply Policy Reform Results

The results of the land supply policy reform are summarized in Table 10 to Table 13. We list the original equilibrium and the counterfactual (with a hat) side-by-side for ease of comparison. Table 10 shows how this counterfactual policy changes net migration and housing costs. First, the policy motivates 17% more workers to move from low productivity cities to high productivity cities, and the increases are the highest in the most productive cities (Tier 1: 22% > Tier 2: 16% > Tier 3: 0%). Meanwhile, because of the land supply redistribution, more land is distributed to cities with more incoming migrants, and housing costs in these cities drop a lot. For Tier 1 and Tier 2 cities, the costs drop to only 70% and 75% of the original equilibrium.

Table 10: **Migration Flow and Housing Cost: Land Supply Reform**

Net Migrant	No. of	Net Migrant			Housing Cost		
Range(2010)	Cities	2010	$\widehat{2010}$	Relative	2010	$\widehat{2010}$	Relative
Overall	233	96m	112m	117%	114	119	104%
(6,13)	5	+45m	+55m	122%	226	158	70%
(1,6)	19	+38m	+44m	116%	136	102	75%
(0, 1)	45	+13m	+13m	100%	118	132	112%
(-1,0)	134	-48m	-48m	100%	87	115	132%
(-4,-1)	30	-48m	-65m	135%	80	105	131%

Notes: This table displays total net stock of migrant workers and population weighted average housing costs for each city group. In the first raw (Overall), we show the number of workers who have migrated and the national population weighted average housing cost. The unit of the net migrant is millions, and the unit of housing costs is Chinese Yuan (RMB) per square meters per year.

We then show how within-city inequality changes in Table 11. The first thing to notice is that the Wage Theil Index effectively does not change. The only noticeable change is that the Theil Index in Tier 2 cities increases by 13%. This is mainly because more high-skill workers move to Tier 1 and Tier 2 cities due to the dramatic drop in housing costs. Nevertheless, for any other city group, the Wage Theil Index is almost identical. However, the population-weighted mean Income Theil Index drops significantly from 0.0184 to 0.0121 (34% drop). Moreover, if we divide by city groups, the drops are much larger for Tier 1 and Tier 2 cities. Since almost 30% of all workers live in these cities, it significantly lowers the average within-city Income Theil Index even though the Income Theil Index rises in cities losing workers. Therefore, the land supply reform helps to reduce within-city income inequality.

Table 11: **Within-city Theil Index: Land Supply Reform**

Net Migrant Range(2010)	No. of Cities	Wage Theil Index 2010	Wage Theil Index $\widehat{2010}$	Wage Theil Index Relative	Income Theil Index 2010	Income Theil Index $\widehat{2010}$	Income Theil Index Relative
Average	233	0.0070	0.0072	103%	0.0184	0.0121	66%
(6,13)	5	0.0097	0.0093	97%	0.0908	0.0428	47%
(1,6)	19	0.0079	0.0089	113%	0.0223	0.0139	62%
(0, 1)	45	0.0083	0.0082	99%	0.0092	0.0098	106%
(-1,0)	134	0.0058	0.0059	101%	0.0052	0.0045	86%
(-4,-1)	30	0.0058	0.0056	97%	0.0062	0.0051	82%

Notes: This table displays population-weighted means of both inequality measures. The original equilibrium is 2010 and the counterfactual equilibrium is $\widehat{2010}$. Relative is calculated via dividing $\widehat{2010}$ by 2010.

We also want to show how the policy changes national inequality and each city's contribution to national inequality in Table 12. Similar to the pattern of within-city inequality, the counterfactual policy does not have much effect on national wage inequality or cities' contributions to national wage inequality. The national Wage Theil Index is unchanged. However, the counterfactual policy significantly lowers national income inequality by 20% measured by the Income Theil Index. By city groups, the positive contributions of Tier 1, 2 and 3 cities and the negative contributions of Tier 4 and Tier 5 cities increases. All these results indicate that the land supply reform lowers national income inequality but not cross-city income inequality since we motivated more high-skill migrants into more productive cities.

Table 12: **Share of National Theil Index: Land Supply Reform**

Net Migrant Range(2010)	No. of Cities	Share of Wage Theil 2010	Share of Wage Theil $\widehat{2010}$	Share of Wage Theil Relative	Share of Income Theil 2010	Share of Income Theil $\widehat{2010}$	Share of Income Theil Relative
National Theil	233	0.062	0.062	100%	0.092	0.074	80%
(6,13)	5	+1.41	+1.46	104%	+1.27	+1.28	101%
(1,6)	19	+0.83	+0.84	101%	+0.70	+0.66	94%
(0, 1)	45	+0.26	+0.23	88%	+0.20	+0.30	150%
(-1,0)	134	-1.00	-0.95	95%	-0.78	-0.73	94%
(-4,-1)	30	-0.49	-0.58	118%	-0.39	-0.50	128%

Notes: This table displays city groups' contribution to the national Theil Index in 2005 and 2010. The first row displays the national wage/income Theil Index in 2005 and 2010.

Finally, we show the skill premium and the housing premium in Table 13. The skill premium is the high-skill wage over the low-skill wage, and the housing premium is the average housing return over the average wage return. The underlying reason why any measures of wage inequality do not change much is that the skill premium does not move at all. The only changes come from the location choices of high-skill workers relative to low-skill, which changes the composition of workers in each city. However, for the housing premium, it is another story. Since the government increases land supply in cities with insufficient land quotas, housing costs drop massively, which dilutes the asset return from property ownership. As a result, housing premia fall by 41% and 27% in Tier 1 and Tier 2 cities. These results help us to better understand the changes in the Theil Indexes.

Table 13: **Skill Premium and Housing Premium**: **Land Supply Reform**

Net Migrant	No. of	Skill Premium			Housing Premium		
Range(2010)	Cities	2010	$\widehat{2010}$	Relative	2010	$\widehat{2010}$	Relative
Average	233	1.40	1.40	100%	0.49	0.45	92%
(6,13)	5	1.39	1.39	100%	1.89	1.12	59%
(1,6)	19	1.40	1.43	102%	0.56	0.41	73%
(0, 1)	45	1.39	1.38	99%	0.35	0.40	114%
(-1,0)	134	1.40	1.39	99%	0.25	0.33	132%
(-4,-1)	30	1.45	1.43	98%	0.31	0.26	84%

Notes: This table displays population-weighted means in 2005 and 2010. Skill Premium is measured as annual high-skill wage over annual low-skill wage for each city, and Housing Premium is measured as average annual housing return over the average annual wage for each city.

6.2 Property Tax and Redistribution

China has no property tax on housing ownership so far. There is a heated debate on whether China should adopt a property tax as redistribution policy. It is widely documented that more than 75% of Chinese household wealth is in housing. Given the approximate ratio of a property tax to rent revenue is roughly 20% in the U.S., this counterfactual taxes property owners' housing income by 20% and redistributes the proceeds to all residences in the same city (think about using the tax revenue to build infrastructures which benefits all residents equally.). For brevity, we only discuss the key results on migrations, housing costs, and inequality in Table 14, Table 15, and Table 16. Other results are presented in appendix C.

Could a reasonable property tax and redistribution give us desirable reductions in income inequality? The answer is yes. This policy can effectively lower income inequality because migrant workers pay property tax for their house in their Hukou city but gain property tax redistribution income from their current working city. The former is usually much lower than the latter for migrants moving from under-developed cities to developed cities. Therefore, property tax allows migrants to share the benefits of the floor space market returns even though they do not own any property in their current working cities.

Table 14: **Migration Flow and Housing Costs: Property Tax**

Net Migrant Range(2010)	No. of Cities	Net Migrant 2010	Net Migrant $\widehat{2010}$	Net Migrant Relative	Housing Cost 2010	Housing Cost $\widehat{2010}$	Housing Cost Relative
Overall	233	96m	97m	101%	114	115	101%
(6,13)	5	+45m	+46m	102%	226	230	102%
(1,6)	19	+38m	+39m	102%	136	137	101%
(0, 1)	45	+13m	+13m	100%	118	118	100%
(-1,0)	134	-48m	-47m	102%	87	87	100%
(-4,-1)	30	-48m	-50m	104%	80	80	100%

Notes: This table displays total net stock of migrant workers and population weighted average housing costs for each city group. In the first row (Overall), we show the number of workers who have migrated and the national population weighted average housing cost. The unit of the net migrant is millions, and the unit of housing costs is Chinese Yuan (RMB) per square meters per year.

Table 15: **Within-city Theil Index: Property Tax**

Net Migrant Range(2010)	No. of Cities	Wage Theil Index 2010	Wage Theil Index $\widehat{2010}$	Wage Theil Index Relative	Income Theil Index 2010	Income Theil Index $\widehat{2010}$	Income Theil Index Relative
Average	233	0.0070	0.0071	101%	0.0184	0.0145	79%
(6,13)	5	0.0097	0.0100	103%	0.0908	0.0670	74%
(1,6)	19	0.0079	0.0080	101%	0.0223	0.0171	77%
(0, 1)	45	0.0083	0.0084	101%	0.0092	0.0081	88%
(-1,0)	134	0.0058	0.0058	100%	0.0052	0.0047	90%
(-4,-1)	30	0.0058	0.0058	100%	0.0062	0.0053	85%

Notes: This table displays population-weighted means of both inequality measures. The original equilibrium is 2010 and the counterfactual equilibrium is $\widehat{2010}$. Relative is calculated via dividing $\widehat{2010}$ by 2010.

Table 16: **Share of National Theil Index: Property Tax**

Net Migrant Range(2010)	No. of Cities	Share of Wage Theil 2010	Share of Wage Theil $\widehat{2010}$	Share of Wage Theil Relative	Share of Income Theil 2010	Share of Income Theil $\widehat{2010}$	Share of Income Theil Relative
National Theil	233	0.062	0.062	100%	0.092	0.074	80%
(6,13)	5	+1.41	+1.42	104%	+1.27	+1.31	103%
(1,6)	19	+0.83	+0.83	101%	+0.70	+0.73	104%
(0, 1)	45	+0.26	+0.26	88%	+0.20	+0.21	105%
(-1,0)	134	-1.00	-0.98	95%	-0.78	-0.82	111%
(-4,-1)	30	-0.49	-0.52	118%	-0.39	-0.44	116%

Notes: This table displays city groups' contribution to the national Theil Index in 2005 and 2010. The first row displays the national wage/income Theil Index in 2005 and 2010.

From Table 14, 15 and 16, we can tell that even though property tax cannot motivate much more migrantions, and barely changes housing costs, it still lowers income inequality (20% drops in the national Theil Index). It works almost exclusively as a redistribution device between local property owners and migrant workers. Therefore, even though it lowers income inequality a lot, it will face a lot of opposition from local property owners in big cities.

6.3 Additional Counterfactual Analysis

In this section, we discuss an additional counterfactual policy reforms: Directly increasing land supply in developed cities. The policy reform is through directly increasing the land supply based on migration inflow (ignoring the additional cost of offering new construction land). Instead of promoting the trade of land quotas across cities, we directly double the land supply increment from 2005 to 2010 and redistribute the additional land supply to cities with positive net migrants. Could we lower income inequality from directly increasing land supply everywhere? The answer is also no. Because the revenue from additional land supply is only redistributed among the local Hukou holders, this policy will only worsen income inequality even though housing costs are dramatically reduced. Detailed results are presented in appendix C.

7 Conclusion

Migration and housing constraints shape income inequality within and across Chinese cities. Along with the nationwide reduction of migration costs and the rapid growth of productivity in more developed cities, we observe a massive reallocation of workers towards these more developed cities, a rapid growth of housing costs in these more developed cities, and a stark increase in income inequality. In a spatial equilibrium model, we explain the mechanism behind these observations and quantify the impacts of the interactions of the massive spatial reallocation of workers with the rapid growth of housing costs on income inequality. The rapid migration to more developed cities and the highly regulated land supply system contribute to housing demand and lift housing costs (rent), which benefits local real estate owners. Housing owners gain more from the rents, and tenants spend more by paying rents. Thus, housing ownership inequality increases inequality in disposable income within the developed cities and across the whole country.

With this understanding of the mechanism, we conduct several feasible counterfactual experiments. Among all counterfactuals, we show that a migration-based land supply reform that allows regions to ”trade” construction land usage quotas could lower within-city income inequality by 34% and national income inequality by 20%. This also encourages more migration to higher productivity cities and improves nationwide productivity.

References

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm, and Nikolaus Wolf. 2015. "The economics of density: Evidence from the Berlin Wall." *Econometrica* 83 (6):2127–2189.

Berman, Eli, John Bound, and Stephen Machin. 1998. "Implications of Skill-biased Technological Change: International Evidence." *The Quarterly Journal of Economics* 113 (4):1245–1279.

Bryan, Gharad and Melanie Morten. 2019. "The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia." *Journal of Political Economy* 127 (5):2229–2268.

Card, David and John E DiNardo. 2002. "Skill-biased Technological Change and Rising Wage Inequality: Some Problems and Puzzles." *Journal of Labor Economics* 20 (4):733–783.

Chao, Chi-Chur, Bharat R Hazari, and Eden SH Yu. 2006. "Rising Wage Inequality in Developing Economies: Privatization and Competition." *Journal of International Trade & Economic Development* 15 (3):375–385.

Chen, Binkai, Dan Liu, and Ming Lu. 2018. "City Size, Migration and Urban Inequality in China." *China Economic Review* .

Chen, Kaiji and Yi Wen. 2017. "The great housing boom of China." *American Economic Journal: Macroeconomics* 9 (2):73–114.

Cuadrado-Ballesteros, Beatriz and Noemí Peña-Miguel. 2018. "The Socioeconomic Consequences of Privatization: An Empirical Analysis for Europe." *Social Indicators Research* 139 (1):163–183.

Dong, Zhiqing, Fangling Wang, and Qingkun Gao. 2013. "Does Skill Premium Derive from Biased Technical Change?" *Statistical Research* 6.

Eaton, Jonathan and Samuel Kortum. 2002. "Technology, Geography, and Trade." *Econometrica* 70 (5):1741–1779.

Fan, Jingting. 2019. "Internal Geography, Labor Mobility, and the Distributional Impacts of Trade." *American Economic Journal: Macroeconomics* 11 (3):252–88.

Fang, Hanming, Quanlin Gu, Wei Xiong, and Li-An Zhou. 2016. "Demystifying the Chinese Housing Boom." *NBER Macroeconomics Annual* 30 (1):105–166.

Garriga, Carlos, Aaron Hedlund, Yang Tang, and Ping Wang. 2017. "Rural-urban Migration, Structural Transformation, and Housing Markets in China." Tech. rep., National Bureau of Economic Research.

———. 2020. "Rural-urban Migration and House Prices in China." *Regional Science and Urban Economics* :103613.

Glaeser, Edward, Wei Huang, Yueran Ma, and Andrei Shleifer. 2017. "A Real Estate Boom with Chinese Characteristics." *Journal of Economic Perspectives* 31 (1):93–116.

Goldberg, Pinelopi K and Nina Pavcnik. 2004. "Trade, Inequality, and Poverty: What Do We Know? Evidence from Recent Trade Liberalization Episodes in Developing Countries." Tech. rep., National Bureau of Economic Research.

Gregorio, Jose De and Jong-Wha Lee. 2002. "Education and Income Inequality: New Evidence from Cross-country Data." *Review of Income and Wealth* 48 (3):395–416.

Han, Jun, Runjuan Liu, and Junsen Zhang. 2012. "Globalization and Wage Inequality: Evidence from Urban China." *Journal of International Economics* 87 (2):288–297.

Han, Libin and Ming Lu. 2017. "Housing Prices and Investment: An Assessment of China's Inland-favoring Land Supply Policies." *Journal of the Asia Pacific Economy* 22 (1):106–121.

———. 2018. "Supply-Demand Mismatch: Solving the Puzzle of China's Housing Price Divergence." *The Journal of World Economy* (10):7.

Hsieh, Chang-Tai and Enrico Moretti. 2019. "Housing Constraints and Spatial Misallocation." *American Economic Journal: Macroeconomics* 11 (2):1–39.

Katz, Lawrence F and Kevin M Murphy. 1992. "Changes in Relative Wages, 1963–1987: Supply and Demand Factors." *The Quarterly Journal of Economics* 107 (1):35–78.

Li, Shi, Hiroshi Sato, and Terry Sicular. 2013. *Rising Inequality in China: Challenges to a Harmonious Society*. Cambridge University Press.

Liang, Wenquan, Ming Lu, and Hang Zhang. 2016. "Housing Prices Raise Wages: Estimating the Unexpected Effects of Land Supply Regulation in China." *Journal of Housing Economics* 33:70–81.

Lu, Ming. 2016. *Great Nation Needs Bigger City*. Shanghai People's Press.

Moore, Mark P and Priya Ranjan. 2005. "Globalisation vs Skill-biased Technological Change: Implications for Unemployment and Wage Inequality." *The Economic Journal* 115 (503):391–422.

Morten, Melanie and Jaqueline Oliveira. 2014. "Migration, Roads and Labor Market Integration: Evidence from a Planned Capital City." *Unpublished manuscript* .

Novotnỳ, Josef. 2007. "On the Measurement of Regional Inequality: Does Spatial Dimension of Income Inequality Matter?" *The Annals of Regional Science* 41 (3):563–580.

Piketty, Thomas, Li Yang, and Gabriel Zucman. 2019. "Capital Accumulation, Private Property, and Rising Inequality in China, 1978–2015." *American Economic Review* 109 (7):2469–96.

Song, Donglin, Linhui Wang, and Zhiqing Dong. 2010. "Is There Skill Biased Technological Change? Evidence from China." *Economic Research Journal* 5:68–81.

Song, Yang. 2014. "What Should Economists Know About the Current Chinese Hukou System?" *China Economic Review* 29:200–212.

Sylwester, Kevin. 2002. "Can Education Expenditures Reduce Income Inequality?" *Economics of Education Review* 21 (1):43–52.

Tombe, Trevor and Xiaodong Zhu. 2019. "Trade, Migration, and Productivity: A Quantitative Analysis of China." *American Economic Review* 109 (5):1843–72.

Verhoogen, Eric A. 2008. "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector." *The Quarterly Journal of Economics* 123 (2):489–530.

Wu, Jing, Joseph Gyourko, and Yongheng Deng. 2016. "Evaluating the Risk of Chinese Housing Markets: What We Know and What We Need to Know." *China Economic Review* 39:91–114.

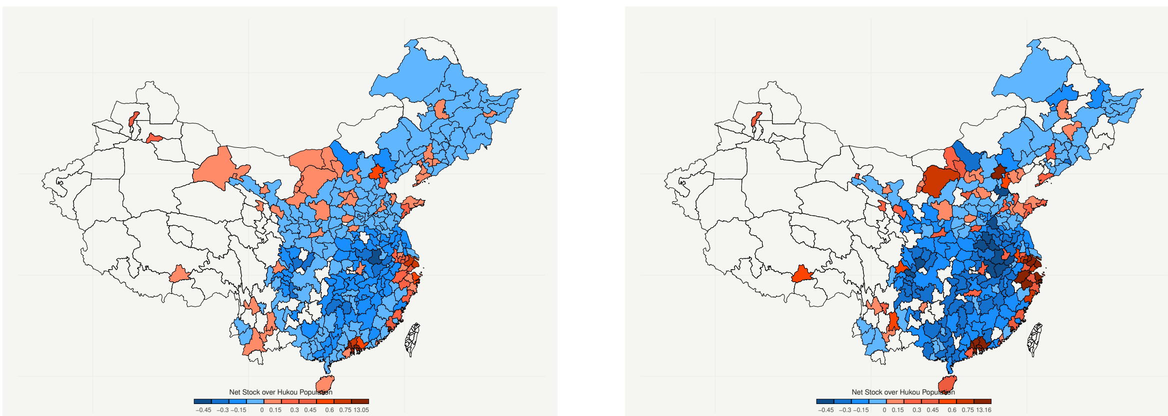
Appendix

A Empirical Appendix

A.1 Map of net stock of migrants workers measured in percentage

Figure 7: Net Stock (%) of migrants by city in China

(a) Net Stock(%) of Workers in 2005 (b) Net Stock(%) of Workers in 2010



A.2 Quality-adjusted Housing Rents and Migration

In this section, we investigate the relation between quality-adjusted housing rents and migration. Using *Census* data, we run a simple regression as follows:

$$rent_{ij} = \beta_0 + \beta_1 NetMig_j + \mathbf{Z_{ij}}'\alpha + \epsilon_{ij} \tag{22}$$

$rent_{ij}$ is the housing rent of house i in city j. $NetMig_j$ is the net stock of migrant workers in city j, with a unit of 10k. $\mathbf{Z_{ij}}$ is a vector of housing characteristics for house i, including whether the house is also used as a business facility, the total area of the house, the number of the floors, the construction structure of the house, the building year of the house, the main cooking equipment, whether it has a tap water system, whether it has an independent kitchen, the type of restroom, and the type of showering system. We run the same regression separately for the year 2005 and the year 2010 using the *Census* data. The results in Table 17 show that a 1 million increase in the number of net stock of migrant workers is correlated with a 13.6 RMB (about 1.9 USD) increase in the annual rent per square meter in 2005. Similarly, a 1 million increase in the number of net stock of migrant workers is correlated with a 4.8 RMB (about 0.7 USD) increase in the annual rent per square meter in 2005. This shows that the positive relation between housing rents and the net stock of migrant workers in the city is robust even when we control for the quality of the houses.

Table 17: The Relation between Housing Rents and Migration

Variables	(1) OLS-2005	(2) OLS-2010
Net Stock of Migrant Workers (10k)	0.0113***	0.00396***
	(0.000173)	(0.0000516)
Observations	81,051	150,298
R-squared	0.207	0.181

Notes: ***p<0.01, **p<0.05, and *p<0.1.

A.3 Additional Results of Inequality from CHIP

In this section, we investigate the inequality between migrants and local residents in more detail. The *Census* is a comprehensive survey, but it does not contain too much information about a household's financial status, income, or expenditure. In the main context, we only have housing rents and wages, which are imputed from the *City Statistic Yearbooks*. We now introduce another dataset called the *Chinese Household Income Project (CHIP)* to further consider this inequality.[20] In 2013, CHIP covers 18,948 households in 15 provinces. After data cleaning in which we keep only urban observations, we have a sample size of 7,400 households. In these 7,400 households, there are 344 rural migrant families (migrant families from rural areas), 223 urban migrant families (migrant families from urban areas), and 6,833 local families.

Table 18: Quantile Statistics

Variable	10%	25%	50%	75%	90%
Non-housing Asset Distribution (RMB)					
Locals	12000	30000	69700	154800	304500
Rural Migrants	7000	18925	40750	98400	185500
Urban Migrants	15000	32500	70000	140000	372000
Net Asset Income Distribution (RMB)					
Locals	-13000	0	10000	39600	66444
Rural Migrants	-10000	0	0	1000	20000
Urban Migrants	-12634	0	0	24000	60000
Expenditure Distribution (RMB)					
Locals	17000	25000	38000	56000	80000
Rural Migrants	12000	20000	30000	48548	77250
Urban Migrants	15200	28000	40500	74000	95000
Savings Rate Distribution					
Locals	3.2%	19.5%	37.4%	53.2%	65.3%
Rural Migrants	11.1%	25.0%	43.2%	60.1%	72.7%
Urban Migrants	6.3%	23.6%	41.4%	53.8%	66.7%

Table 18 shows the distributions of different household-level variables. Non-housing assets is

[20]For more details of this dataset, please refer to Li, Sato, and Sicular (2013).

the total value of the non-housing assets of a household. Net asset income is defined as the difference between total disposable income and wages of the household members. Savings rate is calculated as the ratio of income less expenditure to income. Rural migrants have fewer non-housing assets, less net asset income, and less expenditure. Nevertheless, they save more compared with urban migrants and local residents. In addition, although urban migrants have more non-housing assets, they still have much less net asset income than local residents. This indicates that a very important part of the net asset income of local residents is their housing rent, which results in significant gaps and inequalities in the income and expenditure between local residents and rural migrants.

B Model Appendix

B.1 Estimation of Migration Elasticity

We estimate the migration elasticity (ϵ) from the gravity equation of migration flow (14). We assume $\tau_{ij}^{s} = \tau_{i}^{s} d_{ij}$, where τ_{i}^{s} is the origination-skill fixed component and d_{ij} is the distance index between location i and j. Under these assumptions and given data on migration shares and real incomes, we estimate ϵ using the fixed effect regression:

$$ln(\pi_{ij}^{s}) = \epsilon ln(v_{j}^{s}) + \psi_{ij} + \gamma_{is} + \zeta_{j} + \phi_{ijs}, \text{ for } i \neq j \tag{23}$$

where $\psi_{ij} = -\epsilon ln(d_{ij})$ is the origination-destination pair fixed effect, $\gamma_{is} = -\epsilon ln(\tau_{i}^{s}) - ln(\Phi_{i}^{s})$ is the origination-skill fixed effect, $\zeta_{j} = -\epsilon(1-\beta) ln(Q_{j})$ is the destination fixed effect, and ϕ_{ijs} is the measurement error term. We assume that the error term ϕ_{ijs} is not correlated with $ln(v_{j}^{s})$ after controlling for all these fixed effects.

To estimate ϵ, we need to run a regression estimating (23) with origination-destination pair fixed effects ψ_{ij}, origination-skill fixed effects γ_{is}, and destination fixed effects ζ_{j}. We use migration flows and housing rent data from the *Census* in 2005 and city-skill level average wage data imputed from the *City Statistic Yearbooks*. To calculate π_{ij}^{s} for each origination-destination city pair, we sum up the number of current workers who migrated from each origination city to each destination city by skill groups (with/without a college degree). $ln(v_{j}^{s})$ are different for residents with a local Hukou registration and migrant residents without a local Hukou registration. For migrants, income is the sum of their wages and their housing incomes in their Hukou locations. However, for local incumbents with housing assets, income is a combination of wages and local housing rent incomes. Housing rent incomes are constructed as explained in section 3. Because there are many zero migration flows between small city pairs, $ln(\pi_{ij}^{s})$ actually contains many missing values which are not used in the regression. Hence, we construct $\widehat{ln(\pi_{ij}^{s})}$ by assigning an extremely small value (i.e., 1e-7) to the migration flow and then estimating the same regression with $\widehat{ln(\pi_{ij}^{s})}$[21].

Table 19: Regression of Estimating the Migration Elasticity

Variables	(1)	(2)
$ln(v_j^s)\{Census\}$	1.847***	
	(0.0761)	
$ln(v_j^s)\{CSYB\}$		1.926***
		(0.138)
Origin-Destination FE	YES	YES
Origin-Skill FE	YES	YES
Observations	164,738	137,186
R-squared	0.568	0.577

Notes: Column 1 shows the results when the independent variable is $ln(_j^s)$. Column 2 shows the results when the independent variable is $\widehat{ln(v_j^s)}$, which includes the rebate of housing costs back to local residents. ***p<0.01, **p<0.05, and *p<0.1.

The results are shown in Table 19. Column 1 shows the results of directly regressing migration flows from city i to city j with skill s $\widehat{ln(\pi_{ij}^{s})}$ on destination-skill average income $ln(v_j^s)\{Census\}$ with wages measured from the individual wage in the original *Census 2005*. This gives us a statistically significant estimate of the migration elasticity of 1.847 with a standard error of 0.0761. However, to closely match the model, we run a second regression, which uses the destination-skill average income $ln(v_j^s)$ with wages measured from the City Statistic Yearbook *CSYB 2005*. The results are in column (2) of Table 19, which gives us an estimate of 1.926 with a standard error of 0.138. Our estimates are slightly larger than the estimate of around 1.5 in Tombe and Zhu (2019), which uses province-level data. As our model actually uses the wage data from City Statistics Yearbooks, we prefer to choose ϵ towards the estimation from the second regression using $ln(v_j^s)\{CSYB\}$, therefore, we pick $\epsilon = 1.90$[22].

B.2 Algorithm for Counterfactual Analysis

Given exogenous variables and parameters, we need to calculate the response of the endogenous variables resulting from policy changes. As we have mentioned, we will select the equilibrium that is the closest to the one in the real world. Thus, the variables' initial values will be set equal to the model result in 2010.

[21] The estimation results are robust to the choice of the extreme small value.

[22] The true parameter is very likely to be somewhere between the two estimators. Also, as robustness checks, we solved several models under a variety of parameter choices from 1.5 as in Tombe and Zhu (2019) to 2.0 which is slightly higher than our estimation. In all cases, all the results hold as in the paper, though the magnitudes changes slightly. The results are available upon request.

We first specify the exogenous variables and the model equation system. The exogenous variables are $\{H_i^s, A_j^s, \tau_{ij}^s, L_j, \phi_j\}$ where i indexes origination cities, j indexes destination cities, and s indexes skill. The equation system consists of three blocks. The migration block consists of worker income equation (8), and gravity equation (14), the production block consists of production equation (15) and wage equations (16, 17), and the housing block consists of construction equation (20) and market clearing equation (21).

To calculate the policy counterfactuals, we start with the block in which changes occur and then iterate block by block to update the endogenous variables until all endogenous variables converge. We present the process of calculating a counterfactual, using the relaxation of construction intensity as an example.

Suppose a policy that increases construction intensity by 20%. That is, $\hat{\phi}_j = 1.2 \times \phi_j$ for every city j. We have the following process of updating variables ($\{\hat{X}_j\}^t$ indicates t's iteration of variable X). Starting with the housing block:

$$\{\hat{S}_j\}^1 = \hat{\phi}_j L_j \text{ from eq.(20)} \tag{24}$$

$$\{\hat{Q}_j\}^1 = \frac{1-\beta}{\beta} \frac{w_j^l H_j^l + w_j^h H_j^h}{\{\hat{S}_j\}^1} \text{ from eq.(21)} \tag{25}$$

Now we move to worker's migration choices (migration block):

$$\{\hat{v_{ij}^s}\}^1 = w_j^s + \frac{\{\hat{Q}_i\}^1 \{\hat{S}_i\}^1}{H_i^R} \text{ from eq.(8)} \tag{26}$$

$$\{\hat{\pi_{ij}^s}\}^1 = \frac{(\tau_{ij}^s \{\hat{Q}_j\}^{1^{1-\beta}})^{-\epsilon} (\{\hat{v_{ij}^s}\}^1)^{\epsilon}}{\sum_{k=1}^K (\tau_{ik}^s \{\hat{Q}_k\}^{1^{1-\beta}})^{-\epsilon} (\{\hat{v_{ik}^s}\}^1)^{\epsilon}} \text{ from eq.(14)} \tag{27}$$

Then, combining $\{\hat{\pi_{ij}^s}\}^1$ with $\{H_i^s\}$, we are able to calculate $\{\hat{H_j^s}\}^1$. Finally, we move to the production block to calculate wages:

$$\{\hat{X}_j\}^1 = [(A_j^h \{\hat{H_j^h}\}^1)^{\frac{\sigma-1}{\sigma}} + (A_j^l \{\hat{H_j^l}\}^1)^{\frac{\sigma-1}{\sigma}}]^{\frac{\sigma}{\sigma-1}} \text{ from eq.(15)} \tag{28}$$

$$\{\hat{w_j^l}\}^1 = A_j^{l^{\frac{\sigma-1}{\sigma}}} \{\hat{X}_j\}^{1^{\frac{1}{\sigma}}} \{\hat{H_j^l}\}^{1^{-\frac{1}{\sigma}}} \text{ from eq.(16)} \tag{29}$$

$$\{\hat{w_j^h}\}^1 = A_j^{h^{\frac{\sigma-1}{\sigma}}} \{\hat{X}_j\}^{1^{\frac{1}{\sigma}}} \{\hat{H_j^h}\}^{1^{-\frac{1}{\sigma}}} \text{ from eq.(17)} \tag{30}$$

So far we have updated all the endogenous variables once. We calculate how far $\{\hat{x}_j\}^1$ is from $\{\hat{x}_j\}^0$, where x means any specific variable. If the distance is large, we go back to eq.(24) and eq.(25) to iterate until the distance is small enough. For other counterfactuals, the starting block of iteration may differ, but the general algorithm is identical. The key is to update all the endogenous variables in a loop. We terminate the iteration loop when all the aggregate variables reach an updating error smaller than 1e-7.

C Counterfactual Analysis Appendix

C.1 Property Tax and Redistribution: Additional Tables

Table 20: **Skill Premium and Housing Premium: Property Tax**

Net Migrant	No. of	Skill Premium			Housing Premium		
Range(2010)	Cities	2010	$\widehat{2010}$	Relative	2010	$\widehat{2010}$	Relative
Average	233	1.40	1.40	100%	0.49	0.46	94%
(6,13)	5	1.39	1.39	100%	1.89	1.60	85%
(1,6)	19	1.40	1.43	102%	0.56	0.51	91%
(0, 1)	45	1.39	1.38	99%	0.35	0.34	97%
(-1,0)	134	1.40	1.39	99%	0.25	0.26	104%
(-4,-1)	30	1.45	1.43	98%	0.31	0.21	68%

Notes: This table displays population-weighted means in 2005 and 2010. Skill Premium is measured as annual high-skill wage over annual low-skill wage for each city, and Housing Premium is measured as average annual housing return over the average annual wage for each city.

C.2 Directly Increase Land Supply by Migration Inflows

In Table 21, we consider an alternative counterfactual which directly increases land supply in larger cities proportional to migration inflows but without the trade of land quotas across cities. Since most Chinese cities (except Shenzhen and Dongguan) retain a large portion of farmland, this counterfactual is generally feasible. This counterfactual is to increase the total land supply increment from 2005 to 2010 proportional to positive migration inflows. As a result, cities with positive net inflows keep the same worker-land ratio as in 2005, while cities losing workers do not lose the land quotas.

Table 21: **Counterfactual Construction Land Supply (Directly Increment in Land)**

Net Migrant	No. of	Land Supply (Data)				Counterfactual		
Range(2010)	Cities	2005	2010	Relative	Changes	$\widehat{2010}$	$\widehat{Relative}$	$\widehat{Changes}$
National	233	24,277	31,705	131%	+7,428	39,133	161%	+14,856
(6,13)	5	5,135	5,648	110%	+513	10,389	202%	+5,254
(1,6)	19	3,801	5,912	155%	+2,111	10,461	275%	+6,660
(0, 1)	45	5,555	7,250	131%	+1,695	8,103	145%	+2,548
(-1,0)	134	7,950	10,363	130%	+2,413	8,026	101%	+76
(-4,-1)	30	1,836	2,532	138%	+696	1,836	100%	+0

Notes: This table displays the total land supply data by migration groups in 2005 and 2010, as well as the counterfactual land supply in 2010 (unit: km^2). Range is classified by net stock of migrant workers in 2010 as in the data (unit: millions). Each Net Migrant Range Group consists of the same cities in 2005 and 2010.

Table 22: **Migration Flow and Housing Cost: Direct Land Supply Increment**

Net Migrant	No. of	Net Migrant			Housing Cost		
Range(2010)	Cities	2010	$\widehat{2010}$	Relative	2010	$\widehat{2010}$	Relative
Overall	233	96m	111m	116%	114	93	82%
(6,13)	5	+45m	+54m	120%	226	145	64%
(1,6)	19	+38m	+46m	121%	136	84	62%
(0, 1)	45	+13m	+12m	108%	118	98	83%
(-1,0)	134	-48m	-48m	100%	87	87	100%
(-4,-1)	30	-48m	-63m	131%	80	72	90%

Notes: This table displays total net stock of migrant workers and population weighted average housing costs for each city group. In the first row (Overall), we show the number of workers who have migrated and the national population weighted average housing cost. The unit of the net migrant is millions, and the unit of housing costs is Chinese Yuan (RMB) per square meters per year.

Table 23: **Within-city Theil Index: Direct Land Supply Increment**

Net Migrant	No. of	Wage Theil Index			Income Theil Index		
Range(2010)	Cities	2010	$\widehat{2010}$	Relative	2010	$\widehat{2010}$	Relative
Average	233	0.0070	0.0072	103%	0.0184	0.0245	133%
(6,13)	5	0.0097	0.0092	95%	0.0908	0.1189	131%
(1,6)	19	0.0079	0.0088	111%	0.0223	0.0275	123%
(0, 1)	45	0.0083	0.0083	100%	0.0092	0.0097	105%
(-1,0)	134	0.0058	0.0059	101%	0.0052	0.0051	98%
(-4,-1)	30	0.0058	0.0056	97%	0.0062	0.0066	106%

Notes: This table displays population-weighted means of both inequality measures. The original equilibrium is 2010 and the counterfactual equilibrium is $\widehat{2010}$. Relative is calculated via dividing $\widehat{2010}$ by 2010.

Table 24: **Share of National Theil Index: Direct Land Supply Increment**

Net Migrant Range(2010)	No. of Cities	Share of Wage Theil 2010	Share of Wage Theil $\widehat{2010}$	Share of Wage Theil Relative	Share of Income Theil 2010	Share of Income Theil $\widehat{2010}$	Share of Income Theil Relative
National Theil	233	0.062	0.067	108%	0.092	0.104	113%
(6,13)	5	+1.41	+1.44	102%	+1.27	+1.25	98%
(1,6)	19	+0.83	+0.85	102%	+0.70	+0.68	97%
(0, 1)	45	+0.26	+0.25	96%	+0.20	+0.17	85%
(-1,0)	134	-1.00	-0.95	95%	-0.78	-0.68	87%
(-4,-1)	30	-0.49	-0.58	118%	-0.39	-0.42	108%

Notes: This table displays city groups' contribution to the national Theil Index in 2005 and 2010. The first row displays the national wage/income Theil Index in 2005 and 2010.

Table 25: **Skill Premium and Housing Premium: Direct Land Supply Increment**

Net Migrant Range(2010)	No. of Cities	Skill Premium 2010	Skill Premium $\widehat{2010}$	Skill Premium Relative	Housing Premium 2010	Housing Premium $\widehat{2010}$	Housing Premium Relative
Average	233	1.40	1.40	100%	0.49	0.64	131%
(6,13)	5	1.39	1.39	100%	1.89	2.78	147%
(1,6)	19	1.40	1.43	102%	0.56	0.63	112%
(0, 1)	45	1.39	1.38	99%	0.35	0.35	100%
(-1,0)	134	1.40	1.39	101%	0.25	0.25	100%
(-4,-1)	30	1.45	1.43	102%	0.31	0.18	58%

Notes: This table displays population-weighted means in 2005 and 2010. Skill Premium is measured as annual high-skill wage over annual low-skill wage for each city, and Housing Premium is measured as average annual housing return over the average annual wage for each city.