Urban Housing Prices and Migration's Fertility Intentions : Based on the 2018 China Migrants' Dynamic Survey*

Jingwen Tan[a,**,1], Shixi Kang[a,*,1]

[a]School of Economics, Henan University, Jinming Avenue, Kaifeng, 475004, China

ARTICLE INFO

Keywords:
Migrant
Housing Prices
Fertility Intentions
Moderating Effects
China

Abstract

The size of China's migrant population continues to expand, the fertility rate is significantly lower than the stable generation replacement level of the population, and the potential structural imbalance in human resource supply has attracted widespread attention. This paper estimates the impact of house prices on the fertility intention of the migrant population using data from the 2018 China Migrants Dynamic Survey. The results show that: for every 100% increase in the ratio of house price to household income of the migrant population, the fertility intention of the female migrant population of the right age in the inflow area will decrease by 4.42%; the sensitivity of the fertility intention of the migrant population to house price is influenced by the moderating effect of infrastructure construction in the inflow area; the fertility intention of the female migrant population of the right age with lower age, smaller family size and higher education is higher. Based on the above findings, the study tries to provide a new practical perspective for the mainline institutional change and balanced economic development in China's economic transition stage.

1. Introduction

Human capital is in an important position in the process of economic development (Gennaioli et al., 2013). In recent years, the stock and increment of human capital in China has gradually shown a tendency to be insufficient, and the average age of the working population is 38.8 years, with a significant aging phenomenon (CUFE, 2021). During the same period, China's total fertility rate was between 1.5 and 1.6, significantly lower than the level of generational replacement needed to achieve population stability (Li et al., 2021). One of the reasons for this phenomenon is the long-standing family planning policy in China (Tan and Kang, 2021). Although the Chinese government revised its family planning policy in 2016 and 2021, China's fertility rate remains low, with a natural population growth rate of 0.34 per 1,000 in 2021 and expected to enter negative growth in 2022 (NBS, 2021a). At the same time, China's economy is expanding in size and market and still has a high labor demand (Wu and Zhang, 2020). The potential structural imbalance of human resource supply arising from low fertility rate and high labor demand in China is a matter of concern.

The migrant population is an important part of China's labor supply. Since the reform and opening up of China solved the institutional barriers to population mobility, the size of the migrant population has been expanding, up to 375.82 million by 2020, an increase of 69.73% in 10 years (Zhang and SONG, 2003). China's rural outflow population has provided important human resources for urban development and played a great role in the urbanization process (NBS, 2021b). However, due to the long-standing dualistic household registration system in China, the outflowing population often cannot receive medical and educational resources from the local inflow, which makes it difficult for them to integrate socially and affects their fertility intentions (Hao and Tang, 2015). In China, household registration is highly tied to housing, and the high housing prices prevent the labor force from settling down and obtaining a household registration in the inflowing cities, and some scholars have proposed the concept of semi-urbanization (Wang, 2006).

Improving the fertility intentions of the migrant population, alleviating the structural imbalance in human resource supply, and providing a stable labor supply for economic development are important conditions for promoting sustained economic growth in China. This paper uses data from the 2018 China Migrants Dynamic Survey to study the impact

*The authors thank the Migrant Population Service Center, National Health Commission P. R. China for data support.

*Corresponding author
**Principal Corresponding author
tjw@henu.edu.cn (J. Tan); ksx@henu.edu.cn (S. Kang)
ORCID(s): 0000-0002-3452-959X (J. Tan); 0000-0003-4847-5171 (S. Kang)
[1]These authors contributed equally to this work.

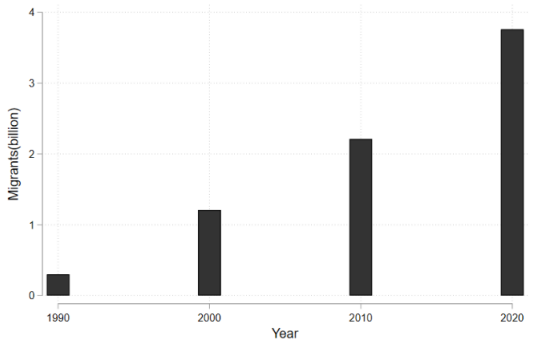

Figure 1: The Number of Migrants.
Notes: Data from CBS National Population Census Bulletin (4th, 5th, 6th and 7th) and China Statistical Yearbook 2021

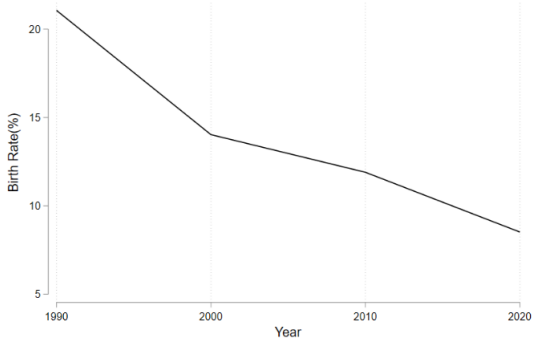

Figure 2: Changes in fertility rates in China in recent years.
Notes: Data from CBS National Population Census Bulletin (4th, 5th, 6th and 7th) and China Statistical Yearbook 2021

of housing prices on the fertility intentions of the migrant population, and analyzes the impact of infrastructure development in the inflow areas and the housing pressure of the migrant population on this mechanism.

2. Research background

The study of human capital is an important part of neoclassical economics (Dettling and Kearney, 2014). Whether it is aimed at the micro level of increasing individual well-being and improving the family environment (Petch and Halford, 2008) or the macro perspective based on promoting the reproduction of human capital and thus the proper functioning of social institutions (Gough, 2017), the study of fertility intentions has obvious relevance. A number of studies have already explored this topic from a multidimensional perspective. For example, Kim and Yeo (2019) found in a study of low fertility in Korea that socio-environmental factors can influence fertility intentions through individual environmental factors. As another example, Arai (2007) further argues for the indirect role of social relationships on fertility intentions through an empirical analysis of the potential impact of community on fertility intentions. Among them, studies using house prices as a source of influence are more extensive. Dettling and Kearney (2014) define fertility prices as housing costs and argue that the real estate market has a more direct effect on fertility intentions. Lino (2002) finds a significant negative relationship between house prices and fertility intentions, suggesting that housing costs are the largest expense of raising children.

Regarding the mechanism of the effect of housing prices on the fertility intention of the migrant population, domestic and foreign research focuses on both the willingness to integrate and the financial ability. On the one hand, the high or low housing price is related to the strength of the social integration intention of the migrant population. According to Entzinger and Biezeveld (2003), social integration includes psychological integration, cultural integration, economic integration, and political integration. House price, as a specific quantification of property rights, is an important indicator of the migrant population's integration into the society of the place of migration in

many ways, which in turn affects the willingness to stay. Since the migration of the migrant population is a short-term decision-making behavior(Du et al., 2014), lowering the threshold of ownership of their property rights is a key initiative to promote long-term residence of the migrant population and thus increase the likelihood of having children. High housing prices mean that it is more difficult for the migrant population to integrate subjectively into the local area, and accordingly, their willingness to have children decreases.

On the other hand, housing prices are also a test of the economic capacity of migrant populations. The current direction of population mobility is mainly from rural to urban areas (Lewis et al., 1954). Under the household registration system, which was born in the context of a dualistic economic system, this test is mainly in terms of job market disadvantages, financing constraints, and lack of public protection. First, migrant populations with the same endowments have less access to permanent jobs in the legitimate sector than local residents. The household registration system provides a natural barrier to entry into the job market, and self-employment becomes a rational choice for this group to cope with employment discrimination compared to wage employment(Banerjee, 1983; Roberts, 2001; Song and Appleton, 2008). Second, the migrant population suffers from a very pronounced financial disincentive to start a business. Under the imperfect financial mechanism, their entrepreneurship tends to rely more on their own capital, which further reduces the possibility of capital circulation and expanded reproduction (Paulson and Townsend, 2004). Third, the entrepreneurial risk and cost of living increase because the migrant population does not enjoy a series of social security brought by the household registration system. For example, Oates (1969) emphasizes the role of public services in a utility model of population migration, arguing that a good social security system will reduce expenditures in many ways and is an important driver of migration decisions. In summary, the series of thresholds imposed by the household registration system make the migrant population relatively less able to afford housing prices and lead to a greater reluctance to bear the various costs arising from childbearing.

There are still some gaps in academic research on fertility intentions of specific groups, and most of the benchmark indicators refer to absolute house prices. This paper takes the migrant population as the main research subject and relative house prices as the core explanatory variable, which provides a new research perspective and policy basis for pulling the benign population growth.

3. Data sources and variable selection

(1) Data sources.

This paper selects data from the "China Migrants Dynamic Survey (CMDS)" organized by the National Health Care Commission in 2018 for empirical analysis. The survey covered 31 provinces (municipalities and autonomous regions) in mainland China, and the sample was selected from the migrant population who had stayed in the local area for more than one month. The average age of the sample is in the 15-59 years old range, and a stratified, multi-stage, large-scale PPS sampling method is adopted to investigate the development, individual characteristics, social integration, and employment of China's migrant population in detail, providing data support for social science research in the direction of labor economics and demography. The data related to urban control variables are obtained from the statistical yearbooks of various cities.

(2) Variable selection.

1. Core variables

The explanatory variable of this paper is the fertility intention of the migrant population, and we choose the question "Do you have a fertility plan at the moment?" in the CMDS questionnaire. The core explanatory variable in this paper is the relative house price, which is the ratio of house price and household income in the city where the migrant population lives. Household income is measured by "what is your total monthly household income" in the CMDS questionnaire, which is a clearer indicator of the difficulty of acquiring property rights in the city than the relative house price constructed by per capita disposable income in the city.

2. Control variables screening and identification strategy

The control variables in this paper include both urban control variables and individual control variables. In existing studies on fertility, factors such as years of education, family size, age, and insurance participation are frequently found as demographic control variables. Variables such as urban health care level and disposable income per capita are frequently occurring urban control variables. In order to reduce the endogeneity problem arising from omitted variables in the research design phase, this paper selected 28 previously occurring variables from previous studies on

Table 1
Variable description and lasso regression results.

Variables	Variable Explanation	Lasso factor
intention	Respondents have fertility intention is 1, no is 0	-
houseincome	House prices in the inflow area/respondents' household income	-0.0026868
family	Number of household members of the respondent	-0.1037128
edu	Years of education of the respondents	0.0053022
health	Respondents considered themselves unhealthy as 1 and healthy as 4	0.0049792
hukou	Respondents held a residential household registration of 1 and an agricultural household registration of 0	0.0028327
job	Respondents' jobs are 2 if they are civil servants or career units, 1 if they are other units, and 0 if they are jobless	0.0023393
age	Age of respondents	-0.0099228
tgdp	Inflow to the city tertiary sector share	0.0003028
green	Inflow to urban green space area ratio	0.0000128
pbook	Per capita book collection flowing into the city	-0.0023456
pteacher	Number of elementary school teachers per capita flowing into the city	0.0018002
physician	Number of physicians per capita flowing into the city	-0.0002457

fertility intentions and controlled for variables with strong explanatory power and wide coverage from the variables to be selected through a machine learning LASSO model to improve the estimation results(Dettling and Kearney, 2014; Flavin and Yamashita, 2011; Jin et al., 2016; Li et al., 2019). LASSO, as a form of penalized regression, is based on the principle of imposing certain constraints on the regression coefficients to avoid the introduction of too many explanatory variables in the equation, retaining those variables that have significant influence on the explanatory variables, which helps to control the screening of variables and obtain better prediction results. Its objective function is:

Theorem 1.

$$\min_{\beta_0,\beta} \frac{i}{2N}\left(\sum_{i=1}^{N}\left(y_i-\beta_0-x_i^T\beta\right)^2+\lambda\sum_{j=1}^{k}\left|\beta_j\right|\right) \tag{1}$$

In equation (1), y_i is the dependent variable, x_i denotes the independent variables including the core explanatory variables, beta denotes the regression coefficient, k denotes the number of explanatory variables, lambda is a non-negative regularization parameter, and $\sum_{i=1}^{N}\left(y_i-\beta_0-x_i^T\beta\right)^2$ denotes the prediction error of the multiple regression results, which $\lambda\sum_{j=1}^{k}\left|\beta_j\right|$ corresponds to a penalty function excluding variables with weak explanatory power for the model in a given case so that the coefficients corresponding to these variables take the value of zero.

In this paper, 28 variables were selected for Lasso regression from both urban and individual perspectives, and 12 variables were selected from them for inclusion in the control variables after comprehensive consideration. Due to the problem of space, only the Lasso estimation results of the screening variables are reported in this paper.

(3) Descriptive statistics of variables.

Table 2 reports the results of descriptive statistics for all variables in the study. It can be seen that about 12% of the total sample of all women of reproductive age have the desire to have children, and that the average household income of two months of the migrant population can offset the price of a square meter of property, which has a large variance. The average number of family members in the sample is three, the average number of years of education is around 10 years, most of them are in agricultural households, and the average age is about 37 years.

4. Research design

(1) Baseline model setting and research assumptions.

Table 2
Descriptive statistics of variables.

Introduction to variables	Meaning	Obs	Mean	Std. Dev.	Min	Max
implicit variable	Fertility intentions	45593	.1198664	.3248082	0	1
Core explanatory variables	Relative to house prices	45593	2.115211	65.32412	0	24275.6
Individual control variables	Number of family members	45593	3.140605	1.210802	1	12
	Number of years of education	45593	10.29905	3.451954	0	19
	health status	45593	1.153993	.4195443	1	4
	Nature of household	45593	.3121184	.4633594	0	1
	Nature of work	45593	.8089934	.5557937	0	2
	age	45593	37.0503	11.20359	15	90
Urban control variables	Share of tertiary sector	45593	54.64655	11.21521	26.54	80.98
	Percentage of green space	45593	41.83624	3.467361	23	93.81
	Book collection per capita	45593	1.620107	1.532317	.0625	8.741936
	Number of primary school teachers per capita	45593	46.57175	19.34495	18.54151	163.3362
	Number of physicians per capita	45593	66.02081	22.85496	17.82246	130.3405

1. Baseline model setting

In this paper, the explanatory variable fertility intention is dichotomous and models that can be used include ordinal Probit, binary Logit, etc. The least squares estimation assumptions in the regression process are less and the results are more robust in statistical significance compared to methods such as likelihood estimation, and the OLS method is still recommended when the explanatory variable is a dummy variable or an ordered variable(Nunn and Wantchekon, 2011). In this paper, the Probit model is used as the main regression model, while the results of Linear Probability Model (LPM) are reported. The basic expression for the Probit model is.

Theorem 2.

$$\Pr\left(\mathrm{T}_{ij}=1 \mid hi, X\right)=F\left(hi, \beta_1\right)=\frac{\exp\left(\beta_0+\beta_1 hi_{ij}+\beta X+\varepsilon_{ij,t}\right)}{1+\exp\left(\beta_0+\beta_1 hi_{ij}+\beta X+\varepsilon_{ij,t}\right)} \tag{2}$$

In equation (2), T_{ij} the intention of the migrant population to have children is a binary variable. If the migrant population i chooses to give birth in the j inflow city, the value is 1, otherwise it is 0. $Pr\left(\mathrm{T}_{ij}=1 \mid hi, X\right)$ is the probability that the migrant population i has the intention to give birth in the inflow city j. F is the cumulative distribution function of the standard normal. $hi$is house price to income ratio, X are other control variables, $\varepsilon_{ij,t}$ and are random disturbance terms.

The basic expression for the LPM model is:

Theorem 3.

$$T_{ij}=\alpha+\beta hi_{ij}+X_{ij}'\gamma+\varepsilon_{ij}, \varepsilon_{ij} \sim N\left(0, \quad \sigma^2\right) \tag{3}$$

In equation (3),T_{ij} is the fertility intention of the migrant population;hi is the relative house price of mobile households, $X_{ij}'\gamma$ is the individual, urban control variables. ε_{ij} is the random disturbance term.

(2) Research hypothesis.

This paper proposes the following hypothesis on the transmission mechanism by which relative house price house prices affect the fertility intentions of the migrant population.

Hypothesis 1: The relative house prices in the cities where the mobile population is located and their fertility intentions are negatively related.

The house price in the incoming city is related to the strength of the migrant population's willingness to settle and integrate socially. House price, as a quantitative indicator of the difficulty of acquiring residential property rights, is an

important indicator of whether the migrant population can be institutionally integrated into the place of migration. In China, housing ownership is linked to household registration and children's eligibility for schooling, medical insurance and other related social benefits. Lowering the threshold of ownership is a key measure to promote stable residence in the inflow area and thus increase the likelihood of having children. High property prices imply that it is more difficult for the mobile population to integrate economically and institutionally into the inflowing region, leading to a decrease in fertility intentions.

Hypothesis 2: Mobility's fertility intentions are sensitive to relative house prices influenced by local infrastructure development.

That is, the preference of the mobile population for local property rights is to some extent due to local factors such as education and environment. When the infrastructure service conditions of the inflowing place are better, the inflowing population is more sensitive to the property price, which is also reflected in their fertility intention, and there is an interaction effect between the infrastructure service construction on the fertility intention of the mobile population and the sensitivity of the property price.

Hypothesis 3: The sensitivity of fertility intention to housing price is influenced by the affordability of housing. Housing price is a test of the financial ability of the mobile population. Housing expenses put great financial pressure on the mobile population and squeeze their ability to bear all costs arising from childbirth. The fertility intentions of the mobile population with housing pressure are more affected by housing prices, while the fertility intentions of the mobile population without housing pressure are not necessarily sensitive to housing prices.

Hypothesis one can be proved by the results of the benchmark regression. For hypothesis two, the interaction term between infrastructure variables and relative house prices is used to examine the possible moderating effect. the basic expression of the Probit model is.

Theorem 4.

$$\Pr\left(\mathrm{T}_{ij}=1 \mid hi, X\right)=F\left(hi, \beta_1\right)=\frac{\exp\left(\beta_0+\beta_1 hi_{ij}+\beta_2 \inf_{ij}+\beta_1 hi_{ij}\cdot\beta_1 \inf_{ij}+\beta X+\varepsilon_{i,t}\right)}{1+\exp\left(\beta_0+\beta_1 hi_{ij}+\beta_2 \inf_{ij}+\beta_1 hi_{ij}\cdot\beta_1 \inf_{jj}+\beta X+\varepsilon_{ij,t}\right)} \tag{4}$$

In equation (4), T_{ij} is the fertility intention of the migrant population, $Pr\left(\mathrm{T}_{ij}=1 \mid hi, X\right)$ is the probability of the migrant population i having fertility intention in the inflow city j. F is the cumulative distribution function of the standard normal. hi is the house price to income ratio, and $\inf_{ij}$ is the infrastructure variable.X is other control variables, $\varepsilon_{ij,t}$ and are random disturbance terms.

In response to hypothesis three, this paper regresses the migrant population in groups according to ownership status. In the CMDS2018 questionnaire, there is no question related to the property ownership status of the migrant population, and this paper uses household housing expenditure as a proxy variable for property ownership status - the group with zero household housing expenditure is regressed in a separate group to examine the effect of property ownership on the mechanism of action.

(3) Endogeneity issues and robustness tests.

The core explanatory variable in the model setting of this paper: fertility intention of the migrant population is influenced by a number of factors. In this paper, variables are screened by Lasso penalized regression in the variable selection stage, but it is still difficult to control all potential influencing factors. Meanwhile, high fertility intention of the migrant population in a region may raise the house price, creating a two-way causality problem and affecting the unbiased assumption of the estimation. Instrumental variables are effective methods in dealing with omitted variables and endogeneity problems due to two-way causality. Common instrumental variables in empirical studies targeting house prices include land development area, the product of long-term interest rate and land supply elasticity, and land sale price (Peng and Du, 2016; Zhang et al., 2018; Chaney et al., 2012; Waxman et al., 2020). In this paper, we choose lagged land price as the instrumental variable of house price. Land sale prices directly affect house prices, which is consistent with the instrumental variable correlation hypothesis; meanwhile, the migrant population does not actively consider solving the land sale prices of previous years when considering fertility, which is consistent with the instrumental variable exogeneity hypothesis. To conform to the data structure of relative house prices, the lagged land price to lagged disposable income per capita ratio is constructed as the instrumental variable in this paper.

As above, this paper reports regression results for both IVProbit and IVLPM. In the two-stage least squares method, the hypothesis of weak instrumental variables can in principle be rejected by analyzing the first-stage F-statistic, and to

Table 3
Baseline regression results.

	(1) OLS	(2) OLS	(3) OLS	(4) OLS	(5) OLS	(6) Probit	(7) Probit
houseincome	-0.00271***	-0.00129*	-0.00690***	-0.00269***	-0.00302**	-0.0140**	-0.0218**
	(0.000898)	(0.000674)	(0.00219)	(0.00102)	(0.00118)	(0.00573)	(0.0110)
job	-	0.00410	-	0.00262	0.00211	-	-0.0192
	-	(0.00261)	-	(0.00272)	(0.00278)	-	(0.0165)
health	-	0.00642**	-	0.00560*	0.00561*	-	-0.0249
	-	(0.00311)	-	(0.00321)	(0.00327)	-	(0.0301)
old	-	-0.0100***	-	-0.00995***	-0.00983***	-	-0.0687***
	-	(0.000193)	-	(0.000200)	(0.000202)	-	(0.00144)
family	-	-0.101***	-	-0.104***	-0.107***	-	-0.696***
	-	(0.00178)	-	(0.00189)	(0.00193)	-	(0.0182)
edu	-	0.00493***	-	0.00531***	0.00555***	-	0.0412***
	-	(0.000460)	-	(0.000490)	(0.000510)	-	(0.00344)
hukou	-	0.00408	-	0.00305	0.000380	-	0.0289
	-	(0.00310)	-	(0.00323)	(0.00382)	-	(0.0229)
green	-	-	0.000145	0.0000614	0.110***	-	0.761***
	-	-	(0.000451)	(0.000422)	(0.0324)	-	(0.263)
pbook	-	-	-0.00114	-0.00261*	-0.0999***	-	-0.691***
	-	-	(0.00152)	(0.00135)	(0.0288)	-	(0.231)
pteacher	-	-	0.000696***	0.00182***	0.0351***	-	0.235***
	-	-	(0.000179)	(0.000165)	(0.00958)	-	(0.0795)
physician	-	-	0.0000226	-0.000263***	-0.00227***	-	-0.0146**
	-	-	(0.0000951)	(0.0000872)	(0.000714)	-	(0.00594)
gdp	-	-	0.00137***	0.000353*	-0.00773***	-	-0.0536***
	-	-	(0.000214)	(0.000188)	(0.00226)	-	(0.0176)
cons/cut1	0.124***	0.759***	0.0201	0.688***	-4.640***	1.152***	34.39***
	(0.00214)	(0.0137)	(0.0208)	(0.0236)	(1.588)	(0.0120)	(13.02)
Individual control variables	-	√	-	√	√	-	√
Urban control variables	-	-	√	√	√	-	√
Fixed effects	-	-	-	-	√	-	√
N	49453	49453	45593	45593	45593	49453	45593
R^2	0.000	0.157	0.003	0.162	0.170	-	-

ensure the robustness of the conclusions, this paper also chooses the limited information maximum likelihood (LIML) method, which is more accurate in predicting weak instrumental variables, for the regression. In addition, generalized distance estimation (GMM) is relatively more effective if the nuisance terms are heteroskedastic or autocorrelated. To prevent potential nuisance terms from affecting the unbiasedness of the estimates, generalized distance estimation was added for robustness testing with iterative generalized distance estimation (IGMM).

Although the sampling method for monitoring the mobility dynamics is scientific and the sample size is large, there may still be endogeneity problems arising from "self-selection" in the sample selection. A robustness test using propensity score matching can effectively alleviate this situation. In this paper, the 1:1 nearest neighbor matching method is used for analysis, and the average treatment effects of the 1:2 nearest neighbor matching method, radius matching and kernel matching are tested for robustness.

5. Empirical process and discussion

(1) Baseline regression results.

Table 3 shows the results of the baseline regression model. Models (1) and (2) are the results of the LPM regressions. Model (1) indicates that the regression coefficient of the house price to income ratio on fertility intention is -0.00271,

Table 4
Baseline regression results.

	(8) TSLS	(9) IVProbit	(10) LIML	(11) GMM	(12) IGMM
Relative to house prices	-0.0442**	-0.222**	-0.0442**	-0.0442**	-0.0442**
	(0.0200)	(0.0965)	(0.0200)	(0.0200)	(0.0200)
Individual control variables	√	√	√	√	√
Urban control variables	√	√	√	√	√
N	45593	45593	45593	45593	45593
R^2	0.121	-	0.121	0.121	0.121
One-stage regression results			F=47.9163, p=0.0000		

which is significant at 1% level, conditional on not including control variables, indicating that there is a significant negative effect of higher house prices on the fertility intention of the migrant population, and the higher the ratio of house prices to the household income of the migrant population, the lower their fertility intention. Model (2) includes both urban and individual control variables, and the regression coefficient does not change significantly (-0.00269) compared to the regression coefficient under the condition of no control variables included.

The estimated results of the Probit model LPM are similar, i.e., there is a significant negative effect of high house prices on the fertility intentions of the migrant population. Model (3) indicates that the regression coefficient of house price to income ratio on fertility intention is -0.0140 when no other control variables are included. model (4) The regression coefficient of relative house price on fertility intention of the migrant population is relatively lower (-0.00987) and slightly less significant (5%) when individual control variables are included. Model (5) has a relatively higher effect of fertility intentions of the migrant population (-0.0344) when urban control variables are included. Model (6), which includes both urban and individual control variables, has a slightly higher regression coefficient (-0.0195) compared to the regression coefficient under the condition that no other control variables are included.

Table 4 depicts the predicted results after adding IV to the model, and the results of 2SLS in model (7) show some increase in the absolute value of the coefficients compared to the results of the LPM regression, indicating that the underlying endogeneity problem tends to underestimate the effect of house prices on the fertility intentions of the migrant population. For each female age-eligible mobile individual, each 100 percent increase in the relative house price to income ratio will reduce the intention to have children locally by 4.42 percent.

According to the predictions of model (8) IVProbit, the regression coefficient of relative house prices is -0.222, which is significant at 5% level, indicating that given constant individual and urban characteristics, the marginal effect of relative house prices on labor force fertility intentions compared to female age mobile individuals without fertility intentions is exp(-0.222) , i.e. for every 100% increase in relative house prices, each female age mobile individual will have a 22.2 percent reduction in fertility intentions.

According to model (9), the regression results for the two-stage least squares regression and the limited information maximum likelihood estimation are highly similar, again demonstrating that there is no weak instrumental variable problem in the empirical evidence. The results of models (10) and (11) using generalized distance estimation and iterative generalized moments estimation are similar to the 2SLS results as well, indicating that the heteroskedasticity problem of potential nuisance terms does not significantly interfere with the empirical study and that the estimation results are robustly keyed.

In summary, hypothesis 1: "The relative house prices in the cities to which the migrant population flows are negatively related to their fertility intentions" is confirmed.

(2) Propensity score matching

Propensity score matching can effectively mitigate the endogeneity problem caused by sample selection bias. In this paper, we use the matched sample for analysis to estimate the effect of matched relative house prices on the fertility intentions of the migrant population after passing a common support hypothesis test and a stationarity test. Table 5 shows the PSM estimation results, and the average treatment effect of all four matching methods is around -0.24, controlling for individual characteristics of the sample, which is the same conclusion as the baseline regression results:

Table 5
Propensity score matching.

Matching Status	1:1 Matching		1:2 match		radius matching		nuclear matching	
	ATT	t	ATT	t	ATT	t	ATT	t
Pre-match	-.0159***	-5.26	-.0159***	-5.26	-.0159***	-5.26	-.0159***	-5.26
Post-match	-.0241***	-4.08	-.0234***	-4.36	-0.217***	-4.25	-0.224***	-4.35

Table 6
Analysis of Moderating effects.

	(13) Probit	(14) Probit	(15) Probit	(16) Probit
hosbook	-0.00439***	-0.00851*	-	-
	(0.00130)	(0.00477)	-	-
hosgreen	-	-	-0.000263**	0.00329*
	-	-	(0.000116)	(0.00172)
Individual control variables	×	√	×	√
Urban control variables	×	√	×	√
N	45593	45593	45593	45593

i.e., there is a significant negative relationship between relative house prices and fertility intentions of the migrant population.

(3) Moderating effects of infrastructure

The effect of relative house prices on the fertility intentions of the migrant population was estimated above, but is the effect of house price to income ratio on the measures always direct? Table 6 reports the relationship between the moderating effects of education and greening-related facilities on house prices and fertility intentions in mobile inflow locations.

Models (12) and (14) put in only the interaction term of relative house prices and infrastructure public services, and (13) and (15) include urban, individual control variables. In particular, models (12) and (13) are the interactions of book collection per capita and relative house price, with regression coefficients of -0.00439 and -0.00851 without and with the inclusion of control variables, respectively, which are significant at 1% and 5% water level. Models (14) and (15) are the interactions between urban green area and relative house prices, with coefficients of -0.000263 and 0.00329 without and with control variables, respectively, significant at 5% and 10% levels. All these results suggest that there is a moderating effect of the level of infrastructure services on relative house prices. The preference of the migrant population for local property rights is to some extent due to local education and environment, and the sensitivity of the migrant population's fertility intention to relative house prices is affected by local infrastructure development, and hypothesis two is confirmed.

(4) The effect of housing stress on the mechanism of action

In the hypothesis of this paper, the difficulty of acquiring relative house price as a measure of housing ownership is negatively related to the fertility intention of the migrant population. Another intuitive manifestation of this logic is that the transmission mechanism becomes more pronounced when the household housing expenditure of the migrant population is higher. In the CMDS questionnaire, household housing expenditure mainly includes rent and mortgage. In this paper, the migrant population is grouped into regressions according to whether the housing expenditure is zero or not. Table 7 reports the regression results of multiple methods for both groups. It is clear that the group with high household housing expenditure among the migrant population is more sensitive to housing prices, and hypothesis three is confirmed.

(5) Heterogeneity analysis

Table 7
Regression in groups.

	(17) OLS	(18) OLS	(19) OLS	(20) OLS	(21) TSLS	(22) TSLS
houseincome	-0.00231**	-0.00287**	-0.00189*	-0.00308*	-0.0292*	-0.0355
	(0.00105)	(0.00122)	(0.00106)	(0.00157)	(0.0172)	(0.0310)
Individual control variables	-	-	√	√	√	√
Urban control variables	-	-	√	√	√	√
N	11889	37564	11038	34555	10973	34417
R^2	0.000	0.000	0.160	0.163	0.134	0.143

Table 8
Heterogeneity Analysis.

	(23)	(24)	(25)	(26)	(27)	(28)
	Family size		Number of years of education		age	
houseincome	-0.00507***	-0.000562	-0.00143**	-0.0189***	-0.0156***	-0.00127*
	(0.00168)	(0.000498)	(0.000704)	(0.00417)	(0.00336)	(0.000680)
N	24674	20919	36338	9255	13255	32338
R^2	0.221	0.038	0.132	0.183	0.136	0.115
Differences between groups	0.005		0.019		0.015	
P-value	0.000		0.000		0.000	

The overall estimation of the effect of relative house prices on the fertility intention of the migrant population is presented above, but the conclusions drawn in the context of a large sample size and wide coverage are slightly generalized, and there is a certain need for heterogeneity analysis based on some characteristics of the sample. Among the methods for comparing the differences in coefficients between groups in group regressions, common tests include the introduction of cross terms, the seemingly uncorrelated model test, and the Fisher's combination test. The seemingly uncorrelated model assumption is relatively loose, allowing two groups of disturbance terms to be correlated with each other, which is more consistent with the research hypothesis of this paper.

The results are reported with estimates for subgroups below the median on the left and for subgroups above the median on the right. In Table 8, models (20) and (21) report regression results based on grouping by sample household size (3 persons), models (22) and (23) report regression results based on grouping by whether or not they have higher education, and models (24) and (25) report regression results based on grouping by age (30 years). According to the results, the fertility intentions of smaller family size, higher education level, and lower age of female migrant population of school age are more sensitive to house prices. The reasons for larger family size may include supporting the elderly or already having children, in either case, there is a certain burden on the family economy, and also the fertility intention of the second child is significantly lower compared to the first child. The mobile group with high education level tends to consider settling in the inflow area, and their fertility intention will be more sensitive to the house price.

6. Conclusion

China's migrant population continues to expand in size, with fertility rates significantly lower than the level of generation replacement to achieve population stability, and the structural imbalance in human resource supply has attracted widespread attention. This paper uses OLS and Probit models to estimate the impact of house prices on the fertility intentions of the mobile population based on data from the 2018 China Migrants Dynamic Survey. The lagged land sales price is used as an instrumental variable of house price to mitigate the potential endogeneity problem. The results show that for every 100% increase in the ratio of house price to household income of the migrant population, the fertility intention of the female migrant population of working age in the inflow area will decrease by 4.42%, and the fertility intention of each female migrant individual of working age will decrease by 22.2%, i.e., the marginal effect

of relative house price on labor force fertility intention is exp(-0.222); The sensitivity of fertility intention to house price is affected by the moderating effect of infrastructure construction in the inflow area. Female mobile population with lower age, smaller family size and higher education have higher fertility intention in the inflow area. In summary, the government can enhance the fertility intentions of the migrant population by: improving the housing policy of the migrant population to reduce housing costs; addressing the institutional barriers to the education of the children of the migrant population to guarantee the right to education for children of school age; and deepening the reform of the household registration system to guarantee basic public service resources for the migrant population.

References

Nicola Gennaioli, Rafael La Porta, Florencio Lopez-de Silanes, and Andrei Shleifer. Human capital and regional development. *The Quarterly journal of economics*, 128(1):105–164, 2013.

CUFE. China human capital report 2021. *Central University of Finance and Economics Press*, 2021.

Yonghui Li, Bolan Shen, and Xiaoqin Li. How can we raise children if we cannot live in peace? –a study on the impact of housing on fertility intention of people of childbearing age. *China Economic Issues*, 02:68–81, 2021.

Jingwen Tan and Shixi Kang. Finding the instrumental variables of household registration: A discussion of the impact of china's household registration system on the citizenship of the migrant population. *arXiv preprint arXiv:2112.07268*, 2021.

NBS. China statistical yearbook 2021. *National Bureau of Statistics of China*, 2021a. URL http://www.stats.gov.cn/tjsj/tjcbw/202201/t20220112_1826282.html.

Jia Wu and Junsen Zhang. Restrictions on school enrollment of children who move with them, child retention and urban labor supply. *Economic Research Journal*, 11:138–155, 2020.

Kevin Honglin Zhang and Shunfeng SONG. Rural–urban migration and urbanization in china: Evidence from time-series and cross-section analyses. *China Economic Review*, 14(4):386–400, 2003.

NBS. Bulletin of the 7th national population census 2021. *National Bureau of Statistics of China*, 2021b. URL http://www.stats.gov.cn/tjsj/zxfb/202105/t20210510_1817183.html.

Pu Hao and Shuangshuang Tang. Floating or settling down: The effect of rural landholdings on the settlement intention of rural migrants in urban china. *Environment and Planning A*, 47(9):1979–1999, 2015.

Chunguang Wang. Study on the "semi-urbanization" of rural migrant population. *Sociological Research(Chinese)*, 5(7):107–122, 2006.

Lisa J Dettling and Melissa S Kearney. House prices and birth rates: The impact of the real estate market on the decision to have a baby. *Journal of Public Economics*, 110:82–100, 2014.

Jemima Petch and W Kim Halford. Psycho-education to enhance couples' transition to parenthood. *Clinical psychology review*, 28(7):1125–1137, 2008.

Margaret Gough. Birth spacing, human capital, and the motherhood penalty at midlife in the united states. *Demographic Research*, 37:363–416, 2017.

Jun Su Kim and Yeong Hun Yeo. The relationship between family-friendly policy and intention of childbirth. *International Journal of IT-based Social Welfare Promotion and Management*, 6:1–6, 2019.

Lisa Arai. Peer and neighbourhood influences on teenage pregnancy and fertility: Qualitative findings from research in english communities. *Health & Place*, 13(1):87–98, 2007.

Mark Lino. Expenditures on children by families. *Family Economics and Nutrition Review*, 14(2):3, 2002.

Han Entzinger and Renske Biezeveld. Benchmarking in immigrant integration. 2003.

Yang Du, Fang Cai, Xiaobo Qu, and JieChen. Continuing the chinese miracle: Reaping the dividends of the household registration system reform. *Economic Research*, 8:4–13, 2014.

William Arthur Lewis et al. Economic development with unlimited supplies of labour. 1954.

JK Banerjee. An experimental study on the cold rolling of strips. 1983.

Kenneth D Roberts. The determinants of job choice by rural labor migrants in shanghai. *China Economic Review*, 12(1):15–39, 2001.

Lina Song and Simon Appleton. Social protection and migration in china: What can protect migrants from economic uncertainty. *Migration and social protection in China*, 14:138, 2008.

Anna L Paulson and Robert Townsend. Entrepreneurship and financial constraints in thailand. *Journal of Corporate Finance*, 10(2):229–262, 2004.

Wallace E Oates. The effects of property taxes and local public spending on property values: An empirical study of tax capitalization and the tiebout hypothesis. *Journal of political economy*, 77(6):957–971, 1969.

Marjorie Flavin and Takashi Yamashita. Owner-occupied housing: life-cycle implications for the household portfolio. *American Economic Review*, 101(3):609–14, 2011.

Yongai Jin, Jian Song, and Wei Chen. Fertility preferences and fertility plans of chinese urban women in the context of comprehensive two-child policy. *Population Studies*, 40(6):22–37, 2016.

Yonghui Li, Xiaoqin Li, and Bolan Shen. Reunification only when living in peace? –a study on the promotional effect of secure housing on the mobility of family-oriented migration. *Financial Research*, 12, 2019.

Nathan Nunn and Leonard Wantchekon. The slave trade and the origins of mistrust in africa. *American Economic Review*, 101(7):3221–52, December 2011. doi: 10.1257/aer.101.7.3221. URL https://www.aeaweb.org/articles?id=10.1257/aer.101.7.3221.

Dongdong Peng and Yunsu Du. Do high housing prices inhibit firms' exports - a theoretical and empirical analysis from chinese manufacturing firms. *China Economic Issues*, (5):3–15, 2016.

Wei Zhang, Jiayun Xu, and Zhusong Yang. House price, wage and resource allocation efficiency - an empirical analysis based on micro household data. *Financial Research*, 458(8):69–84, 2018.

Thomas Chaney, David Sraer, and David Thesmar. The collateral channel: How real estate shocks affect corporate investment. *American Economic Review*, 102(6):2381–2409, 2012.

Andrew Waxman, Yuanning Liang, Shanjun Li, Panle Jia Barwick, and Meng Zhao. Tightening belts to buy a home: Consumption responses to rising housing prices in urban china. *Journal of Urban Economics*, 115:103190, 2020.