

# Research on the Relationship between House Price and High-education Migrant Population Based on PVAR Model

Gang Xia<sup>1</sup>, Hualing Li<sup>1</sup>, Dibei Cong<sup>2</sup>

<sup>1</sup>School of Management Science and Engineering, Guizhou University of Finance and Economics, Guiyang, Guizhou, China

<sup>2</sup>School of Business Administration, Guizhou University of Finance and Economics, Guiyang, Guizhou, China

**Keywords:** Housing prices; High education; Floating population; PVAR model

**Abstract:** The relationship between housing prices and highly educated floating population has caused widespread discussion, so this paper uses 2010-2016 China's dynamic population survey data and urban residential prices to match the data of 231 cities across the country, and incorporates income and industrial structure to establish a PVAR model for empirical analysis. The empirical results show that: (1) Housing prices have no significant crowding-out effect on the accumulation of highly educated floating population. (2) High-educated people flow into cities, and income is the most important influencing factor; (3) the accumulation of highly-educated populations has clearly caused house prices to rise.

## 1. Introduction

After the reform and opening up, our country experiences the large-scale population flow every year. In recent years, more and more cities have introduced policies related to talent introduction, at the same time. Talent is the most important productive force for urban development, and also a very important and indispensable factor for the development of industries, cities and the country. It is also the foundation of national prosperity. However, since 2016, the housing price has increased too much. Under the background of such a high housing price, whether it will lead to the drain of urban talents and the reduction of human capital stock? Moreover, in 2018, Huawei, a technology-intensive enterprise in Shenzhen, moved to Dongguan, which is a typical example of high-tech industry allocation and labour outflow under high housing price. How to make labour force "live" in the context of high housing price is the key problem facing the government in the process of coping with the new urbanization.

## 2. Literature review

There are many researches on housing price and labour force decision-making in academia. Whether a working family lives in a city depends on the balanced result of the city's comprehensive utility to the labour force, that is, the influence of the push and the pull, among which the housing price plays an important role. Rock (1982) believes that housing price is the main cost for labour force to live and work in the city, and directly affects the housing decision of labour force [1]. Subsequently, Helpman (1988) first introduced housing market factors on the basis of the standard model of new economic geography, and pointed out that excessively high housing prices in a certain region would affect the relative utility of laborers, thus inhibiting the agglomeration of labor forces in the region [2]. Andrew J. Plantinga (2013) studied men with college education in the United States and found that men with higher education tend to choose cities with lower housing prices [3]. Diamond (2016) believes that highly skilled labor force prefers cities with better public environment and is correspondingly willing to bear higher living costs [4].

Gao Bo (2012), Gao Bo et al. (2013) found in the panel data of 35 large and medium-sized cities in China from 2000 to 2009 that the increase of housing prices in cities would have an inhibitory effect on labour force, but this inhibitory effect was mainly reflected in rural labour, but it would also

lead to urban industrial transfer and industrial upgrading [5,6]. Jian-yong Fan et al. (2015) explained from the residential model that the high housing price did not restrain the continuous inflow of floating population, mainly because the floating population with low educational background lived in the collective dormitories of urban manufacturers or “urban villages” and other places with low housing cost [7]. Zhao fang and Yuan Chao (2016), using census micro data with macroeconomic data in the field of city, the floating population thought high and low degree of labour market, prices of highly educated population crowding out effect is not obvious, labour migration is mainly affected by the trend of wages, prices is mainly affected by the number of population to move in, the human capital structure in the Labour market impact on house prices and no significant [8]. Bin-kai Chen and Chuan-chuan Zhang (2016) explained the dramatic rise in housing prices from 2002 to 2009 from the perspective of human capital structure, and the increase in the population with higher education was an important reason for the rise in housing prices [9]. Zhang li et al. (2017) used the dynamic survey data of China’s labour force to empirically find that housing prices first attract and then restrain the floating population, showing an inverted “U” shape. They divided the labour force into two groups with different degrees of education, and concluded that the floating population with higher education is more sensitive to housing prices [10]. Ying-gang zhou et al. used the dynamic monitoring and survey data of China’s floating population from 2014 to 2016 to explain from the micro level that the crowding out effect of housing price on the floating population with high education and no house purchase in the city was particularly significant [11].

Based on the above literatures, it can be known that there is no definite point of view on the influence between housing prices and highly educated floating population, and there are few empirical analyses nationwide, mostly based on the empirical analysis at the provincial level. In order to study the relationship between the housing price and the floating population with higher education. This paper introduces a PVAR model suitable for dealing with endogenous relations for empirical analysis to ensure that the empirical results are more scientific and effective. In addition, two control variables, income and industrial structure, are added to the empirical analysis to make the model more complete.

### **3. Variable selection and model setting**

#### **3.1 Variable source and variable selection**

Urban floating population with higher education: it is measured by the proportion of urban floating population with higher education. Different scholars choose different indicators of human capital structure. Yong-jun Chen et al. use the proportion of the population with college education and above in the total population aged 6 and above to measure [12]. Guang-ping Liu measured the proportion of the population with education level in the total employed population of college graduates and above (including higher vocational colleges, college graduates, undergraduate students and graduate students), while the rest was the proportion of the population with low education level [13]. Zhang Li et al. used the data of China’s labour force dynamic survey (CLDS) to classify the floating population with high school education or above as the group with high education level, and the floating population with high school education or below as the group with low education level. Zhao fang et al. [9] classified individuals with a high school education or above as highly educated labour force, while those with a high school education or below were classified as low-educated labour force based on the micro data from the sampling survey of 1% population in China in 2000 and 2005. In this paper, according to the purpose of this study, the floating population from the national health monitoring data arrangement, the outstanding in elementary school, elementary school, junior high school, high school/secondary to low degree of population, the college, undergraduate, postgraduate and above as a highly educated population, and calculate the highly educated population proportion of the floating population in cities, for high, data from 2011-2017 dynamic monitoring data of the floating population of the country [10].

Housing price index: if the floating population wants to settle down in the destination, it should first consider buying housing. Therefore, this paper chooses urban commercial housing as the

explained variable. It is recorded as HP, and the unit is thousand yuan/square meter. The data mainly come from CEIC China economic database, and the missing data are made up according to the municipal statistics bureau.

Income index: in this paper, the average on-job salary of urban residents is chosen. Considering the influence of floating population on housing price, the increase of the average salary of on-job employees can increase residents' and floating population's expectation of buying houses in time. Therefore, the variable of average wage of on-duty workers is selected as the income indicator, which is marked as the wage unit: ten thousand yuan.

Industrial structure: the ratio of the output value of the tertiary industry (ten thousand yuan) to the output value of the secondary industry, recorded as ind. Due to the different nature of the secondary and tertiary industries, the nature of the jobs provided by them is also different, so it has an impact on the employment opportunities of the labor force.

Due to the fact that each database has a different degree of missing data on cities, 231 valid city samples are finally included after data processing. The descriptive statistical analysis of the samples is shown in Table 1.

Table 1 Descriptive statistics of variables nationwide

| Observations                                   | Number of samples | Mean  | Standard deviation | Minimum | Maximum |
|--|-------------------|-------|--------------------|---------|---------|
| Room rate (1000 yuan/m <sup>2</sup> )          | 1617              | 4.83  | 3.22               | 1.40    | 4.55    |
| Proportion of Population with higher education | 1617              | 0.107 | 0.074              | 0.000   | 0.763   |
| Salary income (yuan)                           | 1617              | 4.61  | 1.48               | 0.50    | 32.06   |
| The industrial structure                       | 1617              | 0.871 | 0.658              | 0.175   | 19.214  |

### 3.2 Model setting

In the empirical analysis, the PVAR model treats all variables as endogenous variables. Based on the commonly used PVAR model and the interaction between housing price and floating population, the following model is set:

Establish the PVAR model of housing price, the proportion of urban floating population with higher education, income and industrial structure.

$$Y_{it} = \alpha_{it} + \beta_j Y_{i,t-j} + \varphi_t + f_i + \varepsilon_{it} \quad (1)$$

Where:  $Y_{it}$  presents the endogenous variable of the  $i$  city in year  $t$ ,  $\varphi_t$  represents the time effect of the model,  $f_i$  represents the individual effect of the model,  $\varepsilon_{it}$  represents the residual term,  $i$ ,  $t$ ,  $j$  represent the number of individual, time and lag periods respectively.

## 4. Empirical analysis

### 4.1 Test of model stability and selection of optimal lag order

#### 4.1.1 Model stability test

Since only 231 cities and seven-year data are selected in this paper, they are typical wide cross-section data with short time, and can be regarded as stable data. At the same time, stata15.0 was used to test the stability of the variables, indicating that the characteristic roots of the national variables [lnhp high lnwage ind] were evenly distributed in the unit circle (as shown in Figure 1), indicating that the model met the stability test.

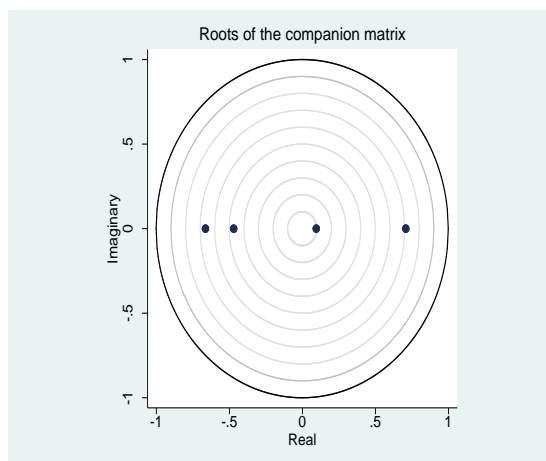


Figure 1 Nationwide adjoint matrix square root test circle

#### 4.1.2 Optimal lag order

The number of lag periods is determined by MBIC, MAIC and MQIC, smaller statistics are selected according to the rules, and the final number of lag periods is selected as the first order, as shown in Table 2.

Table 2 Selection criteria for lag order of PVAR model

| lag | MBIC   | MAIC   | MQIC   |
|-----|--------|--------|--------|
| 1   | -1.842 | -2.385 | -1.534 |
| 2   | -1.013 | -1.445 | -1.896 |
| 3   | -1.361 | -0.961 | -1.349 |

#### 4.2 Parameter estimation of PVAR model

This paper constructs a PVAR model of housing price and the proportion of urban floating population with higher education. The advantage of PVAR is that the impact of each shock on the endogenous variable can be decomposed through the orthogonal response when other variables remain unchanged. Fixed effect and the lagging behind of the dependent variable regression variables, is often used to eliminate the fixed effect coefficient of the mean difference will cause deviation, as a result, this paper USES the forward average difference method, then according to determine the order number, the lagging behind of the lag of the return of the amount as instrumental variable, PVAR GMM estimation, Table 3 lists the PVAR parameter estimation results of the model.

Table 3 PVAR model estimation in China

|                 | <i>lnhp</i>          | <i>high</i>            | <i>lnwage</i>          | <i>ind</i>            |
|-----------------|----------------------|------------------------|------------------------|-----------------------|
| <i>L.lnhp</i>   | 0.228***<br>[0.0327] | -0.0224<br>[0.0160]    | 0.351***<br>[0.0485]   | -2.363***<br>[0.2575] |
| <i>L.high</i>   | 0.201***<br>[0.0442] | -0.276***<br>[0.0427]  | 0.824***<br>[0.1951]   | -0.870***<br>[0.2122] |
| <i>L.lnwage</i> | 0.226***<br>[0.0132] | 0.156***<br>[0.0107]   | 0.389***<br>[0.0434]   | 2.510***<br>[0.1108]  |
| <i>L.ind</i>    | -0.0001<br>[0.0004]  | 0.00470***<br>[0.0004] | 0.00766***<br>[0.0011] | -0.674***<br>[0.1854] |

Note: \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ , L. 1 represents the first order of lag

From the point of price equation, lag of housing prices, lag of urban floating population accounts for higher education than the lag issue of income as is at 1% significance level, suggests that house prices, the urban population of high degree of the floating population, the urban wage income has positive effect to the price, is the main mechanism is due to the housing not only has residential properties, also has the investment properties, property prices in the market, lead people to expect the future prices will rise, so there will be a lot of investment to buy a house, the housing prices to rise.

At the same time, when the income of the floating population is higher, the payment ability for housing is stronger, and the willingness to buy housing is stronger. The demand for housing increases, resulting in the rise of housing prices. Highly educated people are more likely to get higher wages, so they have a stronger ability to pay for housing. At the same time, highly educated people have a stronger willingness to settle in cities, making it easier for them to buy houses, resulting in effective demand for housing, and thus higher housing prices.

From the perspective of the equation of the proportion of the urban floating population with higher education, the housing price has an effect on the inflow of the urban floating population at the national level, but it is not significant. Lagging of high degree of urban floating population has significantly negative effect to itself, it shows that despite the high education background of the labor force will exist positive externality of gathering, but concentration can also lead to increased competition, and many advanced talent cultivation mode in China “homogeneity” undoubtedly makes the competition more intense, the continued in recent years, the problems of college students “employment” also can be seen, is ultimately bad urban human capital structure optimization of the floating population. The optimization of urban income and industrial structure can promote the change of human capital structure of urban floating population. PVAR model regression results with Yong-jun Chen and lee (2019) on property prices, the upgrading of industrial structure and the structure of human capital based on research results in the similar PVAR model, this article USES provincial panel data, the study found that house prices and the structure coefficient of human capital was not significant, the coefficient of human capital structure on house prices significantly, lag of human capital structure has a significant negative effect on their own.

### **4.3 Impulse response**

Compared with other panel models, the PVAR model focuses more on using impulse response analysis to predict the future. In this paper, the impulse response function chart of the whole country over the lag period of 10 was obtained after 500 times of simulation by Monte Carlo method, and a 95% confidence interval was generated at the same time, as shown in Figure 2.

The impact of a standard deviation on the housing price was positive and showed a decreasing trend, and remained stable after the fifth period, indicating that the housing price was self-adjusting and self-stabilizing. To urban education accounted for at the time of lag issue influence as the negative influence, but close to zero, that rising prices for urban highly educated talent flow has a certain effect, but the effect is not strong, at the same time, the effect is not a long time, after the first period and maintain a positive influence for a long time, this is due to the high levels of education and labor tend to have higher incomes, rising house prices “crowding out” of their disposable income is relatively small, on the other hand, the low education level of the labor force because of the low level of wage income, The requirements for housing environment are also relatively low. Many will choose informal housing such as employee dormitory and urban village, so the demand for housing is relatively small.

Given the impact of one standard deviation of the proportion of urban floating population with higher education, from the perspective of the whole country, the population with higher education has a long-term positive effect on the housing price, which lasts for a long time, reaching the maximum in the first phase, and remaining stable and approaching to zero in the sixth phase or so. It maintains a first aggregation and then repulsion effect on itself.

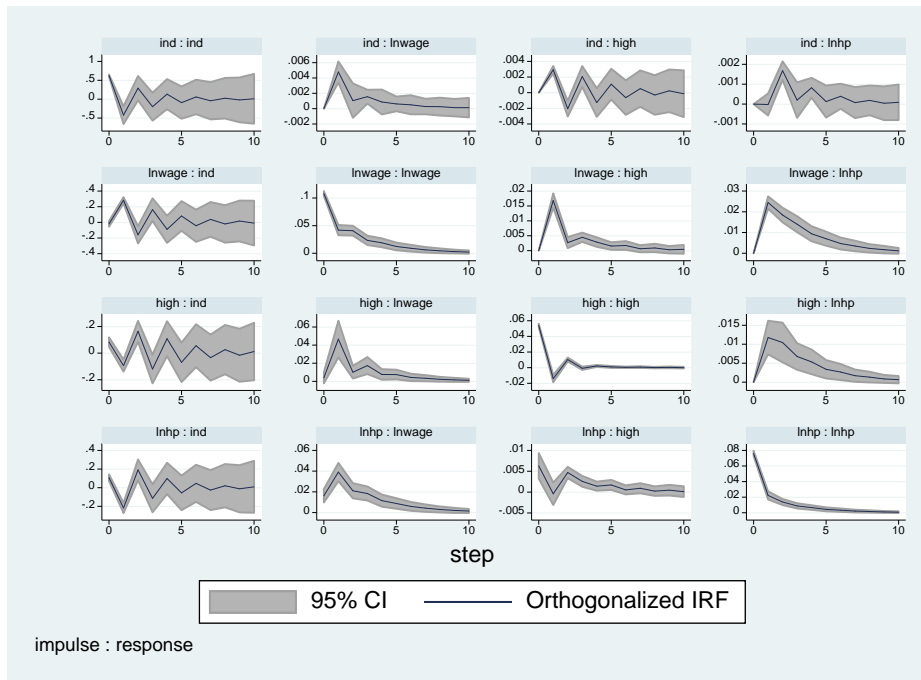


Figure 2 Pulse response function chart of the whole country

#### 4.4 Variance decomposition

In order to more accurately investigate the interaction between housing prices and the proportion of highly educated urban floating population, variance decomposition was conducted to obtain the contribution of shock response of different PVAR equations to endogenous variable volatility. The variance decomposition results of the 5th and 10th prediction periods were given in Table 4.

Table 4 Variance decomposition table of the whole country

| equation      | periods | <i>lnhp</i> | <i>high</i> | <i>lnwage</i> | <i>ind</i> |
|---------------|---------|-------------|-------------|---------------|------------|
| <i>lnhp</i>   | 5       | 0.810       | 0.032       | 0.150         | 0.0004     |
|               | 10      | 0.799       | 0.042       | 0.158         | 0.0004     |
| <i>high</i>   | 5       | 0.019       | 0.888       | 0.088         | 0.0051     |
|               | 10      | 0.021       | 0.884       | 0.089         | 0.0056     |
| <i>lnwage</i> | 5       | 0.126       | 0.124       | 0.748         | 0.0012     |
|               | 10      | 0.130       | 0.125       | 0.740         | 0.0013     |
| <i>ind</i>    | 5       | 0.114       | 0.067       | 0.135         | 0.684      |
|               | 10      | 0.115       | 0.073       | 0.140         | 0.672      |

From national level the variance decomposition of housing prices, housing price variance contribution rates at all from their own, in the fifth stage and the tenth stage their own variance contribution rate and 81.0% and 79.9% respectively, housing prices are expected to have great influences on the housing, income on the contribution rate of the housing price is greater than the urban floating population education contribution ratio of housing prices, and the stage from zero to 10, the contribution rate of income to the price rise gradually, urban industrial structure to house prices of variance accounted for less than is, can be neglected. In terms of the variance decomposition of the proportion of the floating population with higher education, the contribution rate of the floating population is the largest, with another 88.4% in the 10th period. This indicates that the aggregation of the highly educated population has a positive externality, and the larger proportion of the highly educated population in the city itself is more likely to attract the highly educated talents. The contribution rate of income to the proportion of urban floating population with higher education reached 8.9% in the 10th period, and there was no significant change from the 5th period to the 10th period. The contribution rate of housing price to the proportion of urban floating population with higher education is less than 2%, and the industrial structure has no significant contribution to the human capital structure of urban floating population. In terms of the variance

decomposition of the income equation, the contribution of income itself to the difference is particularly large. In the tenth period, there is still 74%. The contribution of the ratio of high education to income variance among the floating population in the city is not significantly different from that of the housing price. From the perspective of the industrial structure equation, the variance contribution rate of income and housing price to the industrial structure was 14.0% and 11.5% respectively in the tenth period, and the contribution rate of the highly educated population was 7.3%.

## 5. Research conclusions

This paper studies the relationship between housing prices and the proportion of floating population with higher education, establishes a PVAR model, and draws the following conclusions: (1) nationwide, housing prices have no significant impact on the optimization of the human capital structure of the floating population. From the impulse response function, housing prices have an attractive effect on the highly educated population; (2) from the perspective of PVAR model estimation results and variance decomposition, income is the most important factor affecting the inflow of highly educated people into cities; (3) the upgrading of urban industrial structure can also have an impact on the selection of the highly educated population, and the optimization of the industrial structure can also be influenced by the highly educated population in the city; (4) the concentration of highly educated population will also cause the rise of housing prices. The policy implication of this paper is that although the high housing price means the unaffordability of housing, it does not have an obvious crowding out effect on the floating population with higher education, but has a gathering effect. Labor force is one of the important factors of economic development, and the vigorous development of local economy depends on the continuous supply of high-quality labor force to some extent. Government should make reasonable housing policy, housing prices control in a reasonable scope, but also to adjust income distribution, increase the residents' income level and optimize the industrial structure of the city, the city should be through price subsidies and tax breaks or form such as "room" to provide them with more stable habitation, weaken the extrusion effect of high prices.

## Reference

- [1] R. J. Wage, "Rents the Quality of Life", *Journal of Political Economy*, vol. 90, no. 6, pp. 433-440, 1989.
- [2] E. Helpman, "The Size of Regions", In: Pines, E., Zilcha, I., and H. Kaufman, (eds.) *Topics in Public Economics: Theoretical and Applied Analysis*, Cambridge University Press, 1988.
- [3] A. J. Plantinga, C. Détang-Dessendre, G. L. Hunt, V. Piguët, "Housing prices and inter-urban migration", *Regional Science and Urban Economics*, vol. 43, no. 2, pp. 296-306, 2013.
- [4] R. Diamond, "The Determinants and Welfare Implications of US Workers Diverging Location Choices by Skill: 1980-2000", *American Economic Review*, vol. 65, no. 8, pp. 435-471, 2016.
- [5] B. Gao, J. Chen and L. H. Zhang, "Regional housing price differences, labor mobility and industrial upgrading", *Economic research*, vol. 1, pp. 66-79, 2012.
- [6] B. Gao, W. L. Wang and X. Li, "The riddle of "scissors gap" between expectation, income gap and housing price and rent in Chinese cities", *Economic research*, vol. 6, pp. 101-113+127, 2013.
- [7] J. Y. Fan, J. W. Mo, J. P. Zhang, "Human capital and housing prices in Chinese cities", *China social sciences*, vol. 4, pp. 131-142, 2016.
- [8] B. K. Zhang, C. C. Zhang, "Human capital and housing prices in Chinese cities", *China social sciences*, vol. 5, pp. 43-64, 2016.
- [9] L. Zhang J. He, R. H. Ma, "How do housing prices affect labor mobility?" *Journal of economic*

research, vol. 8, pp. 155-170, 2017.

[10]F. Zhao, C. W. Yuan, “China’s urbanization development -- research based on spatial equilibrium model”, *Economics (quarterly)*, vol. 16, no. 4, pp. 411-436, 2016.

[11]Y. Zhou, L. N. Meng, Q. Lu, “Who are high housing prices crowding out? – from the micro perspective of China’s floating population”, *Economic research*, vol. 9, pp. 108-124, 2019.

[12]Y. J. Chen, H. H. Li, “Housing price, industrial structure upgrade and human capital structure -- dynamic analysis based on PVAR model”, *Industrial technology economy*, vol. 38, pp. 51-58, 2019.

[13]G. P. Liu, Y. Y. Lian, Y. M. Tian, “Crowding out effect or agglomeration effect? – a study on the influence of housing price on human capital and regional differences”, *Journal of Yunnan university of finance and economics*, vol. 11, pp. 28-36, 2018.