Housing Prices, Internal Migration, and Intergenerational Mobility*

Qingyuan Chai[†]

September 10, 2024

Abstract

This paper examines how housing affordability affects intergenerational mobility in China by influencing internal migration. Using the Housing Purchase Restriction (HPR) policy as a natural experiment and employing an instrumental variable (IV) approach, I find that rising housing costs deter migration, with a more pronounced effect on children from disadvantaged families. Consequently, these children are less likely to migrate and subsequently earn lower incomes than their counterparts from more affluent families, thereby reducing intergenerational mobility. These findings are consistent with a migration decision model where housing expenses disproportionately burden individuals from less affluent backgrounds. To further explore the policy implications, a bare-bone spatial equilibrium model is developed and estimated to evaluate the effects of various housing policies. Overall, the results underscore the crucial role of parental background in determining children's access to labor market opportunities, even when educational attainment and economic conditions are held constant.

Keywords: Housing Market, Migration, Inequality, Intergenerational Mobility, Structural Transformation

JEL Codes: J62, R23, R31

*I am indebted to my committee members, Kevin Lang, Daniele Paserman, and Raymond Fisman, for their guidance and encouragement. For detailed comments and suggestions, I also thank Jérôme Adda, Huiyao Chen, Oren Danieli, Hanming Fang, Adam Guren, Douglas Gollin, Joseph Gyourko, Patrick Kline, Dilip Mookherjee, Yuhei Miyauchi, Sangyoon Park, Maisy Wong for helpful comments and discussions. All errors are my own.

[†]PhD Candidate, Department of Economics, Boston University, email: qchai@bu.edu

1 Introduction

Factors that influence migration decisions are crucial in shaping intergenerational mobility, as they distinctly impact families based on their economic status. If children from lower-income families are less able to migrate and access labor market opportunities, intergenerational mobility may decline. This paper focuses on one such factor: housing prices. I examine the role of housing prices in shaping internal migration and intergenerational mobility in China. The results imply that, with elevated housing prices, migration decreases, and the decrease is larger for children from less advantaged families. Consequently, these disadvantaged children earn lower incomes compared to their counterparts from more affluent families, leading to a reduction in intergenerational mobility.

While the literature extensively documents intergenerational persistence, its determinants, especially in developing countries, remain unclear (Genicot et al., 2024). Most research focuses on correlates of mobility, which, although indicative of institutional and policy influences, do not establish causality. Yet, understanding the determinants is essential for improving intergenerational mobility through public policies. This paper moves beyond descriptive associations, providing causal evidence that rising housing prices decrease intergenerational mobility. The framework developed here is relevant, given the recent surge in housing prices in many countries.

The impact of housing prices on intergenerational mobility is ex-ante ambiguous. On the one hand, elevated housing prices can reduce intergenerational mobility by disproportionately preventing disadvantaged children from migrating and earning higher incomes. Migration involves upfront residential costs, as migrants often face a job search period before securing employment. These costs can be especially burdensome for low-income families, who often face credit constraints (Eggleston et al., 2018; Cai, 2020; Gai et al., 2021). Consequently, even when children receive similar education and encounter comparable economic conditions, their labor market outcomes can still be significantly influenced by their family backgrounds. This suggests that housing affordability can amplify existing inequalities rooted in family background by restricting migration opportunities.

Alternatively, because migration is very costly, most children from poor backgrounds cannot afford to migrate. The cost barrier may be so high that small changes in migration costs have little impact on their decisions. In contrast, children from wealthier families might be closer to the decision threshold, making them more responsive to fluctuations in housing prices.[1]

[1]Another possibility is that children from wealthier families expect higher housing quality. As price

Overall, the impact of housing prices on intergenerational mobility is theoretically ambiguous. I develop a migration decision model to clarify the conditions under which housing prices may either facilitate or hinder intergenerational mobility.

In this paper, I use China as a testing ground to examine how real estate shocks influence intergenerational mobility through the migration channel. China offers a unique setting for several reasons. First, the country has experienced rapid housing price appreciation over the past few decades (Fang et al., 2016), highlighting the significant impact of real estate fluctuations. Housing affordability has become a critical issue for migrants, who complain that "We have no home where there is work, and no work where there is home." Second, the real estate boom varies substantially across different prefectures, providing rich variation to analyze its effects. The Housing Purchase Restriction (HPR) policy implemented in some prefectures offers a natural experiment to identify the causal impact of real estate shocks. Finally, China's high return to migration is well-documented, and migration is a crucial means for economic advancement, especially for workers from rural areas (see, e.g., Lagakos et al., 2020).

The phenomenon of rising housing prices impeding migration is also observed in other countries, making this conceptual framework relevant in broader contexts. For instance, recent years have seen rising housing prices in countries such as India and the United States (Olney and Thompson, 2024; Mahadevia et al., 2012). The impact of housing affordability on migration may vary across different parental backgrounds, thereby affecting intergenerational mobility.

The empirical results suggest that rising housing costs decrease intergenerational mobility by disproportionately preventing children from disadvantaged backgrounds from migrating. I find that higher housing costs reduce migration, with the effect being less pronounced for individuals whose fathers have higher education levels. Similarly, the increase in housing cost barriers leads to lower incomes, but this negative impact is again mitigated for those with more educated fathers.

As housing prices are correlated with economic conditions, the OLS results are subject to omitted variable bias. For instance, fathers with higher education levels may be more effective at leveraging their social networks to secure job opportunities for their children in destination prefectures. As a result, the coefficients might reflect fathers' differential capacity to facilitate their children's access to employment rather than the heterogeneous impacts of housing prices.

To address this concern, this study employs a natural experiment facilitated by the Housing Purchase Restriction policy introduced in China around 2010. By restricting

changes for high-quality homes tend to be larger, rising housing prices could translate into higher migration costs for these children, potentially discouraging their migration.

the number of properties that each household or firm could buy, this policy led to an immediate and sharp decrease in housing demand. Since the policy was launched in only 46 prefectures, its impact can be assessed by analyzing outcome variation before and after the policy's implementation across both the affected and unaffected prefectures. This policy has been used in many existing papers as an exogenous shock on housing prices (Chen et al., 2017; Zhao and Zhang, 2022; Liu et al., 2023; Chen et al., 2023). In this paper, the HPR policy is exploited to construct an instrumental variable for housing prices.

The results remain robust after controlling for various economic conditions at both origin and destination prefectures, accounting for the interaction between children's education levels and housing prices, and excluding prefectures with distinct migrant behaviors or those affected by housing price spillovers from megacities. To address the potential concern that fathers' education may not be comparable across prefectures or years, I replicate the analyses using a residualized measure of fathers' education, adjusted for *Hukou* type and prefecture-year fixed effects. Additionally, I test the robustness by replacing fathers' education with imputed fathers' income or household wealth. All robustness checks yield results similar to the main analyses.

In terms of heterogeneity, the results highlight the importance of basic education in overcoming housing cost barriers. The critical factor is whether the father has any schooling at all. Junior high school also seems to help, although the impacts are noisily estimated. Once a father has junior high school education, higher levels, such as junior high school or beyond, provide few additional advantages in overcoming these barriers.

Further analyses reveal that the effects are more pronounced for children with rural parents, suggesting that rural-urban migration is most affected by housing cost barriers. Since rural-urban migration is often associated with non-agricultural employment, the findings also show that housing prices impact agricultural employment differently depending on parental background. This points to an interaction between housing costs, structural transformation, and intergenerational mobility.

The remainder of the paper proceeds as follows. Section 2 discusses the related literature. Section 3 presents background information. Section 4 lays out a model. Section 5 describes the empirical approach. Section 6 introduces the data. Section 7 presents the main results. Section 8 builds and estimates a spatial equilibrium model for counterfactual analyses, and Section 9 concludes.

2 Relation to the Literature

First, it speaks to the literature that investigates the determinants of intergenerational mobility. While the phenomenon of intergenerational persistence is well-documented, the factors driving it are not fully understood. Previous studies highlight the importance of access to education, local labor market conditions, social capital, and economic shocks (for example, Feigenbaum, 2015; Olivetti and Paserman, 2015; Parman, 2011; Zheng and Graham, 2022; Tan, 2023). This paper suggests that despite equal education and economic conditions, a father's socioeconomic status can still shape their children's ability to seize labor market opportunities, thereby impacting their outcomes. Additionally, existing literature predominantly investigates developed countries; this paper focuses on the developing world, where geographic inequality and migration potentially play crucial roles in shaping intergenerational mobility.

Close to this study, Ward (2022) provides evidence that migration enhances intergenerational mobility, especially noting its effectiveness for individuals from lower-income families. Additionally, he finds that migration was far more effective than education at allowing children to escape poverty. This paper complements his work by examining a different setting and focusing on the impact of elevated housing prices, a trend currently witnessed in many countries.

Furthermore, this paper connects with research on policies that promote migration to reduce inequality. For example, Chetty et al. (2016) examine the Moving to Opportunity (MTO) experiment, highlighting the importance of childhood exposure to better environments through family migration.[2] In contrast, this study focuses on policies that could enhance intergenerational mobility by encouraging migration in adulthood. This distinction underscores different life stages where targeted interventions—childhood versus adulthood—could help break the cycle of poverty and dependence on parental socioeconomic status.

Secondly, this study contributes to the literature on the consequences of changes in house prices amidst China's rapid economic transformation. Recent studies have discussed the effects of house price growth on various household decisions in China, including consumption and saving, labor supply, adult children's co-residence with parents, marriage, and fertility (see, e.g., Wei and Zhang, 2011; Wang and Wen, 2012; Li and Wu, 2019; Li et al., 2020; Sun and Zhang, 2020; Alm et al., 2022; Liu et al., 2023).

Thirdly, this paper relates to the literature on the driving forces of internal migration

[2]Similarly, Alesina et al. (2021) find that in Africa, an additional year in a high-mobility region between ages 5 and 12 significantly increases the likelihood of children from uneducated families completing primary school.

in China. Prior research has emphasized the impacts of many factors such as *Hukou* registration system, the land tenure system, and pollution (see, for example, Zhao, 1999; Whalley and Zhang, 2007; Chen et al., 2018; Tombe and Zhu, 2019; Jin and Zhang, 2023; Khanna et al., 2021). This study complements the literature about the impacts of housing prices on internal migration (Meng et al., 2023). The paper also connects to the literature on housing prices and migration decisions in general. Particularly, in the context of the United States, Ganong and Shoag (2017) find that the deceleration of low-skilled workers relocating to high-cost housing areas can help to explain the decline of the geographic wage convergence from 1980 to 2010.[3] This paper reveals that the effects of housing affordability on migration vary according to parental background, which in turn impacts intergenerational mobility.

Closely related to this study, Garriga et al. (2023) investigates the interrelationship between urbanization, structural transformation, and the housing market through the lens of a dynamic spatial equilibrium model. Consistent with this paper, they find that rising house prices stunt migration. This paper focuses on the impacts of housing costs on intergenerational mobility by affecting migration. The results imply an interaction between housing costs, intergenerational mobility, and structural transformation. Moreover, an instrumental variable is constructed to address the endogeneity of housing prices.

Finally, this paper engages with the literature on poverty traps (Azariadis and Stachurski, 2005) and, specifically, how credit constraint significantly impedes migration in developing contexts. Existing papers have documented that credit constraint is an important barrier to both internal migration (Bryan et al., 2014; Angelucci, 2015; Bazzi, 2017). Fewer studies have pointed out the impacts of credit constraints for internal migration (Cai, 2020; Gai et al., 2021). While internal migration may cost less than international migration, this paper suggests that the barrier remains substantial and can affect intergenerational persistence. Due to credit constraints and the risks associated with migration, many households remain trapped in low-income situations and are hesitant to migrate.

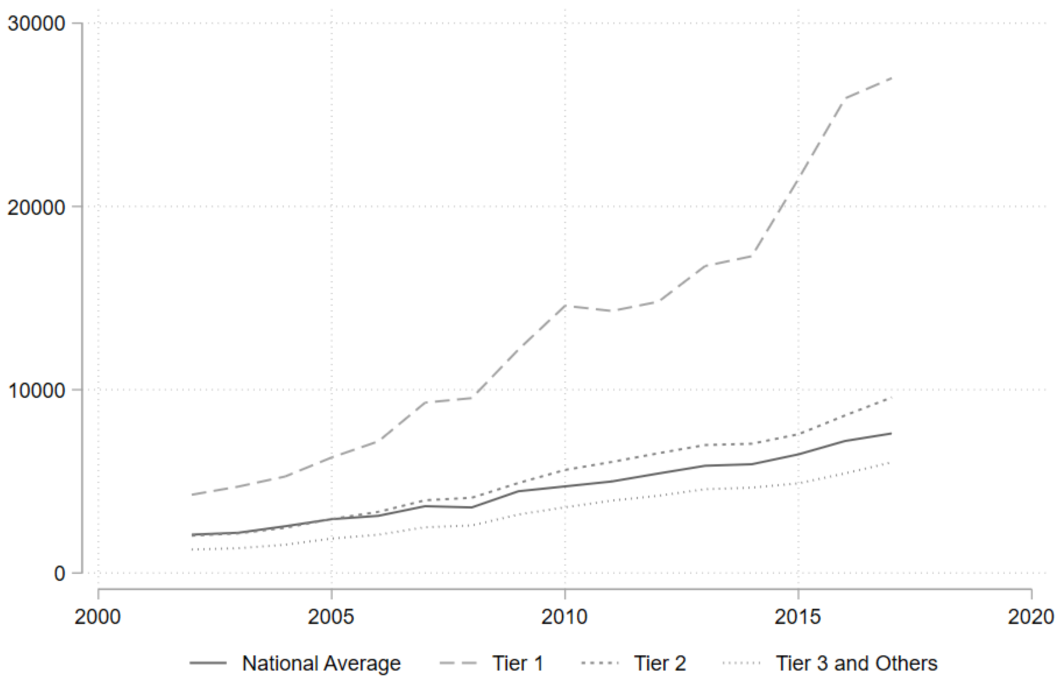
[3] Additionally, some papers explore the heterogeneous impact of housing prices on renters' and owners' migration decisions, which take housing price shocks as income shocks rather than barriers to migration (Zabel, 2012; Foote, 2016).

3 Background

3.1 Housing Boom in China

The 1998 reform, which privatized the housing sector, initiated a rapid growth of the real estate markets in urban China. As shown in Figure 1, housing prices in China have surged dramatically over the past two decades. Even after accounting for housing quality, from 2002 to 2013, the average yearly growth rate of residential housing prices in 35 major Chinese prefectures was 11.4 percent (Sun and Zhang, 2020). Such increases have been attributed to factors such as land supply restrictions, increases in mortgage supply due to fiscal stimulus, expected economic growth, etc. (see, for example, Wang and Wen, 2012; Fang et al., 2016; Sun and Zhang, 2020).

Figure 1: Residential Housing Prices by Prefecture Tier (yuan/sq.m)



NOTE: This figure presents the nominal yearly housing prices across prefectures of different tiers from 2002 to 2017. Tier-1 prefectures are comprised of Beijing, Shanghai, Guangzhou, and Shenzhen. Thirty-one prefectures are classified as Tier-2. Other prefectures fall into Tier 3.
SOURCE: Rogoff and Yang (2020)

In tandem with the rise in housing prices, rental costs also surged, a factor of particular concern to migrants. According to data from the China Migrants Dynamic Monitoring Survey (CMDS), the median monthly rent in Tier 1 prefectures doubled from 500 RMB in 2011 to 1000 RMB in 2017, not taking into account the potential decrease in residential quality that migrants might accept to manage higher costs. After adjusting for the Consumer Price Index (CPI), this rise translates to a jump from 403 to 717 in 2000 RMB.

Despite the overall upward trend in housing prices over the past few decades, there are significant temporal and regional disparities across prefectures. This variation can be attributed to factors such as the differential housing supply elasticity, diverse local house purchasing policies, and spatial inequality in economic development (Chen and Wang, 2015; Saiz, 2010; Gyourko et al., 2022). Appendix Figure B.1 illustrates this variation by showing the residualized housing prices, which are derived from regressing housing prices on prefecture fixed effects and year fixed effects.

3.2 Housing Purchase Restriction Policy

On April 17, 2010, the State Council of China released the "Notice of the State Council on Resolutely Curbing the Excessive Rise of House Prices in Some Prefectures" aimed at moderating the surging housing prices in some urban regions. After this notice, Beijing pioneered the implementation of a Housing Purchase Restriction (HPR) policy on May 1, which limited households to buying no more than one new property. Subsequently, this policy was extended to an additional 45 major prefectures across China from late 2010 to early 2011, as detailed in Table B.1 (dates are from Chen et al. (2017)). While the specific regulations varied among these prefectures, the general framework of the HPR policy typically restricted Chinese households to owning no more than two properties, with a mandatory two-year interval between purchases. The policy was in place until 2014, after which it began to be slowly removed.

The policy was adopted abruptly and unexpectedly. Immediately following the implementation of the HPR policy, housing demand declined noticeably, leading to slower price growth and a reduction in the number of houses sold in the affected prefectures (Du and Zhang, 2015; Li et al., 2017). This paper utilizes this slowdown in the housing market, which affects migrants' housing affordability and, thus, their migration decisions.

4 Model

How do housing prices affect migration decisions and intergenerational mobility? To explore the relevant factors, I construct a model of migration decisions. The model indicates that housing price changes can either increase or decrease migration and intergenerational mobility, depending on parameters and specific distributions of disutility regarding migration.

4.1 Setup

Table 1: Children's Income by Parental Backgrounds and Migration Status

		Y: Stay	Y: Migrate
F	H	Y_H^0	Y_H^1
	L	Y_L^0	Y_L^1

In this section, I describe a discrete version of the model for simplicity. A continuous version can be found in Appendix Section A, with similar main results. Assume fathers' socioeconomic status is high (H) or low (L). If the child stays, their income is Y_H^0 or Y_L^0, depending on the father's status. If the child migrates, their income is Y_H^1 or Y_L^1. Table 1 shows children's income by parental background and migration status.

The model makes the following assumptions. First, it posits that all children are homogeneous, and it is the father's type, F, that affects the children's income, Y. Second, $Y_L^1 > Y_L^0$ and $Y_H^1 > Y_H^0$, meaning that migration increases income. Third, the model assumes that incomes at the origin and destination are realized after the migration decision. That is, the migration decision is made based on the expected income from either staying or migrating rather than on actual income.

Migration incurs a cost, denoted as $M(h, F)$, which depends on the housing cost h and the father's type F. Assume that $\frac{\partial M(h,F)}{\partial h} > 0$, meaning that an increase in housing costs increases migration costs. In addition to financial costs, individuals experience psychological costs associated with migration, referred to as "disutility," denoted by τ. This disutility τ is independent of F and realized Y. The cumulative distribution function (CDF) and probability density function (PDF) for τ are denoted as $F_\tau(\tau)$ and $f_\tau(\tau)$, respectively.

Combining the above benefits and costs of migration, a child will migrate if:

$$Y_F^1 - Y_F^0 - M(h, F) > \tau \tag{1}$$

where F can be either H or L.

Denote the left-hand side as $B(h, F)$, which is the net benefit of migrating. For a child faced with father's type F and housing cost h, the probability that they migrate is $Pr(migrate|h, F) = Pr(B(h, F) > \tau) = F_\tau(B(h, F))$.

In the regression analysis, I regress the indicator for migration on the interaction term of paternal socioeconomic status and housing cost. As I am analyzing a binary model, the sign of the interaction term corresponds to the sign of $\frac{\partial Pr(migrate|h,F)}{\partial h}\Big|_{F=H} - \frac{\partial Pr(migrate|h,F)}{\partial h}\Big|_{F=L}$. It can be derived that:

$$\frac{\partial Pr(migrate|h,F)}{\partial h} = -f_\tau(B(h,F))M'_F(h) < 0 \tag{2}$$

and

$$\frac{\partial Pr(migrate \mid h,F)}{\partial h}\Big|_{F=H} > \frac{\partial Pr(migrate \mid h,F)}{\partial h}\Big|_{F=L} \iff \frac{M'_H(h)}{M'_L(h)} < \frac{f_\tau(B_L)}{f_\tau(B_H)} \tag{3}$$

where B_F denotes $B(h,F)$, $M'_F(h)$ denotes $\frac{\partial M(h,F)}{\partial h} > 0$, and $F = H$ or L.

In the empirical analysis, I also examine the effects on children's income. In the model, the expected income of children, regardless of their migration status, is:

$$\begin{aligned} E[Y|h,F] &= F_\tau(B(h,F))Y_F^1 + [1 - F_\tau(B(h,F))]Y_F^0 \\ &= F_\tau(B(h,F))(Y_F^1 - Y_F^0) + Y_F^0 \end{aligned} \tag{4}$$

The difference in the effect of housing costs on income is $\frac{\partial E[Y|h,F]}{\partial h}\Big|_{F=H} - \frac{\partial E[Y|h,F]}{\partial h}\Big|_{F=L}$. It can be derived that:

$$\frac{\partial E[Y|h,F]}{\partial h} = -f_\tau(B(h,F))M'_F(h)(Y_F^1 - Y_F^0) < 0 \tag{5}$$

and

$$\frac{\partial E[Y|h,F]}{\partial h}\Big|_{F=H} > \frac{\partial E[Y|h,F]}{\partial h}\Big|_{F=L} \iff \frac{M'_H(h)}{M'_L(h)} < \frac{f_\tau(B_L)}{f_\tau(B_H)}\frac{(Y_L^1 - Y_L^0)}{(Y_H^1 - Y_H^0)} \tag{6}$$

4.2 Discussion

Equation (3) suggests that the impact of housing prices on migration decisions depends on two factors: the heterogeneous costs, $\frac{M'_H(h)}{M'_L(h)}$, and the different position at the disutility distribution, $\frac{f_\tau(B_L)}{f_\tau(B_H)}$. First, for the magnitude of $\frac{M'_H(h)}{M'_L(h)}$, if credit constraints exist, housing costs increase migration costs more for children with L type fathers, namely $M'_F(h)$ decreases in F. For $\frac{f_\tau(B_L)}{f_\tau(B_H)}$, $f_\tau(B_F)$ measures the sensitivity of disutility with respect to the net benefit of migration. If more children of H type fathers are at the margin of deciding whether to migrate or not, while fewer children of L type fathers are at the margin, $f_\tau(B_L)$ is smaller than $f_\tau(B_H)$, and vice versa.

Equation (6) suggests that the two mechanisms described above, heterogeneous migration costs and differential position at disutility distribution, also apply to the analysis of the heterogeneous impacts of housing costs on children's income. There is an additional term, $\frac{(Y_L^1 - Y_L^0)}{(Y_H^1 - Y_H^0)}$, which measures the relative benefit of migrating between children with L and H type fathers. I assume that parental background holds more value in the hometown, so $\frac{(Y_L^1 - Y_L^0)}{(Y_H^1 - Y_H^0)}$ is larger than one, meaning that migration is more beneficial for children with low-type fathers. Given this assumption, Equation (6) is a necessary condition of Equation (3). The underlying idea is that children from low-income backgrounds benefit more from migration, but they also suffer more when housing costs rise, leading to greater negative impacts on their income.

Overall, the theoretical signs of the differences in the effects of housing costs on migration decisions and children's income between those with high- and low-type fathers are undetermined. The empirical analysis is essential to determine which factors dominate in the data.

5 Empirical Approach

I begin the analyses by exploring how housing prices influence migration decisions and income, with a focus on how these effects vary according to fathers' years of education. To address potential identification concerns, I then employ a 2SLS estimation, using the Housing Purchase Restriction (HPR) policy to construct the instrumental variable.

5.1 Baseline

I run the following regression to examine the heterogeneous impacts of housing prices on migration and income across different parental backgrounds. For individual i born in year y whose origin is prefecture p, I estimate the following equation:

$$\begin{aligned} Outcome_{ipy} = \beta_1 FatherEduy_{ipy} \times HPgap_{py} + \beta_2 FatherEduy_{ipy} \\ + \beta_3 HPgap_{py} + \Pi Z_{py} + \Omega X_{ipy} + \mu_y + \eta_p + \epsilon_{ipy} \end{aligned} \tag{7}$$

where β_1 is the coefficient of interest. $Outcome_{ipy}$ can be either the indicator for migration or log income.[4] Migration is defined as residing in a different prefecture

[4]One potential concern is that, as there exists a sizeable rural-urban gap in the cost of living, the variation in income may capture mainly the price difference in different areas but not the real difference. To address this concern, I estimate the main results using income adjusted for rural-urban and provincial price differences based on the approach in Brandt and Holz (2006), and the results remain robust. I do not use adjusted income for the main results because, as noted in Albert and Monras (2022) and supported by anecdotal evidence, a significant portion of people's consumption occurs at their place of origin rather

from one's origin, with the origin defined as the prefecture where the individual lived at age 14.[5] $HPgap_{py}$ represents the housing price gap between the origin and potential destinations, which I describe in detail below. $FatherEduy_{ipy}$ denotes the individual's father's years of education. The individual-level control variables, X_{ipy}, include gender, parents' *Hukou* status, and education fixed effects. The origin prefecture-by-birth year level controls, Z_{py}, include employment and wage levels at the origin prefecture, as well as weighted averages of employment and wages at the destination prefectures. μ_y and η_p represent fixed effects for birth year and origin prefecture respectively.

Although I am using cross-section data from one single year, the survey has information on the migration history of the interviewee, including the origin, destination, and year of migration, allowing me to back out the housing prices that they face during the working ages. To accurately measure the effect of the housing market on migration decisions, ideally, I would like to pair each migration decision with the characteristics of the housing market at the time when the individual was contemplating whether to migrate. In practice, following Sun and Zhang (2020), the expected housing price gap the individuals were exposed to is calculated as follows.

First, using China's 2010 census sample data, I calculate the age-specific migration probability, $age\_specific\_migration\_prob_k$, which represents the proportion of individuals who migrated at age k relative to the total number of migrants. Appendix figure B.3 plots the age-specific migration probability rate.

$$age\_specific\_migration\_prob_k = \frac{\#\, of\ migrants\ migrated\ at\ age\ k}{\#\, of\ migrants}$$

To minimize the influence of parental migration decisions, the calculation starts at age 16. Additionally, the calculation ends at age 45 because data suggests that people barely migrate after 45. Adjusting the current thresholds of 16 and 45 yields similar results. If an individual is of age l in the survey year, they were exposed no later than age l, in which case l is the upper bound of the summation.

For an individual born in year y, at age l, originating from prefecture p, the expected housing price gap to which they were exposed is defined by the following expression:[6][7]

than at their destination.

[5]Results are robust using other ages such as 16 and 18.

[6]I assume the age-specific migration rate is homogeneous across prefectures. Results remain robust if I use province-specific migration rate. In addition, I assume the age-specific migration rate does not change over time. To mitigate the concern that the age-specific probability also captures birth cohort effects, I do the same analysis using the probabilities generated leveraging census data in 2000 and 2005. The results are very similar.

[7]For transparency and concerns about the functional form, results are robust using the housing price

$$HPgap_{py} = \frac{\sum_{k=16}^{min\{l,45\}} age\_specific\_migration\_prob_k \times (lndestprice_{y+k,p} - lnprice_{y+k,p})}{\sum_{k=16}^{min\{l,45\}} age\_specific\_migration\_prob_k}$$

$HPgap_{py}$ represents the weighted average of the housing price gap between the origin and potential migration destinations. To calculate the housing prices of the potential destinations, I use a weighted average where existing migration patterns determine the weights. The process starts with census data to identify the number of migrants between each pair of origin and destination prefectures. The proportion of migrants from a specific origin moving to each destination is then calculated and used as the weighting factor. For example, if 50% of migrants from prefecture p move to Shanghai, then Shanghai's housing price is assigned a weight of 50% in the calculation for prefecture p. The formula for calculating the weighted average housing price, $lndestprice_{pt}$, is as follows:

$$lndestprice_{pt} = \sum_d w_{pd} lnprice_{dt}$$

where w_{pd} is the fraction of migrants from origin prefecture p who migrated to prefecture d, and $\sum_d w_{pd} = 1$.[8] $lndestprice_{pt}$ is thus the weighted average of housing price at the potential destinations for origin prefecture p in year t.

Variation in housing price gaps arises from two sources: first, children born in the same prefecture but in different years encounter different housing price shocks at ages crucial for making migration decisions; second, children from different origin prefectures experience disparate housing price shocks given the different origin and destination prefectures.

The control variables at the prefecture times birth year level undergo a similar process of weighting. Log wage and log employment at the origin are calculated as weighted averages based on age-specific migration probabilities. For wage and employment at the destinations, the weighted averages are computed in a manner similar to that for housing price gaps, with an initial weighting according to migration patterns followed by a secondary weighting using age-specific migration probabilities.

As the housing prices in China persistently increased over the last decades, as shown

gap for a certain age, such as 20 or 25.

[8]For the main results, I use census data in 2010 to generate the migration patterns weights, which has both a large sample size and detailed information on origin and destination prefectures of migrants. While these weights are time-invariant in the regression, it could be concerning that the weights may have included impacts on housing prices. I have tried weights generated using census data in 2005 and the CMDS data in 2010-2017, restricting migrants to those who migrated before 2000. All these results are very similar to the main results.

in Figure 1, there is a concern that the $FatherEduy \times HPgap$ may predominantly reflect heterogeneous impacts of fathers' education across birth cohorts. To address this concern, I further control the interaction between fathers' education and fixed effects for birth and estimate the following equation:

$$\begin{aligned} Outcome_{ipy} = \beta_1 FatherEduy_{ipy} \times HPgap_{py} + \gamma^y FatherEduy_{ipy} \\ + \beta_3 HPgap_{py} + \Pi Z_{py} + \Omega X_{ipy} + \mu_y + \eta_p + \epsilon_{ipy} \end{aligned} \tag{8}$$

where, compared to Equation (7), the impact of the fathers' education is allowed to vary for children born in different years. This approach isolates the time-varying effects of fathers' education, thereby refining the estimates of the coefficient of interest, β_1.

5.2 Instrumental Variable

Housing prices are highly correlated with local economic conditions. Although various origin and destination economic conditions have been controlled, unobservables can still drive the results. To address the endogeneity issue, an instrumental variable is needed to capture the impacts of housing prices. The Housing Purchase Restriction (HPR) policy is leveraged as a natural experiment to construct the instrumental variable.

Figure 2 shows the housing prices time series between the prefectures that ever implemented the HPR policy and those that never adopted the policy, measured by log yuan per square meter adjusted to 2000 RMB. To make the comparison more transparent, the housing prices of the never-treated prefectures are shifted upward by 0.88 log points. Visually, the pre-trend was quite parallel before the implementation of the HPR policy around 2010. Since the implementation of the policy, the growth of the treated prefectures immediately got slower, while the housing prices in untreated prefectures kept growing. As the treated prefectures gradually relaxed the restriction in 2014-2016, the differences returned to the original level.

Appendix Figure B.2 shows the event study plot for the policy. Using the prefecture-year level panel data, housing prices are regressed on prefecture fixed effects, year fixed effects, and interaction terms between year dummies and an indicator for ever having adopted the policy. The figure shows the estimates and confidence intervals for the coefficients of the interaction terms, which reflect the changes in differences in housing prices between the ever-treated and never-treated prefectures over time. Estimates suggest that the differences remained stable before the policy, became much smaller during the policy, and got back to the original level once the policy was lifted in most of the treated prefectures.

To construct the instrumental variable, I first create a dummy variable at the

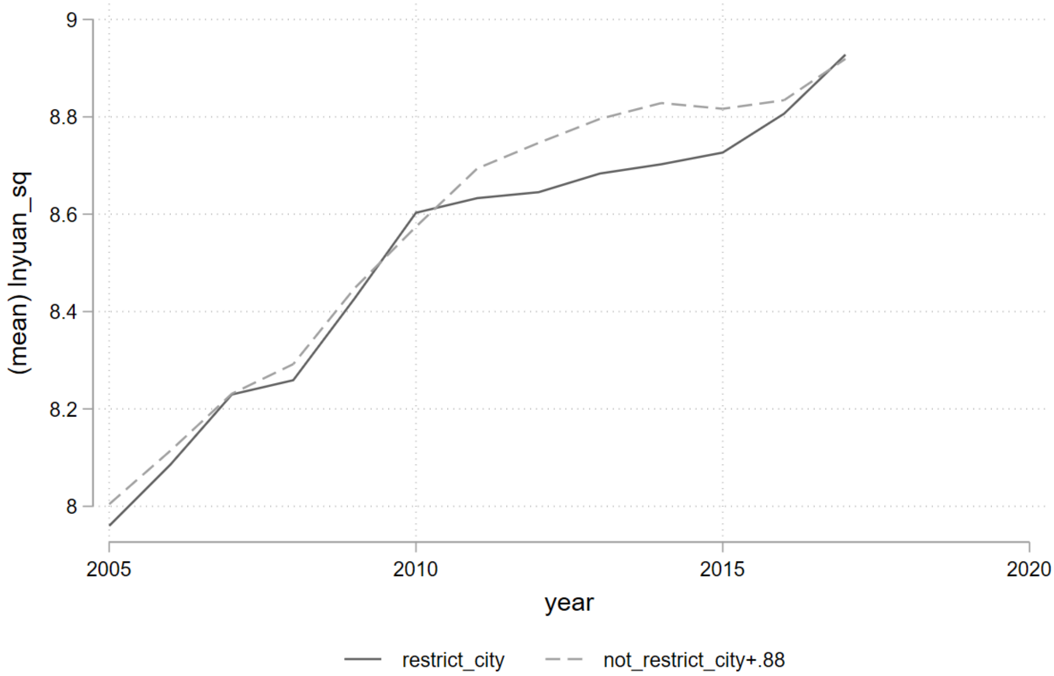


Figure 2: Housing Prices for HPR and non-HPR prefectures

NOTE: The figure shows the housing prices time series between the prefectures that implemented the HPR policy and those that never adopted it, measured by log yuan per square meter adjusted to 2000 RMB. To make the comparison more transparent, the housing prices of the never-treated prefectures are shifted upward by 0.88 log points.

prefecture-year level, indicating whether the HPR policy is in effect. Then, the dummies for the origin and destination prefectures are weighted and summed in the same way as the housing price gaps. The weighted average, HPR_{py}, is then used as the instrumental variable for the housing price gap $HPgap_{py}$.

6 Data

6.1 Housing Prices and Local Economic Conditions

The panel data on housing prices are obtained from the CEIC database, which provides annual selling prices for commercial residential properties in each prefecture. As the individual level data is in 2017, the housing price data used covers the period from 2000 to 2017. Housing price is measured in log yuan per square meter and adjusted to 2000 RMB using the national Consumer Price Index (CPI).

I use housing prices as proxies for housing costs for migrants. Migrants usually do not purchase houses at the destinations due to both *Hukou* restrictions and credit constraints. According to the CMDS dataset, about 10-20% migrants who migrated across prefectures purchased a house at the destination. What matters more for the migration decision is the rental prices at the destination. Unfortunately, data on rental prices are limited in the Chinese context. Data suggests that rental prices and house purchasing prices are highly correlated.[9] Therefore, I use housing prices to proxy the residential costs faced by migrants.

The CEIC database also contains data on local economic conditions, including employment, wage, export, FDI, and fiscal expenditure.

6.2 Individual Characteristics

Individual characteristics data are sourced from the China Household Finance Survey (CHFS), a nationally representative survey project executed by the Research Center for China Household Finance. The CHFS is designed to collect detailed micro-level data on household finance, including comprehensive information on demographic characteristics, employment status, assets and liabilities, income and consumption, social security and insurance, and subjective attitudes. This dataset provides a comprehensive and detailed description of household economic and financial behaviors. Further information about the dataset can be found in Gan et al. (2014).

This study uses data from the 2017 China Household Finance Survey (CHFS), which provides information on respondents' origin and destination prefectures, as well as their migration history. The CHFS collects basic information about each individual's parents, such as education and *Hukou* status, regardless of whether they co-reside with the respondent or are still alive. This feature helps me avoid the selection biases often encountered in studies that rely on surveys focusing only on cohabitating household members. Additionally, the CHFS includes detailed information on income from the informal sector and farming, which helps mitigate selection biases that commonly arise in studies lacking income data for workers in these sectors.

I restrict the sample in the following ways. Firstly, because the migration behavior is potentially different in relatively affluent prefectures compared to the others, I drop observations whose origin is one of the Tier-1 prefectures: Beijing, Shanghai, Shenzhen, and Guangzhou. Second, to ensure the availability of income information, I limit the observations to individuals aged 16-55. The results are not sensitive to different selections of the upper and lower bound of the age range. Appendix Table B.2 shows the descriptive statistics.

[9]The correlation coefficient between log rent and log housing prices is 0.89, indicating a strong correlation between the two stats. The rent data is sourced from WIND.

7 Results

7.1 OLS Results

I estimate Equation (7) using OLS, and the results are shown in Column (1) of Table 2. The coefficient for $HPgap$ is negative, indicating that higher housing costs reduce migration. Since fathers' education is not centered, the magnitude of the coefficient reflects the impact of housing prices on children whose fathers have zero years of schooling. The positive coefficient for $FatherEduy$ suggests that fathers' education can mitigate the adverse effects of housing costs. However, this mitigation is not sufficient to fully counterbalance the challenge posed by housing prices. Overcoming this challenge would require more than 30 years of education, which is beyond what is achievable in this setting.

Table 2: Impacts on Migration Decision

	(1) Migration	(2) Migration	(3) Migration	(4) Migration
FatherEduy	0.005^{***}		0.005^{***}	
	(0.001)		(0.001)	
FatherEduy× HPgap	0.009^{***}	0.009^{***}	0.007^{***}	0.007^{***}
	(0.001)	(0.001)	(0.001)	(0.001)
HPgap	-0.267^{***}	-0.289^{***}	-0.246^{***}	-0.268^{***}
	(0.072)	(0.072)	(0.072)	(0.072)
Eduy× HPgap			0.008^{***}	0.008^{***}
			(0.002)	(0.002)
FatherEduy× birthyearFE	-	Y	-	Y
Obs.	15,128	15,128	15,128	15,128
Adj. R-sq	0.423	0.424	0.424	0.425
Mean(Dep. Var.)	0.201	0.201	0.201	0.201

NOTE: This table shows the OLS results. Compared to the odd columns, the even columns control for the fathers' years of education interacted with birth year fixed effects, so the level term of fathers' education year is absorbed, and the estimated coefficient is not shown in the table. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

To address the concern that the $FatherEduy \times HPgap$ may predominantly reflect heterogeneous impacts of fathers' education across birth cohorts, I estimate Equation (8) where the effect of the fathers' education is allowed to vary for children born in different years, thus isolating the time-varying impact of a father's education. The results are shown in Column (2) of Table 2. The coefficients of interest barely change.

Because fathers' education also affects children's education, one potential explanation for the findings is that fathers' education proxies children's education, and it is children's higher education attainment that is overcoming the housing costs. While such a channel also affects intergenerational mobility, the policy implications differ depending on whether the focus should be on enhancing educational opportunities or improving labor market prospects for workers.

To assess how much of the observed impact is attributable to children's education, I include the interaction term between children's years of education and the housing price gap, $Edu \times HPgap$, in columns (3) and (4). The coefficients of $Edu \times HPgap$ are positive and significant, suggesting that children's education attainment helps overcome the housing costs barrier. The coefficients of $FatherEduy_{ipy} \times HPgap_{py}$ decrease slightly but remain large and significant, suggesting that although the impact of fathers' education on children's education partly explains how fathers' education helps overcome barriers related to housing costs, fathers' education continues to help through channels other than influencing children's education.

Table 3: Impact of housing price on intergenerational mobility

	(1) lnIncome	(2) lnIncome	(3) lnIncome	(4) lnIncome
FatherEduy	0.023***		0.023***	
	(0.003)		(0.003)	
FatherEduy× HPgap	0.029***	0.028***	0.021***	0.020***
	(0.006)	(0.006)	(0.007)	(0.007)
HPgap	-1.068***	-1.031***	-0.982***	-0.943***
	(0.333)	(0.342)	(0.334)	(0.343)
Eduy× HPgap			0.024***	0.024***
			(0.009)	(0.009)
FatherEduy× birthyearFE	-	Y	-	Y
Obs.	12,189	12,189	12,189	12,189
Adj. R-sq	0.352	0.353	0.352	0.353
Mean(Dep. Var.)	9.793	9.793	9.793	9.793

NOTE: This table shows the OLS results. Compared to the odd columns, the even columns control for the father's year education interacted with birth year fixed effects, so the level term of fathers' education year is absorbed, and the estimated coefficient is not shown in the table. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table 3 shows the results estimated in the same way as in Table 2, but with the outcome variable replaced by the log of children's income. The sample sizes are smaller, as only observations with available income information are included. The coefficient of $FatherEduy$ is positive and significant, suggesting that fathers' education is positively correlated with the children's income. The magnitude is close to the findings in the literature. For example, Lee et al. (2024) find that an exogenous one-year increase in parents' schooling increases children's lifetime earnings by 1.2 percent on average. Our estimates are somewhat larger.

The negative coefficients for $HPgap$ suggest that higher housing costs are associated with lower income, which aligns with expectations given that migration often leads to income gains, and increased housing costs can hinder migration. Since $FatherEduy$ is not centered, the magnitude of the $HPgap$ coefficient reflects its impact on children whose fathers have zero years of education. Additionally, the positive and significant coefficients for the interaction term $FatherEduy \times HPgap$ indicate that higher levels of fathers' education can mitigate the adverse effects of housing costs on children's income.

Columns (2) and (4) control for fathers' years of education interacted with birth year fixed effects, leading to little change in the estimates. In Columns (3) and (4), after controlling for $Edu \times HPgap$, the coefficients of interest decrease slightly but remain large, positive, and statistically significant.

7.2 IV Results

Unobserved factors, such as job opportunities, might confound the relationship between housing prices and migration. As a booming economy attracts migrants while higher housing costs deter them, the coefficient of $HPgap$ is likely underestimated in the OLS results. The positive coefficient for $FatherEduy \times HPgap$ could indicate that fathers with higher education levels are more effective at using their social networks to secure job opportunities for their children in destination prefectures during economic booms. As a result, these coefficients may reflect fathers' varying abilities to help their children access employment rather than the differential impacts of housing prices.

Table 4 shows the results using the instrumental variable approach. Row "K-P F-stat" shows the first-stage *F-stats*, which are larger than 10, the rule of thumb, suggesting strong first stages. The "LM test" row shows the statistic testing whether the matrix of coefficients from the first-stage regressions has full column rank. Large values of the test statistic indicate that the instruments provide sufficient independent variation to account for the endogenous variables. The results of the first stages are shown in Appendix Table B.3. The impacts of the HPR policy on housing costs are negative, as expected. The policy reduces housing demand, which in turn slows the growth of housing prices.

Table 4: IV Results

	(1) Migration	(2) Migration	(3) lnIncome	(4) lnIncome
FatherEduy× HPgap	0.012***	0.009***	0.024***	0.030***
	(0.002)	(0.002)	(0.009)	(0.010)
HPgap	-0.923***	-0.802**	-3.388**	-3.763***
	(0.354)	(0.345)	(1.409)	(1.440)
Eduy× HPgap		0.009***		-0.018
		(0.003)		(0.014)
FatherEduy× birthyearFE	Y	Y	Y	Y
K-P F-stat	30.203	20.600	58.633	39.776
LM test	55.953	57.091	110.914	112.987
Obs.	15,128	15,128	12,189	12,189
Mean(Dep. Var.)	0.201	0.201	9.793	9.793

NOTE: This table shows the main results using the HPR policy treatment as the instrumental variable. Fathers' years of education interacted with birth year fixed effects are controlled for in all columns. In Columns (1) and (3), the endogenous variables are $HPgap$ and $FatherEduy \times HPgap$; the instrumental variables are HPR and $FatherEduy \times HPR$. In Columns (2) and (4), the endogenous variables are $HPgap$, $FatherEduy \times HPgap$, and $Edu \times HPgap$; the instrumental variables are HPR, $FatherEduy \times HPR$, and $Edu \times HPR$. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. The K-P F-stat row shows the F-stats of the first stages. The LM test row shows the statistics for the underidentification test. A rejection of the null indicates that the matrix of first-stage coefficients has full column rank, i.e., the model is identified. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

In Columns (1) and (2), the negative coefficient for $HPgap$ suggests that larger housing price gaps reduce migration, while the positive coefficient for the interaction term indicates that fathers' education can help mitigate the deterring effect of higher housing costs. In Columns (3) and (4), the negative coefficient for $HPgap$ suggests that higher housing price gaps reduce income, although these coefficients are not statistically significant. The positive coefficient for the interaction term again indicates that fathers' education can mitigate the adverse effects of housing costs.

To interpret the magnitude of the coefficients, I first need to understand the scale of $HPgap$. $HPgap$ is a weighted average of the difference between the log housing prices at potential destinations and the log housing price at the origin, measured in yuan per square meter. The mean value of $HPgap$ is 0.55, indicating that housing prices at the destinations are approximately 173% of those at the origin. I then calculate the impact of a 0.1 change in $HPgap$. Since $HPgap$ is demeaned, the magnitude of the coefficient reflects intergenerational persistence at the average housing price gap. The coefficient for $FatherEduy \times HPgap$ in Column (4) suggests that a 0.1 increase in $HPgap$ could raise intergenerational persistence by 13%.[10] For comparison, Feigenbaum (2015) finds that a downturn during the Great Depression that is one standard deviation more severe increases intergenerational persistence by 39%.

As the data is cross-sectional, children's income is from the year 2017. Income data from a single year might be subject to transitory shocks or measurement errors. However, this measurement error is not a big concern here, as income is the dependent variable, not the independent variable. While measurement error in the dependent variable may increase noise, it does not introduce bias. The independent variable, fathers' education, is stable and less susceptible to transitory shocks or measurement errors. To further address this concern, I follow Nybom and Stuhler (2017) and average income across the 2015, 2017, and 2019 waves of the CHFS, provided income data are available in these waves.

Appendix Table B.7 presents the results in Columns (3) and (4), where children's income is calculated as a multi-year average. For comparison, Columns (1) and (2) of the same table reproduce Columns (3) and (4) from the main results in Table 4. The estimated coefficients using single-year and multi-year income calculations are very similar.

7.3 Robustness

7.3.1 Additional Controls

Although I have controlled for employment and wages at the origin and destination, there may still be concerns that some unobservable economic conditions can potentially confound the results. To address this concern, I include origin × birth year fixed effects. The housing price gap is also at the origin times birth year level, where the potential destinations and the corresponding weights are specific to each origin. Therefore, the origin × birth year fixed effects would capture the time-variant characteristics in the origin and the origin-specific weighted destination shocks. The results are shown in Table 5 Columns (1) and (2). $HPgap$ is at the origin × birth year level, so they are absorbed by the fixed effects. The estimated coefficients of interest remain close to the main results.

Recent econometric literature (Feigenberg et al., 2023) has pointed out that, for interacted models, omitted variable bias may exist even when the level terms of confounders

[10] Although not shown in the table, I ran regressions without controlling for the interaction between fathers' education and children's birth year fixed effects to estimate the coefficient for $FatherEduy$, which is approximately 0.022. The calculation for the change in intergenerational persistence due to a 0.1 change in $HPgap$ is $0.03 \times 0.1 / 0.023 = 13\%$.

Table 5: Robustness: Additional Controls

	(1) Migration	(2) Migration	(3) Migration	(4) Migration
FatherEduy× HPgap	0.010***	0.009***	0.009***	0.006**
	(0.002)	(0.002)	(0.003)	(0.003)
HPgap			-0.801**	-0.702**
			(0.336)	(0.329)
Eduy× HPgap		0.004		0.009***
		(0.004)		(0.003)
OriginFE× BirthyearFE	Y	Y	-	-
Interaction terms	-	-	Y	Y
Obs.	15,128	15,128	15,128	15,128
Mean(Dep. Var.)	0.201	0.201	0.201	0.201

NOTE: This table shows the results of adding more controls using the HPR policy treatment as the instrumental variable. Fathers' years of education interacted with birth year fixed effects are controlled for in all columns. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level.

are controlled for. For this paper, one concern could be that the housing price gap is capturing the economic development at the destination. The impact of the economic development at the destination can be heterogeneous across children with different parental backgrounds. For example, children with more affluent fathers may respond more to economic growth at potential destinations due to information advantage. This concern can be addressed by controlling for the interaction term between fathers' education and economic conditions at the destination. In Table 5 Columns (3) and (4), I control for the interaction term between fathers' education and the employment, wage, fiscal expenditure, export, FDI at the destination prefectures, and the level terms of these measures of economic conditions. The results are close to the main results in Table 4

7.3.2 Alternative Measure of Parental Background

Because our dataset includes children from various prefectures and birth years, the same level of fathers' educational attainment may reflect different parental backgrounds. For instance, a father with a high school education in a remote area may have a very different socioeconomic status compared to a father with the same education level in Beijing. To address this concern, I replace the original measure with residualized fathers' years of education, obtained by regressing fathers' education on their *Hukou* status and origin prefecture-by-child birth year fixed effects.

The results are shown in Appendix Table B.8. The estimates are similar to the main results in Table 4, suggesting that the results are not sensitive to the measurement of fathers' education.

While education is an important indicator of parental background, it does not capture all aspects. Fathers with the same educational attainment can still have very different incomes and levels of household wealth. To address this concern, I replace fathers' education with imputed fathers' income and household wealth in the following robustness tests.

To address the lifecycle bias and transitory shocks in income, researchers commonly employ proxies such as parental education or occupation to proxy parental lifetime income (Solon, 1992; Gong et al., 2012). In line with this methodology, I utilize parental education, *Hukou* status, Communist party membership, job position, and coastal region indicators to impute parental lifetime income. These variables are relatively stable throughout the life cycle and strongly correlate with lifetime earnings, thereby reducing lifecycle bias. Moreover, the use of estimated income, rather than actual income, helps to decrease the attenuation bias caused by temporary income fluctuations (Fan et al., 2021).

The imputation takes the following steps. First, I use CHFS 2015 data, restrict the sample to males, and estimate the following equation:

$$ln(Income)_i = \alpha_0 + \alpha_1 Age_i + \alpha_2 Age\_squared_i + X_i' \alpha_X + \epsilon_i \tag{9}$$

X is a set of demographic and socioeconomic variables, including fixed effects for education, *Hukou* type, Communist party membership, job position, and birth cohort. There is also a coastal dummy, which accounts for regional differences.[11] Then, based on Equation (9), I calculate the predicted income for our main sample by applying the estimated coefficients to the available information on their fathers' characteristics.[12]

[11] Communist party membership is an indicator of whether the father is a member of the Communist Party. Job position refers to administrative levels including 1) ordinary worker, 2) department manager, 3) general manager, 4) (deputy) team leader/section chief, 5) (deputy) division head, 6) (deputy) director, 7) (deputy) bureau head and above, 8) village cadre, 9) township cadre, etc. The coastal dummy equals one if the residential province was a coastal province when the child was age 14. For the birth cohort, I separate fathers into 10-year age groups based on their birth years. When the father's birth year is not available, I use the average birth year of fathers given the child's birth year, which I generated using the 2013 China Health and Retirement Longitudinal Study (CHARLS) data, a nationally representative dataset focusing on the old population.

[12] I do not use the coefficients of age, age squared, or birth cohort because I am using a cross-sectional dataset to impute fathers' income and want to avoid income variation caused by age differences. Instead of excluding age-related variables, an alternative approach involves constructing the imputed income for a worker at a representative age. This alternative approach involves calculating the imputed income under the assumption that all individuals are at the mean age, effectively adding a constant to all observations.

Table 6: Robustness: Impute Fathers' Income

	(1) Migration	(2) Migration	(3) lnIncome	(4) lnIncome
lnFatherIncome× HPgap	0.188***	0.138***	0.404***	0.497***
	(0.034)	(0.038)	(0.148)	(0.163)
HPgap	-0.967***	-0.837**	-3.293**	-3.647**
	(0.357)	(0.347)	(1.397)	(1.432)
Eduy× HPgap		0.009***		-0.017
		(0.003)		(0.014)
lnFatherIncome× birthyearFE	Y	Y	Y	Y
K-P F-stat	30.454	20.781	60.250	40.795
LM test	56.034	57.198	112.350	114.492
Obs.	15,107	15,107	12,172	12,172
Mean(Dep. Var.)	0.200	0.200	9.793	9.793

NOTE: This table shows the results using imputed fathers' income in place of fathers' education. Imputed fathers' incomes interacted with birth year fixed effects are controlled in all columns. In Columns (1) and (3), the endogenous variables are $HPgap$ and $FatherIncome \times HPgap$; the instrumental variables are HPR and $FatherIncome \times HPR$. In Columns (2) and (4), the endogenous variables are $HPgap$, $FatherIncome \times HPgap$, and $Edu \times HPgap$; the instrumental variables are HPR, $FatherIncome \times HPR$, and $Edu \times HPR$. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. The K-P F-stat row shows the F-stats of the first stages. The LM test row shows the statistics for the underidentification test. A rejection of the null indicates that the matrix of first-stage coefficients has full column rank, i.e., the model is identified. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

I re-estimate the main results using imputed fathers' income in place of fathers' education. Table 6 presents the results. The F-statistics of the first stages remain above 10, indicating strong first-stage regressions. The coefficients for $HPgap$ are negative across all columns, suggesting that higher housing costs act as a barrier to migration and reduce income. The coefficients for the interaction term, $FatherIncome \times HPgap$, are positive and significant in all columns, indicating that fathers' income can help mitigate the negative impacts of housing costs.

To evaluate the goodness of the imputation, I compare our estimates of the intergenerational income elasticity (IGE) to the existing literature. I show results in Appendix Table B.9 where the $FatherIncome \times birthyear$ fixed effects are dropped so that I observe the estimated coefficient of fathers' log income, i.e., the estimates of IGE. Our IGE estimate is approximately 0.384, closely aligning with Fan et al. (2021), who report an

The results will be the same, as the imputed income of fathers is centered (demeaned) before estimating the regression models. I also check the results incorporating age-related variables when imputing income, and they remain robust.

IGE of 0.390 for the 1970-1980 birth cohort and 0.442 for the 1981-1988 birth cohort.[13]

A similar procedure is applied to impute household assets. Appendix Table B.10 shows the results where fathers' education is replaced with imputed household assets. The results remain robust.

7.3.3 Additional Robustness Tests

Characteristics of the origin prefecture may affect people's migration response to housing price changes. For example, the migration behaviors of individuals from economically more developed origin prefectures or popular destination prefectures may differ from those of others. To address this concern, I exclude from the sample the top ten prefectures by GDP per capita or the top ten most popular destinations and rerun the main regressions.[14] Additionally, people born in prefectures adjacent to megacities such as Beijing or Shanghai may choose to live in origin and commute rather than rent or purchase houses in the megacities. Another concern is that housing prices may have geographical spillovers, which could affect migrations through an income effect. To check whether these concerns matter for our result, I run the main regressions excluding the prefectures that are adjacent to the Tier-1 prefectures: Beijing, Shanghai, Shenzhen, or Guangzhou.[15] Appendix Table B.4 shows the results restricting the origin prefectures as described above. The results are close to the main findings.

The housing price data used in the main results comes from the CEIC database. To verify that the observed patterns are not specific to CEIC, I also use data from Anjuke, a leading online real estate platform in China which is known for its extensive listings and comprehensive services for property buying, selling, and renting. The results are shown in Appendix Table B.6. The estimates are close to the main results using the CEIC data, suggesting that I am not capturing dataset-specific patterns.

Another potential concern is that affluent destinations are more likely to adopt the HPR policy, and the varying effects of parental background may differ for origin prefectures with higher migration rates to affluent areas. Consequently, the estimated coefficient of interest might capture the heterogeneous impact of parental background

[13]The measurement and sample period in our study differ slightly from Fan et al. (2021). They use imputed parents' income rather than fathers' income and focus on individuals born between 1970-1988, while our sample includes individuals born between 1962-2000.

[14]The top ten prefectures in terms of GDP per capita in 2000 were Shenzhen, Shanghai, Zhuhai, Wuxi, Suzhou, Guangzhou, Beijing, Xiamen, Dongying, and Hangzhou. Aside from megacities like Beijing and Shanghai, the top ten prefectures vary across different years. Using the top ten prefectures from other years yields very similar results. According to 2010 census sample data, the top ten most popular destinations are Beijing, Shanghai, Shenzhen, Guangzhou, Dongguan, Hangzhou, Suzhou, Chengdu, Ningbo, and Foshan.

[15]The adjacent prefectures are Baoding, Shaoguan, Suzhou, Zhongshan, Tianjin, Dongguan, Zhangjiakou, Chengde, Jiaxing, Qingyuan, Langfang, Foshan, Huizhou.

driven by specific origin characteristics that are correlated with higher migration rates to affluent destinations. To address this concern, I construct indicators for whether a prefecture is among the 46 that have ever adopted the HPR policy. These indicators for potential destinations are then weighted by the migration patterns and interacted with fathers' education. This measure captures characteristics of the origin that are correlated with higher migration rates to destinations that have adopted the policy. The results remain robust controling for this interaction term as shown in Appendix Table B.5.

7.4 Heterogeneous Analysis

7.4.1 Fathers' Education Levels

This section investigates heterogeneous impacts across different demographic groups. First, I examine which level of father's education matters most for overcoming housing price barriers.

Table 7: Heterogeneity: Fathers' Education Levels

	OLS		IV	
	(1)	(2)	(3)	(4)
	Migration	lnIncome	Migration	lnIncome
$1\{\text{FatherEduLevel} \geq 2\} \times$ HPgap	0.046***	0.142**	0.072***	0.208*
	(0.011)	(0.070)	(0.020)	(0.115)
$1\{\text{FatherEduLevel} \geq 3\} \times$ HPgap	0.029*	0.095	0.028	0.144
	(0.015)	(0.070)	(0.021)	(0.090)
$1\{\text{FatherEduLevel} \geq 4\} \times$ HPgap	-0.002	-0.023	0.004	-0.006
	(0.019)	(0.077)	(0.026)	(0.100)
HPgap	-0.263***	-0.934***	-0.705**	-3.461**
	(0.072)	(0.347)	(0.357)	(1.471)
K-P F-stat			11.598	22.778
LM test			54.348	111.976
Obs.	15,128	12,189	15,128	12,189
Mean(Dep. Var.)	0.200	9.793	0.200	9.793

NOTE: This table shows the IV results for heterogeneity across children with parents of different education levels. The interaction term between children's education and the housing price gap is included in all columns. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table 7 shows the results interacting different levels of fathers' education attainment with the housing price gaps. Fathers' educational attainment, $FatherEduLevel$, is a category of four levels: 1) no schooling (32.36% of the sample), 2) primary school (32.91%), 3) junior high school (19.21%), and 4) senior high or above (15.52%). In the regression, "no schooling" is taken as the reference group. To assess the incremental impact of each education level, I use indicators for reaching or exceeding a particular educational threshold. For instance, the coefficient of the interaction term $1\{FatherEduLevel \geq 3\} \times HPgap$ reflects the additional impact of having a junior high school education, compared to only primary school education, in overcoming housing barriers. A significant and positive coefficient indicates that higher educational attainment provides an advantage, while an insignificant coefficient suggests that further education beyond the previous level does not significantly help in overcoming the housing price barrier.

Columns (1) and (3) of Table 7 show the impacts on migration. The results reveal that, compared to fathers with no schooling, having a primary school education significantly aids in overcoming housing barriers, as evidenced by the positive and statistically significant coefficients. The magnitudes of the coefficients of having a junior high school education are negligible, although noisily estimated. Further education beyond junior high school, such as senior high school or higher, does not yield statistically significant additional benefits in mitigating the challenges posed by housing expenses.

Columns (2) and (4) display the impacts on income, which are consistent with the migration findings. Fathers with a primary school education provide the most substantial advantage in mitigating the adverse effects of housing costs on their children's income, as shown by the positive and significant coefficients. The magnitudes of the coefficients of having a junior high school education are negligible, although noisily estimated. In contrast, Further educational attainment beyond junior high school does not result in significant additional income benefits, suggesting that the critical threshold for overcoming the adverse effects of housing prices on income lies at the basic education level.

The findings suggest that fathers with low education attainment, who likely have limited access to credit, are much less able to help their children overcome housing barriers, thereby limiting the children's opportunities for migration and economic advancement. These results align with the broader literature on poverty traps and credit constraints, which emphasize how such constraints severely restrict people's ability to migrate, even within their own country. Although migration is often viewed as a pathway out of poverty, it remains inaccessible for many low-income households due to the associated costs and uncertainties. These findings highlight the critical need for policies that provide affordable housing, especially for those at the bottom of the income distribution, as these measures can significantly enhance economic mobility and reduce intergenerational poverty.

7.4.2 Rural vs Urban

I also examine heterogeneity across children with parents of different *Hukou* types. I did not use children's *Hukou* type directly, as it is correlated with their migration behavior. Columns (1) and (2) in Table 8 show the results for children with at least one rural *Hukou* parent, while Columns (3) and (4) show the results for children with urban *Hukou* parents. The findings suggest that the impact of parental background in overcoming housing cost barriers is particularly significant for children with rural *Hukou* parents. The coefficients for the interaction term $FatherEduy \times HPgap$ are positive and statistically significant in both the migration and income models for rural *Hukou* families, indicating that higher parental education helps these children overcome housing barriers, facilitating migration and improving income outcomes.

In contrast, for children with urban *Hukou* parents, the coefficients are much smaller and not statistically significant, suggesting that parental education has a negligible effect on overcoming housing cost barriers in urban contexts. The differences in the coefficients between rural and urban *Hukou* groups are statistically significant, supporting the notion that housing cost barriers impact these two groups differently.

These findings highlight the distinct challenges rural families face in navigating the housing market and underscore the critical role of parental background in mitigating these barriers for children from rural areas. Given that most rural workers migrate from rural to urban areas, the results suggest that rural-urban migration is the most affected by housing cost barriers compared to other types of migration.

7.5 Impacts on Structural Transformation

In the heterogeneous analyses, I find that the impacts of housing costs are more pronounced for individuals with rural *Hukou* parents. Given that migration is often linked to non-agricultural employment, housing prices may have varied effects on agricultural employment depending on parental backgrounds. This variation highlights a complex interaction between housing costs, structural transformation, and intergenerational mobility. To explore this interaction further, I examine the impacts of housing costs on agricultural employment.

Table 9 shows the results of this analysis, where the outcome variable is whether the individual works in the agricultural sector. Columns (1) and (2) display the results

Table 8: Heterogeneity: Rural vs Urban

	Rural		Urban	
	(1)	(2)	(3)	(4)
	Migration	lnIncome	Migration	lnIncome
FatherEduy× HPgap	0.013***	0.046**	0.005	-0.012
	(0.003)	(0.014)	(0.004)	(0.013)
K-P F-stat	16.785	36.129	13.365	16.073
LM test	48.475	102.421	39.327	49.862
Obs.	10,205	8,400	4,923	3,789
Mean(Dep. Var.)	0.208	9.511	0.179	10.417

NOTE: This table shows the IV results for heterogeneity across children with parents of different *Hukou* types. The interaction term between children's education and the housing price gap is included in all columns. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

for the full sample, while Columns (3) and (4) focus on children with at least one rural *Hukou* parent. The findings suggest that higher housing costs increase the likelihood of working in the agricultural sector. This effect is particularly evident for those with rural *Hukou* backgrounds, as shown by the significant positive coefficients for $HPgap$.

Moreover, the interaction between fathers' education and housing costs reveals that higher levels of fathers' education mitigate the impact of housing costs on the probability of working in agriculture. The negative and significant coefficients for the interaction term $FatherEduy \times HPgap$ indicate that as fathers' education increases, the influence of housing costs on agricultural employment decreases. This suggests that families with more educated fathers are better equipped to navigate the challenges posed by housing costs, enabling their children to pursue non-agricultural employment opportunities.

These heterogeneous effects of housing costs on agricultural employment across different parental backgrounds underscore the role of housing costs in shaping structural transformation within the economy. The findings imply that housing costs can reinforce existing inequalities and contribute to the persistence of intergenerational poverty by limiting economic mobility and trapping individuals in agricultural work. This interaction between housing costs, migration, and agricultural employment highlights the need for policies that address housing affordability, particularly for rural populations, to foster more inclusive economic growth and reduce intergenerational inequality.

Table 9: Impacts on Agricultural Employment

	Rural Sample	
	(1)	(2)
FatherEduy× HPgap	-0.008**	-0.010***
	(0.003)	(0.004)
HPgap	1.456**	1.554**
	(0.695)	(0.711)
Eduy× HPgap	-	Y
FatherEduy× birthyearFE	Y	Y
K-P F-stat	24.926	16.619
LM test	49.059	49.036
Obs.	8,223	8,223
Mean(Dep. Var.)	0.294	0.294

NOTE: This table shows the results using agricultural employment as the outcome. Fathers' years of education interacted with birth year fixed effects are controlled for in all columns. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

8 Spatial Equilibrium Model

In the reduced-form analyses, I aggregate the impacts of housing prices at potential destinations into a weighted average, treating migration as a binary decision—whether to migrate or not. However, in reality, migration is not a binary choice but a selection from multiple potential destinations, where the housing price at one destination can influence migration flows to other destinations. A discrete choice model allows me to capture this complexity by modeling migration decisions as a set of discrete choices.

Moreover, while the current analyses focus on the household side and do not discuss the impacts of migration on the destinations, migration certainly has general equilibrium effects. For example, when housing prices decrease, migration increases; housing prices at the destination will then be endogenously driven up by the inflow of migrants, preventing further migration.

Through counterfactual analyses, the spatial equilibrium model can be used to simulate the outcomes of various policy interventions, such as housing subsidies, changes in land use regulations, or increases in housing supplies in high-demand areas for migrants. Additionally, it enables me to evaluate how the policy impacts would differ depending on which destination prefectures and which subset of populations they target.

By comparing the simulated outcomes with the current equilibrium, the model can provide valuable insights into the most effective strategies for promoting greater economic mobility through enhanced housing affordability. Moreover, I would be able to assess the extent to which intergenerational persistence, geographical inequality, and barriers to structural transformation may be reduced with various housing policies.

In this section, I develop a spatial equilibrium model to generate counterfactuals. The model follows a standard framework, with one key tweak: it incorporates a migration cost related to housing prices, with impacts that vary based on parental backgrounds.

8.1 Setup

8.1.1 Labor Supply / Worker's Choice of Destination and Housing Demand

Each worker chooses to live in the prefecture that offers him the most desirable bundle of wages, housing prices, and moving costs. A worker living in prefecture p inelastically supplies one unit of labor and earns a wage of W_p. Living in a prefecture, the worker consumes housing M, which has a local price of rH_p, where r is the discount factor that transfers housing purchase prices to rental prices, and a national good O, which has a national price that I normalize to one. The worker has Cobb-Douglas preferences for housing and national good, which they maximize subject to their budget constraint:

$$\max_{M,O} \quad \ln\left(M^{\zeta}\right) + \ln\left(O^{1-\zeta}\right)$$

$$\text{s.t.} \quad O + rH_pM \leq W_p.$$

Workers' relative taste for national versus local goods is governed by ζ, where $0 \leq \zeta \leq 1$. ζ is assumed to be constant across workers. The worker's optimized utility function from consumption can be expressed as an indirect utility function for living in prefecture p. If the worker were to live in prefecture p, his indirect utility from consumption, V_p, would be:

$$V_p = ln(W_p) - \zeta ln(rH_p) = w_p - \zeta h_p - \zeta lnr$$

where $w_p = ln(W_p)$ and $h_p = ln(H_p)$. The worker's optimized utility function also leads to his housing demand (D_p):

$$D_p = \frac{\zeta W_p}{rH_p} \tag{10}$$

Additionally, worker i from origin prefecture o pays financial and psychological moving costs C_{iop} when relocating to destination prefecture p:

$$C_{iop} = \beta_f^h h_p + \beta^d lnDistance_{op} + \beta^r ruralHukou_i \times \mathbf{1}\{o \neq p\}$$
$$+ \beta^m male_i \times \mathbf{1}\{o \neq p\} + \pi_f + \xi_p$$

where $\beta_f^h h_p$ captures the transient residential cost that occurs when the worker searches for a job once they arrive at the destination. The impacts of h_p differentiate across parental backgrounds, f, as workers from more affluent families are more capable of paying the upfront residential cost. $\beta^d lnDistance_{op}$ captures the transportation cost or the utility cost of being far away from home. β^r measures the deferential cost for workers from rural areas if they migrate. Rural workers may find migrating more beneficial than their urban counterparts, as they often access significantly improved job opportunities and amenities compared to what is available in their rural hometowns. β^r measures the deferential cost between male and female workers. Additionally, π_f are fixed effects capturing taste differences for workers from different parental backgrounds. ξ_p is the residual, capturing destination-specific unobserved characteristics such as local amenities. Housing prices can be correlated with local amenities and are, therefore, endogenous. I instrument housing prices using the HPR policy as in the main analyses.

Each worker has an idiosyncratic taste for each prefecture, which is measured by ϵ_{iop}. ϵ_{iop} is drawn from a Type I Extreme Value distribution. To simplify notation and discussion of estimation, I renormalize the utility function by dividing each worker's utility by the standard deviation of ϵ_{iop} so that ϵ_{iop} is drawn from a standard Type I Extreme Value distribution. With a slight abuse of notation, I redefine the parameters of the re-normalized utility function using the same notation of the utility function measured in wage units. β^w denotes the coefficient for utility from wages. Combined, the utility a worker obtains by choosing to live in prefecture p is:[16]

$$U_{iop} = \beta^w(w_p - \zeta h_p) + \beta_{fedu}^h h_p + \beta^d lnDistance_{op} + \beta^r ruralHukou_i \times \mathbf{1}\{o \neq p\}$$
$$+ \beta^m male_i \times \mathbf{1}\{o \neq p\} + \pi_f + \xi_p + \epsilon_{iop} \tag{11}$$

I rewrite this equation as

$$U_{iop} = \delta_p^f + \chi_{iop} + \epsilon_{iop}$$

where $\delta_p^f = (w_p - \zeta h_p) + \beta_f^h h_p + \pi_f + \xi_p$ and $\chi_{iop} = \beta^d lnDistance_{op} + \beta^r ruralHukou_i \times \mathbf{1}\{o \neq p\} + \beta^m male_i \times \mathbf{1}\{o \neq p\}$.

[16] $-\zeta lnr$ is dropped as it is a constant.

$lnDistance_{op}$ and $ruralHukou_i$ are taken as exogenous while h_p is endogenous. w_p is determined by exogenous productivity, as detailed below. This setup is the conditional logit model, first formulated in this utility maximization context by McFadden (1973). Aggregate population differences of workers across prefectures represent differences in these workers' mean utility values for these prefectures. The total expected population of prefecture p is the probability that each worker lives in the prefecture, summed over all workers:

$$L_p = \sum_{i \in L} \frac{exp(\delta_p^f + \chi_{iop})}{\sum_{n \in N} exp(\delta_n^f + \chi_{ion})}$$

where L is the set of workers in the nation.

8.1.2 Labor Demand

I assume that within each prefecture, firms produce a homogenous tradeable good and share identical production technology with constant returns to scale production functions. Therefore, firm-level labor demand translates directly to prefecture-level aggregate labor demand. For simplicity of notation, I assume that each prefecture has a representative firm that maximizes profits. Let the price of output be normalized to one. In each prefecture p, the representative firm takes productivity A_p as given and produces goods according to the production function:

$$Y_p = A_p L_p$$

L stands for labor in the prefecture. Since there are a large number of firms and no barriers to entry, the labor market is perfectly competitive, and firms hire such that wages equal the marginal product of labor. Namely:

$$w_p = lnA_p$$

8.1.3 Housing Supply

Following Diamond (2016), housing prices, h_p, are determined by the equilibrium in the housing market. The production of housing relies on inputs such as construction materials and land. Developers act as price-takers and sell homogenous houses at the marginal cost of production.

$$h_p = MC(\kappa_p, \eta_p)$$

The function $MC(\kappa_p, \eta_p)$ represents the relationship between local construction costs, κ_p, and local land costs, η_p, and the marginal cost of constructing a home. The cost of land η_p is a function of the aggregate demand for housing. I parameterize the log housing supply equation as follows:

$$h_p = ln(\kappa_p) + \gamma ln(D_p) \tag{12}$$

The elasticity of housing prices γ with respect to housing demand will be calibrated based on the existing literature. $ln(\kappa_p)$ is unobserved and is included in the residuals.

8.1.4 Equilibrium

Equilibrium in this model is defined by a menu of wages and housing prices (w_{pt}, h_{pt}) with populations (L_{pt}), such that:

- The labor demand equals labor supply:

$$L_p = \sum_{i \in L} \frac{exp(\delta_p^f + \chi_{iop})}{\sum_{n \in N} exp(\delta_n^f + \chi_{iop})} \tag{13}$$

  where $\delta_p^f = (w_p - \zeta h_p) + \beta_f^h h_p + \pi_f + \xi_p$ and $\chi_{iop} = \beta^d lnDistance_{op} + \beta^r ruralHukou_i \times \mathbf{1}\{o \neq p\} + \beta^m male_i \times \mathbf{1}\{o \neq p\}$ and $w_p = lnA_p$.

- Housing demand equals housing supply:

$$h_p = \Gamma lnL_p + \Gamma w_p + \varepsilon_p \tag{14}$$

  where $\Gamma = \frac{\gamma}{1+\gamma}$ and $\varepsilon_p = \Gamma(ln\zeta - lnr) + (1 - \Gamma)ln\kappa_p$

The model does not allow me to solve for equilibrium wages and housing prices analytically, but this setup is useful in estimation.

8.2 Estimations

For this simple setup, I only need to estimate the labor supply side. The key parameter Γ is calibrated based on the existing literature.

In practice, I have a panel of data on housing prices in different prefectures across the years. This allows me to use not only cross-sectional variation but also intertemporal variation for estimation. Specifically, the utility a worker born in year t obtains by choosing to live in prefecture p is:[17]

$$\begin{aligned} U_{iopt} &= \beta^w w_{pt} + \beta^h_f h_{pt} + \beta^d lnDistance_{op} + \beta^r ruralHukou_i \times \mathbf{1}\{o \neq p\} \\ &+ \beta^m male_i \times \mathbf{1}\{o \neq p\} + \pi_{f,t} + \mu_p + \xi_{fpt} + \epsilon_{iopt} \end{aligned} \tag{15}$$

where I add time-invariant prefecture fixed effects μ_p.

The magnitudes of the coefficient on wages and housing prices represent the elasticity of workers' demand for a small prefecture with respect to its local wages and housing prices, respectively.[18]

The probability that worker i choose to migrate to prefecture p is:

$$Pr_{iopt} = \frac{\exp\left(\delta_{fpt} + \chi_{iop}\right)}{\sum_{n \in N} \exp\left(\delta_{fnt} + \chi_{ion}\right)} \tag{16}$$

where N is the set of all destination prefectures.

I estimate the model in two steps. First, I estimate Equation (16) using MLE. Secondly, I use the estimated $\hat{\delta}_{fpt}$ and apply a 2SLS estimation utilizing the implementation of HPR policy as the instrument for housing prices h_{pt} to estimate β^h_f.

In practice, there is not enough data to estimate δ_{fpt} for each birth year and each level of fathers' education. Instead, I divide the sample into early and late birth cohorts based on birth years and categorize fathers' education into high and low. This approach allows me to estimate four δ's for each prefecture. I then use these to estimate the following equation with a 2SLS approach:

$$\begin{aligned} \hat{\delta}_{fpt} &= \beta_1 FatherEduHigh \times HousingPrices_{pt} + \beta_2 HousingPrices_{pt} \\ &+ M'_{pt}\Theta + \mu_p + \pi_{f,t} + \xi_{fpt} \end{aligned} \tag{17}$$

where $FatherEduHigh$ is an indicator for fathers' education being high type. M_{pt} are macroeconomic conditions, which consist of employment rate, GDP per capita, and log wage. μ_p are prefecture fixed effects, and $\pi_{f,t}$ are fathers' education times birth cohort fixed effects. ξ_{fpt} is the residual. Housing price is instrumented using the HPR policy, and the interaction term $FatherEduHigh \times HousingPrices$ is instrumented by

[17] I abuse notation slightly by combining $-\beta^w \zeta h_{pt}$ and $\beta^h_f h_{pt}$ to $\beta^h_f h_{pt}$ as they are not identified separately in estimation.

[18] Due to the functional form assumption for the distribution of workers' idiosyncratic tastes for prefectures, the elasticity of demand of workers with father's education f and birth cohort t for a prefecture p with respect to local housing prices, for example, is $(1 - s_{pt})\beta^h_f$. s_{pt} is the share of all workers of birth cohort t, living in prefecture p. For a small city, where the share of all workers of birth cohort t living in prefecture p is close to zero, the demand elasticity for rent is simply β^h_f.

$FatherEduHigh \times HPR$.

The results of the second step are shown in Table 10. Columns (1) and (2) show the OLS estimates, both without and with macroeconomic controls. Columns (3) and (4) show the IV estimates, also with and without these controls. The coefficients for housing prices are negative and significant, indicating that higher housing prices in a prefecture reduce a worker's utility of residing there. The positive and significant coefficients for the interaction term between fathers' education and housing prices suggest that higher levels of fathers' education can partially offset the adverse effects of housing prices. However, this mitigation is not sufficient to fully counterbalance the adverse impacts of high housing costs.

Table 10: Conditional logit

	OLS		IV	
	(1)	(2)	(3)	(4)
FatherEduHigh× HousingPrices	0.380***	0.384***	0.747**	0.758**
	(0.107)	(0.110)	(0.322)	(0.326)
HousingPrices	-1.181***	-1.011***	-4.050***	-3.606**
	(0.347)	(0.376)	(1.230)	(1.436)
FatherEduHigh× birthcohort FE	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y
Controls	-	Y	-	Y
K-P F-stat			12.880	7.772
LM test			19.100	9.569
Obs.	607	603	607	603

NOTE: Standard errors are clustered at the prefecture level. Columns (1) and (2) show the OLS estimates, both without and with macroeconomic controls, which consist of employment rate, GDP per capita, log wage, and log export. Columns (3) and (4) show the IV estimates, again with and without these controls.

The coefficients on housing prices align with the findings in the existing literature. The estimates in Column (4) indicate that the elasticity of workers' demand for a small prefecture with respect to local housing prices is approximately -3.606 for workers with low-education fathers and -2.848 for workers with high-education fathers. In comparison, Diamond (2016) reports an elasticity of around -2.496 for non-college workers and -1.312 for college-educated workers, though her estimates refer to elasticity with respect to rent rather than housing purchase prices.

The results of the first stage of the 2SLS estimation are shown in Appendix Table B.11. The coefficients exhibit the expected signs. The coefficients of the HPR policy, when housing price is the independent variable, are negative. This indicates that the implementation of the HPR policy leads to a reduction in housing prices.

The results are robust when the first step is made more complex by adding the worker's education and existing migration patterns between origin and destination and by adjusting the thresholds for father's education and birth cohorts.

8.3 Counterfactual

I use iteration to obtain counterfactuals. I focus on policies on Tier-1 prefectures as they are the most important destinations. Among migrants, approximately one-third relocate to Tier-1 prefectures. However, this overall figure masks significant heterogeneity across parental backgrounds: the proportion is about one-quarter for children with low-education fathers and 41% for those with high-education fathers.

I first analyze the impact of a housing price regulation policy that reduces housing construction costs by 10% in Tier-1 prefectures. In practice, I subtract the residuals of the housing market equation (14) by 0.1. Table 11 shows the counterfactual results. Δ% refers to percentage point changes before and after the policy. For example, for Fedutype = L, 85.15% stayed in the origin before the policy and 84.36% stayed after, the change is 84.36 - 85.15 = -.79. The results imply that, while the fraction of workers with low-educated fathers who migrate to Tier-1 prefectures increase by .95 percentage points, the increase is much larger for workers with high-education fathers. This is because the high initial housing prices in Tier-1 prefectures make workers with disadvantaged backgrounds unlikely to consider migrating there, while workers with high-education fathers are more likely to be at the margin of making that decision. This corresponds to the discussion in Section 4 about how many children with low-education fathers are at the margin relative to those with high-education fathers.

Table 11: Counterfactual Results I

Fedutype	**Stay** (Δ%)	**Tier-1** (Δ%)	**Others** (Δ%)
L	-0.79	1.01	-0.22
H	-1.20	1.70	-0.49

I then analyze the impact of a targeted housing subsidy policy offering a 10% discount for migrants in Tier-1 prefectures whose fathers have low education levels. Migrants with low-educated fathers now face housing prices at 90% of the actual prices, while those with highly educated fathers pay the full price. This policy is much for effective at increasing the migration of workers with low-education fathers. Some workers with high-education fathers are crowded out, but the magnitude is not very big.

Table 12: Counterfactual Results II

Fedutype	**Stay** ($\Delta\%$)	**Tier-1** ($\Delta\%$)	**Others** ($\Delta\%$)
L	-0.92	1.18	-0.26
H	0.23	-0.29	0.06

Overall, the two counterfactual analyses suggest that a targeted policy is more effective in increasing migration among workers from poor families. Although not in the model, a targeted policy is also likely more cost-efficient.

9 Conclusion

This paper examines the relationship between housing affordability, internal migration, and intergenerational mobility in China, highlighting the multifaceted impact of rising housing costs on socioeconomic disparities. Utilizing the Housing Purchase Restriction policy as a natural experiment and an instrumental variable approach, this paper documents that elevated housing costs significantly deter internal migration, with disproportionately adverse effects on children from less privileged backgrounds. This dynamic exacerbates income disparities and diminishes intergenerational mobility, reinforcing the cycle of economic disadvantage.

This paper proposes that in China, a country characterized by geographical inequality (Xie and Zhou, 2014) and high intergenerational persistence, implementing housing price regulations has the potential to alleviate both concerns by encouraging greater internal migration. This implication somewhat relates to the literature explaining the mechanism behind the Great Gatsby Curve, which refers to the positive correlation between inequality and intergenerational persistence observed in many countries. While the literature focuses on education attainment as a main mechanism, this paper points out that even with the same education, access to opportunities is different with different parental backgrounds.

This paper's conceptual framework is relevant in many developing countries where migration plays crucial roles in social mobility. For future research, comparative studies across countries with varying housing markets and migration policies could offer a broader perspective on the global implications of housing affordability for intergenerational mobility.

In conclusion, this research underscores the critical need for targeted policies to alleviate housing affordability issues and promote economic equality and mobility. As housing markets continue to evolve, ongoing research and policy innovation will be essential to mitigating the economic and social impacts of housing affordability.

A Appendix: Continuous Model

This section shows a continuous version of the model where, instead of High and Low type, both children's income and fathers' socioeconomic status are continuous. The model is based on Borjas (1987). The model suggests that housing price changes may increase or decrease migration and then intergenerational mobility, depending on parameters and specific distributions of disutility with respect to migration.

I assume that if a child stays at the origin, they get $Y_{stay} = \mu_0 + \delta_0 F$; if the children migrate, they get $Y_{migrate} = \mu_1 + \delta_1 F$. Migration requires a cost of $M(h, F)$, which depends on the housing price gap h and the father's socioeconomic status F. I assume that $\frac{\partial M(h,F)}{\partial h} > 0$, meaning that an increase in housing price gaps increases migration cost. People also have mental costs for migration, which I call "disutility" and denote as τ. τ is independent of F and the realized Y. I denote the CDF and PDF function of τ as $F_\tau(\tau)$ and $f_\tau(\tau)$

This model makes several assumptions. First, instead of analyzing the heterogeneity with respect to the potential migrants' education as in Borjas (1987), I assume here that each individual has the same level of education, and the father's socioeconomic status, F, affects the children's income, Y. Second, the disutility function is independent of F and the realized Y.

Combining the above benefit and cost, a child will migrate if:

$$Y_{migrate} - M(h, F) - \tau > Y_{stay} \tag{18}$$

Rearrange the equation and substitute $Y_{migrate}$ and Y_{stay}, the above equation is equivalent to:

$$(\mu_1 - \mu_0) + (\delta_1 - \delta_0)F - M(h, F) > \tau \tag{19}$$

Denote the lefthand side as $B(h, F)$, which is the net benefit of migrating. For a child faced with their father's socioeconomic status F and housing price gap h, the probability that they migrate is $Pr(migrate|h, F) = Pr(B(h, F) > \tau) = F_\tau(B(h, F))$.

In our main regression analysis, I regress the migration indicator on the interaction of paternal socioeconomic status and the housing price gap. The coefficient of interest, the cross derivative of migration probability with respect to the housing price gap and the father's socioeconomic status, can be derived from the model.

$$\frac{\partial^2 Pr(migrate|h, F)}{\partial h \partial F} = \underbrace{[-f_\tau(B)M''_{Fh}]}_{\text{heterogeneous cost}} + \underbrace{[-f'_\tau(B)B'_F M'_h]}_{\text{position at disutlity distribution}}$$

Similarly, I derive the cross derivative of the expected income with respect to h and F:

$$\frac{\partial^2 E[Y|h,F]}{\partial h \partial F} = \underbrace{[-(B+M)f_\tau(B)M''_{Fh}]}_{\text{heterogeneous cost}} + \underbrace{[-(B+M)f'_\tau(B)B'_F M'_h]}_{\text{position at disutility distribution}} + \underbrace{[-(\delta_1-\delta_0)f_\tau(B)M'_h]}_{\text{heterogeneous benefit}}$$

B Appendix: Figures and Tables

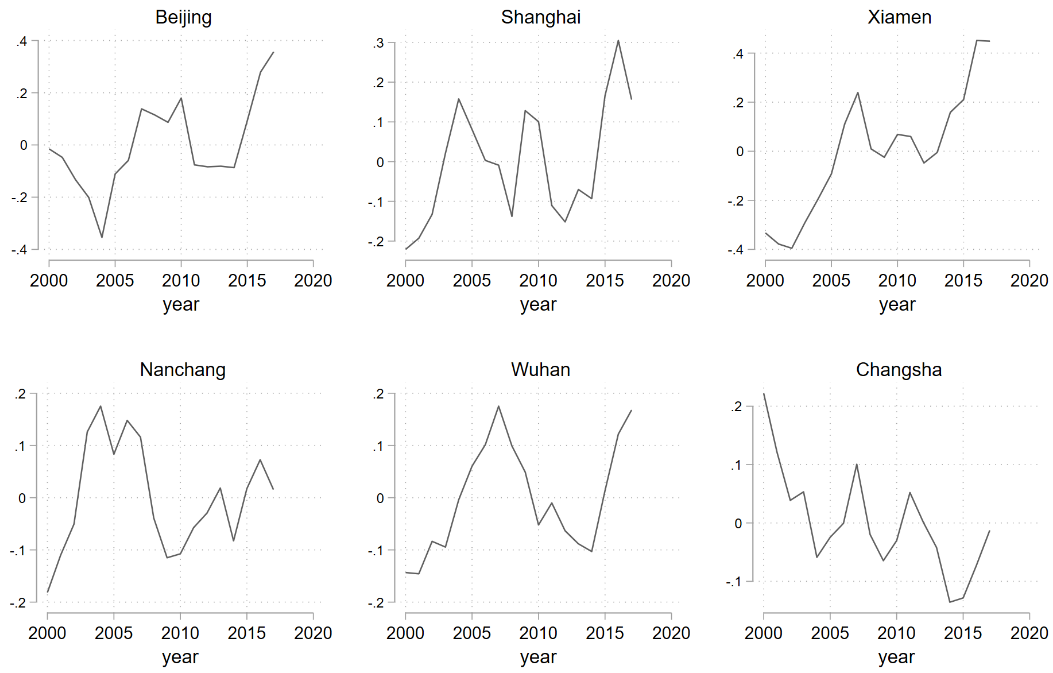


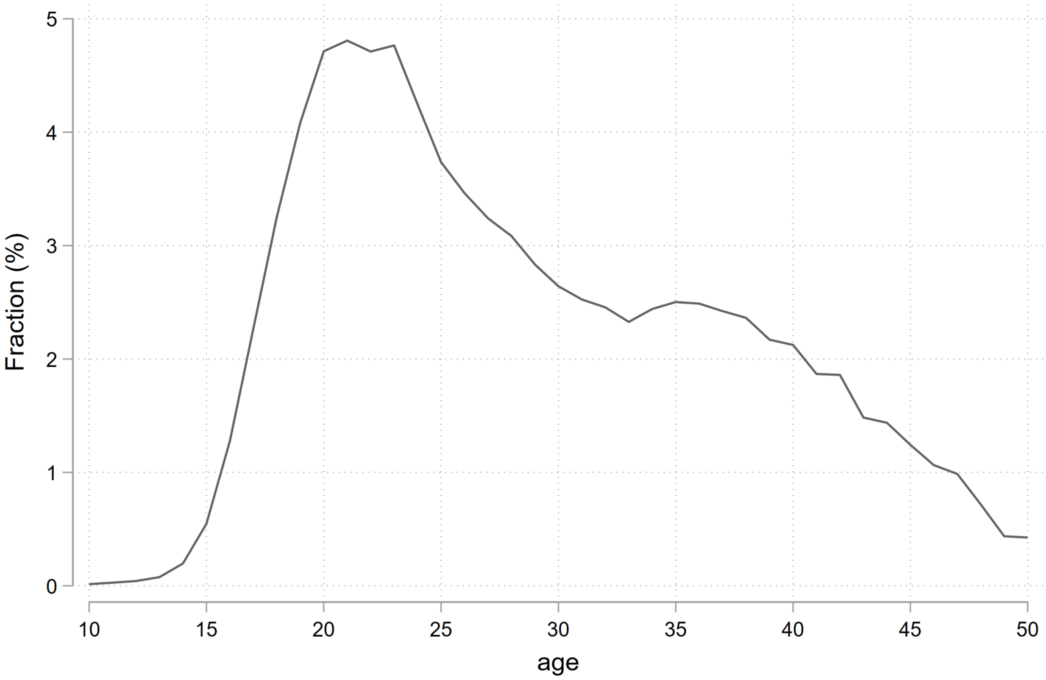
Figure B.1: Housing price residuals in six major prefectures in China

NOTE: The figure shows the residualized housing prices obtained from regressing housing prices (measured by log yuan per square meter in 2000RMB) on prefecture fixed effects and year fixed effects. The figure underscores the considerable temporal and regional variation across prefectures.

Figure B.2: HPR Event Study Plot

NOTE: The figure shows the event study plot for the policy. Using the prefecture-year level panel data, housing prices are regressed on prefecture fixed effects, year fixed effects, and interaction terms between year dummies and an indicator for ever having adopted the policy. The figure shows the estimates and confidence intervals for the coefficients of the interaction terms, which reflect the changes in differences in housing prices between the ever-treated and never-treated prefectures over time.

Figure B.3: Age-specific Migration Probability



NOTE: The figure shows the age-specific migration probability rate, which is the proportion of people who migrated at a specific age out of the total number of migrants.

C Solving for the Counterfactuals

I use iteration to obtain counterfactuals. Specifically, from the estimation above, I derive the estimates for parameters $\beta^d, \beta^r, \beta^w, \beta^h_{fedu\_L}, \beta^h_{fedu\_H}$ and fixed effects $\{\mu_p\}$ and $\{\pi_{fedu,t}\}$. $\{\xi^{fedu}_{pt}\}$ and $\{\varepsilon_{pt}\}$ for all the prefectures and birth cohorts are calculated using the following equations:

$$\xi^{fedu}_{pt} = \delta^{fedu}_{pt} - (\beta^w w_{pt} + \beta^h_{fedu\_L} h_{pt} + \beta^h_{fedu\_H} h_{pt} + \mu_p + \pi_{fedu,t})$$

$$\varepsilon_{pt} = h_{pt} - \Gamma(lnL_{pt} + w_{pt})$$

With the estimated parameters, fixed effects, and data on worker characteristics and prefectures, the counterfactual equilibrium is obtained through the following steps:

1. I denote the set of labor supply, housing prices, prefecture average attractiveness for all the prefectures and birth cohorts as $\{L^0_{pt}\}$, $\{h^0_{pt}\}$, and $\{\delta^0_{fedu,pt}\}$.

2. Increase the housing prices of the prefectures of interest, e.g., Beijing, by 10% via increasing the $\varepsilon_{Beijing}$, capturing changes in the housing supply regulations. Alternatively, I could decrease the housing prices faced by children of low-education fathers by 10%, representing a housing substitution policy.

3. Change in $\varepsilon_{Beijing}$ leads to change in housing price $h_{Beijing}$, which then leads to change in Beijing's attractiveness as a residential prefecture $\delta^{fedu}_{Beijing}$, and eventually affect L_p for all the prefectures, as workers change the migration destinations from Beijing to other prefectures. I denote these new labor supplies in different destinations $\{L^1_{pt}\}$.

4. Changes in labor supply will then change the housing prices in the housing market according to Equation (14). A decrease in labor supply in Beijing decreases its housing prices and vice versa for prefectures other than Beijing. A new set of housing prices $\{h^1_{pt}\}$ can be obtained by inserting $\{L^1_{pt}\}$, $\{w_{pt}\}$ and $\{\varepsilon_{pt}\}$ into Equation (14).

5. Changes in housing prices will then change the attractiveness of the prefectures according to Equation (17). A new set of attractiveness $\{\delta^1_{fedu,pt}\}$ can be obtained by inserting $\{h^1_{pt}\}$, $\{w_{pt}\}$, $\{\mu_{pt}\}$, $\{\pi_{fedu,t}\}$, and $\{\xi^{fedu}_{pt}\}$ into Equation (17).

6. Changes in attractiveness $\{\delta^1_{fedu,pt}\}$ of the destinations will affect labor supply again. A new set of labor supply $\{L^2_{pt}\}$ can be obtained by inserting $\{\delta^1_{fedu,pt}\}$

into Equation (15).

7. I then compare $\{L^1_{pt}\}$ and $\{L^2_{pt}\}$. If they are close enough, I stop the loop and claim I have reached a new equilibrium. Otherwise, I keep iterating Steps 4-6 until the two subsequent labor supplies $\{L^s_{pt}\}$ and $\{L^{s+1}_{pt}\}$ are close enough.

Table B.1: The List of 46 Prefectures that Adopted the Housing Purchase Restriction Policy with Dates of Announcing and Abolishing the Policy

Prefecture	Start Year	Start Month	End Year	End Month
Beijing	2010	4		
Tianjin	2010	10	2014	10
Shijiazhuang	2011	2	2014	9
Taiyuan	2011	1	2014	8
Huhehaote	2011	4	2014	6
Shenyang	2011	3	2014	9
Dalian	2011	3	2014	9
Changchun	2011	5	2015	6
Haerbin	2011	2	2014	8
Shanghai	2010	10		
Nanjing	2010	10	2014	9
Wuxi	2011	2	2014	8
Xuzhou	2011	5	2014	8
Suzhou	2011	3	2014	9
Hangzhou	2010	10	2014	8
Ningbo	2010	10	2014	7
Wenzhou	2011	3	2013	8
Shaoxing	2011	8	2014	8
Jinhua	2011	3	2014	8
Quzhou	2011	9	2014	7
Zhoushan	2011	8	2013	1
Taizhou	2011	8	2014	8
HeFei	2011	1	2014	8
Fuzhou	2010	10	2014	8
Xiamen	2010	10	2014	8
Nanchang	2011	2	2014	7
Jinan	2011	1	2014	7
Qinghai	2011	1	2014	9
Zhengzhou	2011	1	2014	8
Wuhan	2011	1	2014	7
Changsha	2011	3	2014	8
Guangzhou	2010	10		
Shenzhen	2010	9		
Zhuhai	2011	11	2016	3
Foshan	2011	3	2014	8
Nanning	2011	3	2014	10
Haikou	2010	10	2014	7
Sanya	2010	10		
Chengdu	2011	2	2015	1
Guiyang	2011	2	2014	9
Kunming	2011	1	2014	8
Xian	2011	3	2014	9
Lanzhou	2011	3	2014	7
Xining	2011	8	2014	9
Yinchuan	2011	2	2014	8
Wulumuqi	2011	3	2014	8

NOTE: This table lists the dates of announcing and abolishing the Housing Purchase Restriction (HPR) Policy from Chen et al. (2017)

Table B.2: Descriptive Statistics

	Mean	SD
Migration	0.20	0.40
lnIncome	9.79	1.67
FatherEduy	5.72	4.51
HPgap	0.55	0.49
Male	0.46	0.50
Eduy	10.04	3.76
Rural Hukou	0.67	0.47
Age	43.17	8.94

Table B.3: IV Results: First Stage

	(1) HPgap	(2) FatherEduy× HPgap	(3) HPgap	(4) FatherEduy× HPgap
HPR	-0.157***	-13.224***	-0.153***	-13.405***
	(0.015)	(0.583)	(0.015)	(0.589)
FatherEduy× HPR	0.000	2.721***	-0.001	2.765***
	(0.001)	(0.058)	(0.001)	(0.059)
FatherEduy× birthyearFE	Y	Y	Y	Y
Eduy× birthyearFE	-	-	Y	Y
Obs.	15,128	15,128	15,128	15,128
Adj. R-sq	0.995	0.812	0.995	0.813

NOTE: The figure shows the first-stage results. When included, Eduy×HPgap is also instrumented using Eduy×HPR. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table B.4: Robustness: Sample Construction

	(1) Migration	(2) Migration	(3) Migration	(4) Migration	(5) lnIncome	(6) lnIncome	(7) lnIncome	(8) lnIncome
FatherEduy× HPgap	0.009***	0.009***	0.010***	0.010***	0.030**	0.030**	0.035**	0.033**
	(0.003)	(0.003)	(0.003)	(0.003)	(0.013)	(0.014)	(0.015)	(0.015)
HPgap	-0.802*	-1.038**	-0.788*	-0.769*	-3.763**	-4.165**	-3.403**	-3.688**
	(0.432)	(0.503)	(0.429)	(0.435)	(1.698)	(1.782)	(1.709)	(1.814)
Eduy× birthyearFE	Y	Y	Y	Y	Y	Y	Y	Y
Exclude	-	adjacent to tier-1	high GDP per capita	popular destinations	-	adjacent to tier-1	high GDP per capita	popular destinations
Obs.	15,128	14,408	14,677	14,200	12,189	11,621	11,839	11,422
Mean(Dep. Var.)	0.201	0.202	0.205	0.210	9.793	9.782	9.770	9.745

NOTE: This table shows the robustness results for different samples using the HPR policy treatment as the instrumental variable. Fathers' years of education interacted with birth year fixed effects are controlled for in all columns. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table B.5: Robustness: IV Validity

	(1) Migration	(2) Migration	(3) Migration	(4) Migration	(5) lnIncome	(6) lnIncome	(7) lnIncome	(8) lnIncome
FatherEduy× HPgap	0.012***	0.011***	0.009***	0.008***	0.024***	0.021*	0.030***	0.027**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.009)	(0.011)	(0.010)	(0.012)
HPgap	-0.923***	-0.911**	-0.802**	-0.788**	-3.388**	-3.373**	-3.763***	-3.748***
	(0.354)	(0.355)	(0.345)	(0.346)	(1.409)	(1.405)	(1.440)	(1.436)
Eduy× HPgap			0.009***	0.009***			-0.018	-0.018
			(0.003)	(0.003)			(0.014)	(0.014)
FatherEduy× Evertreated		0.004		0.005		0.018		0.017
		(0.006)		(0.006)		(0.031)		(0.031)
K-P F-stat	30.203	30.308	20.600	20.680	58.633	58.562	39.776	39.733
Obs.	15,128	15,128	15,128	15,128	12,189	12,189	12,189	12,189
Mean(Dep. Var.)	0.201	0.201	0.201	0.201	9.793	9.793	9.793	9.793

NOTE: This table shows the robustness results controlling for the average weighted ever-treated indicators and its interaction term with fathers' education. Fathers' years of education interacted with birth year fixed effects are controlled for in all columns. Odd columns are copied directly from Table 4, and even columns are corresponding specifications controlling for the average weighted ever-treated indicators and its interaction term with fathers' education. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table B.6: Robustness: Alternative Housing Price Dataset

	(1) Migration	(2) Migration	(3) lnIncome	(4) lnIncome
FatherEduy× HPgap	0.011***	0.008***	0.026**	0.039***
	(0.002)	(0.003)	(0.010)	(0.011)
HPgap	-2.007***	-1.767***	-4.739**	-6.137***
	(0.657)	(0.625)	(2.207)	(2.317)
Eduy× HPgap		0.009***		-0.042***
		(0.003)		(0.016)
FatherEduy× birthyearFE	-	Y	-	Y
K-P F-stat	16.666	11.601	27.654	18.942
Obs.	15,103	15,103	12,171	12,171
Mean(Dep. Var.)	0.199	0.199	9.794	9.794

NOTE: This table shows the results leveraging the Anjuke data using the HPR policy treatment as the instrumental variable. Fathers' years of education interacted with birth year fixed effects are controlled for in all columns. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table B.7: Robustness: Average Income Across Waves

	(1) lnIncome	(2) lnIncome	(3) lnIncome multi-year	(4) lnIncome multi-year
FatherEduy× HPgap	0.024***	0.030***	0.024***	0.030***
	(0.009)	(0.010)	(0.008)	(0.009)
HPgap	-3.388**	-3.763***	-3.712***	-4.131***
	(1.409)	(1.440)	(1.330)	(1.359)
Eduy× HPgap		-0.018		-0.019
		(0.014)		(0.012)
FatherEduy× birthyearFE	Y	Y	Y	Y
K-P F-stat	58.633	39.776	56.749	38.474
LM test	110.914	112.987	108.468	110.525
Obs.	12,189	12,189	12,977	12,977
Mean(Dep. Var.)	9.793	9.793	9.312	9.312

NOTE: This table presents results using the log of average income across waves, as opposed to using children's income solely from 2017. The average income calculation only includes data from waves where the worker's residential prefecture remained the same as that of 2017. Notably, about 66% of the observations incorporate income data from multiple waves. For these observations, the income is averaged across the relevant waves and adjusted for the Consumer Price Index (CPI). The analysis controls for fixed effects associated with the observed waves, such as 'observed in both 2015 and 2017'. Additionally, all columns include controls for fathers' education interacted with birth year fixed effects. Columns (1) and (2) replicate results from Table 4 Columns (3) and (4). Differences in observations between Columns (3) and (4) versus (1) and (2) arise when income data from 2015 or 2019 are available for workers missing income data in 2017. Standard errors are clustered at the prefecture-birth year level, with significance levels indicated as follows: ***p<0.01, **p<0.05, *p<0.1.

Table B.8: Robustness: Residualized Fathers' Education

	(1) Migration	(2) Migration	(3) lnIncome	(4) lnIncome
FatherEduy_r× HPgap	0.008***	0.008***	0.025**	0.028**
	(0.002)	(0.002)	(0.012)	(0.012)
HPgap	-0.897**	-0.897**	-3.054**	-3.032**
	(0.363)	(0.363)	(1.463)	(1.464)
Eduy× HPgap		0.004		-0.014
		(0.004)		(0.018)
FatherEduy_r× birthyearFE	Y	Y	Y	Y
K-P F-stat	28.646	19.096	51.546	34.360
LM test	50.500	50.500	96.228	96.270
Obs.	13,187	13,187	10,600	10,600
Mean(Dep. Var.)	0.146	0.146	9.737	9.737

NOTE: This table shows the results using residualized fathers' years of education to measure parental background, using the HPR policy treatment as the instrumental variable. The number of observations differs from the main analyses because some prefecture-by-birth-year cells contain singletons. These singletons are dropped because the residualized father's education cannot be obtained for them. Fathers' years of education interacted with birth year fixed effects are controlled for in all columns. In Columns (1) and (3), the endogenous variables are $HPgap$ and $FatherEduy\_r \times HPgap$; the instrumental variables are HPR and $FatherEduy\_r \times HPR$. In Columns (2) and (4), the endogenous variables are $HPgap$, $FatherEduy\_r \times HPgap$, and $Edu \times HPgap$; the instrumental variables are HPR, $FatherEduy\_r \times HPR$, and $Edu \times HPR$. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table B.9: Intergenerational Income Elasticity (IGE)

	(1)	(2)
	lnIncome	lnIncome
lnFatherIncome	0.382***	0.385***
	(0.058)	(0.058)
lnFatherIncome× HPgap	0.328*	0.421**
	(0.171)	(0.188)
HPgap	-3.672***	-3.985***
	(1.359)	(1.387)
Eduy× HPgap		-0.016
		(0.014)
K-P F-stat	64.569	43.517
LM test	118.704	120.277
Obs.	12,172	12,172
Mean(Dep. Var.)	9.793	9.793

NOTE: This table shows the results using imputed fathers' income in place of fathers' education. In Column (1), the endogenous variables are $HPgap$ and $FatherIncome \times HPgap$; the instrumental variables are HPR and $FatherIncome \times HPR$. In Columns (2), the endogenous variables are $HPgap$, $FatherIncome \times HPgap$, and $Edu \times HPgap$; the instrumental variables are HPR, $FatherIncome \times HPR$, and $Edu \times HPR$. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. The K-P F-stat row shows the F-stats of the first stages. The LM test row shows the statistics for the underidentification test. A rejection of the null indicates that the matrix of first-stage coefficients has full column rank, i.e., the model is identified. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table B.10: Robustness: Impute Household Assets

	(1) Migration	(2) Migration	(3) lnIncome	(4) lnIncome
lnHHAssets× HPgap	0.131***	0.093***	0.287***	0.351***
	(0.025)	(0.027)	(0.105)	(0.116)
HPgap	-1.007***	-0.866**	-3.340**	-3.684***
	(0.357)	(0.346)	(1.380)	(1.418)
Eduy× HPgap		0.010***		-0.016
		(0.003)		(0.014)
lnHHAssets× birthyearFE	Y	Y	Y	Y
K-P F-stat	30.568	20.870	61.072	41.335
LM test	56.113	57.304	113.179	115.435
Obs.	15,107	15,107	12,172	12,172
Mean(Dep. Var.)	0.200	0.200	9.793	9.793

NOTE: This table shows the results using imputed household assets in place of fathers' education. Imputed household assets interacted with birth year fixed effects are controlled in all columns. In Columns (1) and (3), the endogenous variables are $HPgap$ and $HHAssets \times HPgap$; the instrumental variables are HPR and $HHAssets \times HPR$. In Columns (2) and (4), the endogenous variables are $HPgap$, $HHAssets \times HPgap$, and $Edu \times HPgap$; the instrumental variables are HPR, $HHAssets \times HPR$, and $Edu \times HPR$. Control variables include gender, parental *Hukou* status, log wage and log employment at the origin prefecture, weighted average of the log wage and log employment at the potential destination prefectures, fixed effects for origin prefecture, birth year, and education level. The K-P F-stat row shows the F-stats of the first stages. The LM test row shows the statistics for the underidentification test. A rejection of the null indicates that the matrix of first-stage coefficients has full column rank, i.e., the model is identified. Standard errors are clustered at prefecture times birth year level; ***p<0.01, **p<0.05, *p<0.1.

Table B.11: Conditional Logit: First Stage

	(1) HousingPrices	(2) FatherEduHigh× HP	(3) HousingPrices	(4) FatherEduHigh× HP
HPR	-0.140***	-0.419***	-0.121***	-0.407***
	(0.027)	(0.038)	(0.030)	(0.039)
FatherEduHigh× HPR	0.002	0.697***	0.002	0.692***
	(0.003)	(0.069)	(0.003)	(0.069)
FatherEduHigh× birthcohort FE	Y	Y	Y	Y
Prefecture FE	Y	Y	Y	Y
Controls	-	-	Y	Y
Obs.	607	607	603	603

NOTE: Standard errors are clustered at the prefecture level. HP is the abbreviation for housing prices. Columns (1) and (2) show the estimates without macroeconomic controls. Columns (3) and (4) show the results with macroeconomic controls including employment rate, GDP per capita, log wage, and log export.

References

Christoph Albert and Joan Monras. Immigration and spatial equilibrium: the role of expenditures in the country of origin. *American Economic Review*, 112(11):3763–3802, 2022.

Alberto Alesina, Sebastian Hohmann, Stelios Michalopoulos, and Elias Papaioannou. Intergenerational mobility in africa. *Econometrica*, 89(1):1–35, 2021.

James Alm, Weizheng Lai, and Xun Li. Housing market regulations and strategic divorce propensity in china. *Journal of Population Economics*, 35(3):1103–1131, 2022.

Manuela Angelucci. Migration and financial constraints: Evidence from mexico. *Review of Economics and Statistics*, 97(1):224–228, 2015.

Costas Azariadis and John Stachurski. Poverty traps. *Handbook of economic growth*, 1: 295–384, 2005.

Samuel Bazzi. Wealth heterogeneity and the income elasticity of migration. *American Economic Journal: Applied Economics*, 9(2):219–255, 2017.

George J Borjas. Self-selection and the earnings of immigrants. Technical report, National Bureau of Economic Research, 1987.

Loren Brandt and Carsten A Holz. Spatial price differences in china: Estimates and implications. *Economic development and cultural change*, 55(1):43–86, 2006.

Gharad Bryan, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak. Underinvestment in a profitable technology: The case of seasonal migration in bangladesh. *Econometrica*, 82(5):1671–1748, 2014.

Shu Cai. Migration under liquidity constraints: Evidence from randomized credit access in china. *Journal of Development Economics*, 142:102247, 2020.

Ting Chen, Laura Liu, Wei Xiong, Li-An Zhou, et al. Real estate boom and misallocation of capital in china. *Work. Pap., Princeton Univ., Princeton, NJ*, 9, 2017.

Yu Chen and Jufen Wang. Social integration of new-generation migrants in shanghai china. *Habitat International*, 49:419–425, 2015.

Yuanyuan Chen, Le Wang, and Min Zhang. Informal search, bad search?: the effects of job search method on wages among rural migrants in urban china. *Journal of Population Economics*, 31:837–876, 2018.

Yunjie Chen, Yongzheng Liu, and Hailun Wei. The unintended consequences of home-buying restriction in china. *Available at SSRN 4637419*, 2023.

Raj Chetty, Nathaniel Hendren, and Lawrence F Katz. The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. *American Economic Review*, 106(4):855–902, 2016.

Rebecca Diamond. The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *American Economic Review*, 106(3):479–524, 2016.

Zaichao Du and Lin Zhang. Home-purchase restriction, property tax and housing price in china: A counterfactual analysis. *Journal of Econometrics*, 188(2):558–568, 2015.

Karen Eggleston, Ang Sun, and Zhaoguo Zhan. The impact of rural pensions in china on labor migration. *The World Bank Economic Review*, 32(1):64–84, 2018.

Yi Fan, Junjian Yi, and Junsen Zhang. Rising intergenerational income persistence in china. *American Economic Journal: Economic Policy*, 13(1):202–230, 2021.

Hanming Fang, Quanlin Gu, Wei Xiong, and Li-An Zhou. Demystifying the chinese housing boom. *NBER macroeconomics annual*, 30(1):105–166, 2016.

James J Feigenbaum. Intergenerational mobility during the great depression. 2015.

Benjamin Feigenberg, Ben Ost, and Javaeria A Qureshi. Omitted variable bias in interacted models: A cautionary tale. *Review of Economics and Statistics*, pages 1–47, 2023.

Andrew Foote. The effects of negative house price changes on migration: Evidence across us housing downturns. *Regional Science and Urban Economics*, 60:292–299, 2016.

Qingen Gai, Naijia Guo, Bingjing Li, Qinghua Shi, Xiaodong Zhu, et al. Migration costs, sorting, and the agricultural productivity gap. 2021.

Li Gan, Zhichao Yin, Nan Jia, Shu Xu, Shuang Ma, Lu Zheng, et al. Data you need to know about china. *Springer Berlin Heidelberg. https://doi*, 10(978):3, 2014.

Peter Ganong and Daniel Shoag. Why has regional income convergence in the us declined? *Journal of Urban Economics*, 102:76–90, 2017.

Carlos Garriga, Aaron Hedlund, Yang Tang, and Ping Wang. Rural-urban migration, structural transformation, and housing markets in china. *American Economic Journal: Macroeconomics*, 15(2):413–440, 2023.

Garance Genicot, Debraj Ray, and Carolina Concha-Arriagada. Upward mobility in developing countries. Technical report, National Bureau of Economic Research, 2024.

Honge Gong, Andrew Leigh, and Xin Meng. Intergenerational income mobility in urban china. *Review of Income and Wealth*, 58(3):481–503, 2012.

Joseph Gyourko, Yang Shen, Jing Wu, and Rongjie Zhang. Land finance in china: Analysis and review. *China Economic Review*, 76:101868, 2022.

Zhangfeng Jin and Junsen Zhang. Access to local citizenship and internal migration in a developing country: Evidence from a hukou reform in china. *Journal of Comparative Economics*, 51(1):181–215, 2023.

Gaurav Khanna, Wenquan Liang, Ahmed Mushfiq Mobarak, and Ran Song. The productivity consequences of pollution-induced migration in china. Technical report, National Bureau of Economic Research, 2021.

David Lagakos, Samuel Marshall, Ahmed Mushfiq Mobarak, Corey Vernot, and Michael E Waugh. Migration costs and observational returns to migration in the developing world. *Journal of Monetary Economics*, 113:138–154, 2020.

Sang Yoon Tim Lee, Nicolas Roys, and Ananth Seshadri. The causal effect of parents' education on children's earnings. *Manuscript. https://www. ssc. wisc. edu/aseshadr/WorkingPapers/spillover. pdf [accessed 26 October 2021]*, 2024.

Han Li, Jiangyi Li, Yi Lu, and Huihua Xie. Housing wealth and labor supply: Evidence from a regression discontinuity design. *Journal of Public Economics*, 183:104139, 2020.

Lixing Li and Xiaoyu Wu. Housing price and intergenerational co-residence in urban china. *Journal of Housing Economics*, 45:101596, 2019.

Victor Jing Li, Andy Wui Wing Cheng, and Tsun Se Cheong. Home purchase restriction and housing price: A distribution dynamics analysis. *Regional Science and Urban Economics*, 67:1–10, 2017.

Hong Liu, Lili Liu, and Fei Wang. Housing wealth and fertility: evidence from china. *Journal of Population Economics*, 36(1):359–395, 2023.

Darshini Mahadevia, Zhiyan Liu, and Xiuming Yuan. Housing of urban migrants in india and china. *Environment and Urbanization Asia*, 3(1):143–163, 2012.

Lina Meng, Xiao Xiao, and Yinggang Zhou. Housing boom and household migration decision: new evidence from china. *The Journal of Real Estate Finance and Economics*, 67(3):453–479, 2023.

Martin Nybom and Jan Stuhler. Biases in standard measures of intergenerational income dependence. *Journal of Human Resources*, 52(3):800–825, 2017.

Claudia Olivetti and M Daniele Paserman. In the name of the son (and the daughter): Intergenerational mobility in the united states, 1850–1940. *American Economic Review*, 105(8):2695–2724, 2015.

William W Olney and Owen Thompson. The determinants of declining internal migration. Technical report, National Bureau of Economic Research, 2024.

John Parman. American mobility and the expansion of public education. *The Journal of Economic History*, 71(1):105–132, 2011.

Kenneth S Rogoff and Yuanchen Yang. Peak china housing. Technical report, National Bureau of Economic Research, 2020.

Albert Saiz. The geographic determinants of housing supply. *The Quarterly Journal of Economics*, 125(3):1253–1296, 2010.

Gary Solon. Intergenerational income mobility in the united states. *The American Economic Review*, pages 393–408, 1992.

Ang Sun and Qinghua Zhang. Who marries whom in a surging housing market?. *Journal of Development Economics*, 146:102492, 2020.

Hui Ren Tan. A different land of opportunity: The geography of intergenerational mobility in the early twentieth-century united states. *Journal of Labor Economics*, 41(1):77–102, 2023.

Trevor Tombe and Xiaodong Zhu. Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–1872, 2019.

Xin Wang and Yi Wen. Housing prices and the high chinese saving rate puzzle. *China Economic Review*, 23(2):265–283, 2012.

Zachary Ward. Internal migration, education, and intergenerational mobility: Evidence from american history. *Journal of Human Resources*, 57(6):1981–2011, 2022.

Shang-Jin Wei and Xiaobo Zhang. The competitive saving motive: Evidence from rising sex ratios and savings rates in china. *Journal of political Economy*, 119(3): 511–564, 2011.

John Whalley and Shunming Zhang. A numerical simulation analysis of (hukou) labour mobility restrictions in china. *Journal of Development Economics*, 83(2):392–410, 2007.

Yu Xie and Xiang Zhou. Income inequality in today's china. *Proceedings of the national academy of Sciences*, 111(19):6928–6933, 2014.

Jeffrey E Zabel. Migration, housing market, and labor market responses to employment shocks. *Journal of Urban Economics*, 72(2-3):267–284, 2012.

Renjie Zhao and Jiakai Zhang. Rent-tax substitution and its impact on firms: Evidence from housing purchase limits policy in china. *Regional Science and Urban Economics*, 96:103804, 2022.

Yaohui Zhao. Leaving the countryside: rural-to-urban migration decisions in china. *American Economic Review*, 89(2):281–286, 1999.

Angela Zheng and James Graham. Public education inequality and intergenerational mobility. *American Economic Journal: Macroeconomics*, 14(3):250–282, 2022.