

# Research on Relationship Between Population Structure and Real Estate Investment with Big Data

Wenjing Xiang\*, Gang Xia\*

School of Business Administration, Guizhou University of Finance and Economics, Guiyang, China

## Email address:

987748368@qq.com (Wenjing Xiang), 1142678284@qq.com (Gang Xia)

\*Corresponding author

## To cite this article:

Wenjing Xiang, Gang Xia. Research on Relationship Between Population Structure and Real Estate Investment with Big Data. *Science Journal of Business and Management*. Vol. 11, No. 1, 2023, pp. 12-17. doi: 10.11648/j.sjbm.20231101.13

**Received:** January 13, 2023; **Accepted:** February 2, 2023; **Published:** February 14, 2023

---

**Abstract:** With the rapid development of "Internet +", the current supply and demand of real estate have changed. The introduction of big data technology to analyze the demographic structure and real estate market-related information improves the scientific nature of real estate investment decisions. Scientifically defining the guiding role of population structure on real estate development investment is helpful to promote the coordinated development of real estate industry. Based on this, we build a theoretical model of the real estate development investment data analysis support system by selecting the panel data of each province from 2002 to 2018 and using the econometric model. The impact of population structure is tested on real estate investment from the national level and sub-regional level. The specific effect of population structure is also analyzed on investment under the urbanization rate. The research result shows that the total dependency ratio of the population and the youth dependency ratio is negatively correlated with real estate development investment at the national level. The elderly dependency ratio is positively correlated with real estate development investment. At the regional level, the elderly dependency ratio is only in the central region. As a positive impact, the juvenile dependency ratio affects real estate development investment, but the degree of influence is different. From the perspective of the urbanization process, the higher the urbanization rate, the less binding the juvenile dependency ratio on real estate development investment. In view of this, from the perspective of population structure, suggestions are put forward on the supply-side structural reform of the real estate industry.

**Keywords:** Population Structure, Real Estate Development Investment, Big Data

---

## 1. Introduction

In the context of the big data, researches on population structure and real estate development investment is largely shown in three respects below. To begin with, many scholars explore the role of big data inside the real estate sales as well as the operation. Scholars such as Ma Muyuan applied big data technology to the housing management system to manage the quantity, files, and configuration of real estate [1]. Luo Yuxia analyzed the shortcomings of traditional real estate marketing models and proposed the relationship between real estate enterprises and big data [2]. The second is the research on the big data concerning the real estate supervision and decision-making. Scholars such as Shen Zhigang proposed to build one real estate marketplace supervision mechanism through data mining, for the sake of realizing early warning and

forecast of the marketplace concerning the real estate, medium, macro and micro dissection, and to visualize the results [3]. Scholars such as Lou Liming put forward the detection idea of the whole process of the marketplace concerning the real estate, focusing on improving the data sharing mechanism, quantifying indicators, and establishing mathematical models [4]. Third, few scholars explore the correlation between population structure and real estate investment. Jin Jianfeng studied the demand for real estate investment [5], and Zhang Hong conducted an empirical study on the factors affecting the changes in the spatial agglomeration of housing advancement investing within China. Consequences manifest that population urbanization exerts an essential impact on the changes in spatial agglomeration level of the real estate investing [6]. Yang Yingying and other scholars studied the relationship between urbanization, population structure and housing supply structure transformation based on statistical data, and put forward

rationalization countermeasures and suggestions to make balanced development between the three [7]. On the basis of constructing the mechanism of the influence of aging on the housing need, Zhu Linlin, along with other scholars used the spatial econometric model and panel threshold regression model to empirically research the influence of aging on the housing need. Also, they concluded that the housing price, family size and elderly dependency ratio were passively linked to the housing sales as well as housing sales area, that is, the higher the house price, the smaller the number of households, and the deeper the degree of aging, all of which will reduce the purchase of residential housing [8]. Long Teng et found that population urbanization in large and medium-sized cities has effectively driven housing demand in the past decade, and has performed in terms of floating population, family size, population urbanization rate, and education level, and promoted housing price increases [9]. Li Dan et found that the size of families is becoming smaller and smaller, and the type of family is becoming more and more simplified; The number of households is also gradually increasing, and the demand for housing is strong in the next decade; The demand for rental private housing and affordable housing is increasing; Household demographics have a significant impact on the type of house purchased, the size of the house purchased, and the type of house [10]. Deng Hongqian's empirical research consequences reflect that the ratio of the dependency concerning children and the elderly population have a significant negative correlation to the housing prices. Also, the population turnover rate is significantly actively linked to the housing prices [11]. Yin Luyang investigates the correlation between the population structure and the housing need, the larger the proportion of children and the percentage of the elderly inside the population, the increase in average household size reflects the increase of the child population, the higher the education level, the higher the unmarried rate and the urbanization rate, the greater the housing demand [12]. Gao yuan research shows that the influence of the old-age dependency ratio on the housing prices is positively correlated. Also, the per capita disposable income and urbanization rate of the residents have an essential influence upon housing prices [13]. Jin wei uses the Hansen non-dynamic panel regression model to explore the correlation between the population structure and urban housing prices during the economic

transformation. These results show that the active effect of the population age structure upon the housing prices shows an "inverted U" trend [14]. Shi Xiaohao analyzed that changes in population structure are important factors affecting housing demand, and population urbanization, population aging, and second-child birth policies will have a significant impact on urban housing demand [15].

Most of the existing literature is about the research on big data technology and real estate sales, supervision and operation systems, but there are few studies on real estate development investment. Little literature includes population structure in the empirical research of real estate development investment. Few scholars extend from the national level to different regional levels to study the distribution effect of the population structure upon the real estate advancement investment. For the new normal of economic growth, there are few studies on the correlation between population structure and the real estate advancement investing, and few scholars have considered the intermediary effect of big data technology. Therefore, using 31 provincial-level data in my country from 2002 to 2018, we estimated the model from three aspects. First, based on national-level data to test whether the population structure (total dependency ratio, old dependency ratio, and juvenile dependency ratio) has any impact on real estate advancement investing. The second is just to further analyze whether the population structure plays a certain part in regional distribution of the real estate investing based on the regional level. The third is to explore whether the allocation of the real estate investing is different from the population structure under the background of different urbanization rates.

## 2. Theoretical Model, Variable Selection, Data Description and Model Setting

### 2.1. Theoretical Model

The population is important to determine the real estate marketplace demand as well as the real estate developers' investing. Big data technology reflects the intermediary effect in real estate development companies' investment decisions. The theoretical pattern of this article is depicted in "Figure 1".

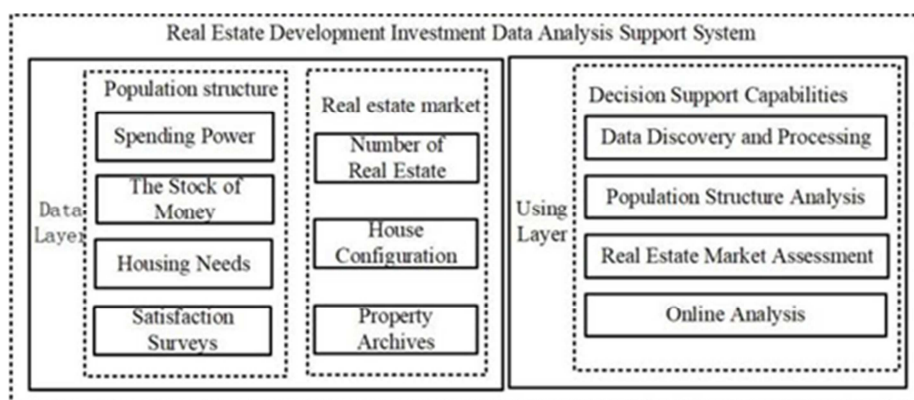


Figure 1. Theoretical model.

The real estate development investment data support system is divided into data layer and application layer. The data layer includes two sections: population structure and real estate market. The population structure section includes spending power, currency storage, housing demand, and consumer satisfaction. The real estate marketplace includes the number of houses, house configuration, as well as property files. The layer of the application is for the capability dissection of decision support. Through the data capture and processing of demographic structure and real estate market, online analysis reports are generated to support real estate investment decisions.

## 2.2. Variable Selection and Data Description

The research data are collected from the National Bureau of Statistics, and all variables are processed in logarithm. Variable selection is as follows.

Real estate development investment ( $hi$ ) stands for the explained variable and expressed by the factual real estate advancement investing in every province. Total dependency ratio ( $tdr$ ) means the rate of non-working-age population to working-age population in whole population, indicating how many non-working-age populations are borne by every 100 working-age people. Old-age dependency ratio ( $odr$ ) indicates the rate of the elderly people to working-age people in the entire population, indicating how many elderly people are borne by every 100 working-age population. Juvenile dependency ratio ( $ydr$ ) refers to the rate of the quantity of the children and children in the total deduction to the quantity of the working-age people, indicating how many children and children are borne by every 100 working-age population. We selected control variables which might impact the behavior of the real estate advancement and investing spatial distribution distinctions as below.

Economic development level ( $gprgdp$ ): The difference in the level of economic development among provinces is measured by the average annual increase ratio of real GDP every capita.

Industrial structure: The industrial structure is gauged by the proportion of the extra value of the tertiary industry in the whole GDP of per province.

Population scale is shown as the permanent resident population of each province at the end of the year. To some degree, it reflects the impact on real estate development investment based on the demand level.

The rate of the urbanization is shown by the proportion of the whole urban population to the whole population of each province. It reflects the process of the transformation of agricultural population into non-agricultural population, as well as the size of the city and the level of infrastructure.

## 2.3. Model Setting

By virtue of the distinctions in the panel figures, a fixed impact model is built for analyzing. Based on the above discussion, the following measurement model is proposed.

$$\ln hi_{it} = \beta_0 + \beta_1 \ln tdr_{it} + \beta_2 \ln odr_{it} + \beta_3 \ln ydr_{it} + \theta \ln X_{it} + \varepsilon_{it}$$

$$\ln hi_{it} = \beta_0 + \beta_1 \ln tdr_{it} + \beta_2 \ln odr_{it} + \beta_3 \ln ydr_{it} + \theta \ln X_{it} + \varepsilon_{it} \quad (1)$$

The subscripts  $i$  and  $t$  represent the province and time, respectively, and  $\varepsilon_{it}$  represents the random disturbance term.  $X$  refers to a suit of the control variables which might impact the real estate advancement and investing behavior, such as economic development level ( $gprgdp$ ), industrial structure (structure), population size (scale), and urbanization rate (urbanrate).

## 3. Estimation Results and Analysis

### 3.1. Estimated Results and Analysis of the Impact of Total Dependency Ratio on Real Estate Investment

Table 1 shows the influence of the whole dependency rate of the population on real estate advancement investing based on estimates at the level of the nation.

**Table 1.** Impact of the Total Dependency Ratio of the Population on Real Estate Development Investment (Based on the Nation Level).

| Variable        | Model 1              | Model 2              | Model 3                 | Model 4                 |
|-----------------|----------------------|----------------------|-------------------------|-------------------------|
| Intdr           | -4.632***<br>(-9.70) | -4.277***<br>(-9.57) | -1.578***<br>(-6.603)   | -0.349<br>(-1.340)      |
| Inodr           |                      |                      |                         |                         |
| Inydr           |                      |                      |                         |                         |
| Ingprgdp        |                      |                      | 0.011<br>(0.434)        | -0.061*<br>(-1.786)     |
| Instructure     |                      |                      | -0.022<br>(-0.097)      | 0.281<br>(1.119)        |
| Inurbanrate     |                      |                      | 5.563***<br>(36.181)    | 4.595***<br>(24.953)    |
| Inscale         |                      |                      | 3.255***<br>(8.856)     | 1.113***<br>(14.782)    |
| constant        | 23.418***<br>(13.60) | 22.138***<br>(13.62) | -35.467***<br>(-11.320) | -19.798***<br>(-12.572) |
| N               | 527                  | 527                  | 517                     | 517                     |
| R <sup>2</sup>  | 0.160                |                      | 0.851                   |                         |
| Inspection type | FE                   | RE                   | FE                      | RE                      |

\* T-statistics in parentheses, \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$  (Same below)

It is shown in Models 1 and 2 that the coefficients of the entire population dependency rate variables are all essential at the level of 1%, and are negatively correlated with real estate development investment. For every 1% grows in the total population dependency rate, real estate development investment decreases by about 4.6%.

The control variables affecting real estate advancement investing are added to the trial of the fixed effect model. The result shows that the goodness of fit reaches 0.85. The total dependency ratio of the population is still significantly negatively linked to the real estate development investing at the level of 1%. This shows that the growth in the total population dependency rate brings about a decrease in level

of social purchasing power and a decrease in housing demand, thereby weakening the investment intensity and scale of real estate development. The impact of industrial structure and economic development level on real estate development investing isn't significant. It might be caused by the fact that the upgrading of the structure of the industry has little change in the housing scale, and the short-term fluctuation of economic development level is small. Urbanization rate and population size are both significantly positively correlated with the real estate investing at the 1% level. For a 1% growth in urbanization rate, the real estate advancement investing increased by 5.6%, which reflects the relationship between new urbanization construction and housing demand.

### 3.2. Estimated Results and Analysis of the Impact of Juvenile Dependency Ratio and Old Age Dependency Ratio on Real Estate Investment

Table 2 shows the impact of juvenile dependency ratio and old age dependency ratio on real estate investment.

**Table 2.** Impact of Juvenile Dependency Ratio and OLD Age Dependency Ratio on Real Estate Development Investment (Based on the National Level).

| Variable        | Model 1                | Model 2                | Model 3                 | Model 4                 |
|-----------------|------------------------|------------------------|-------------------------|-------------------------|
| lnodr           | 2.674***<br>(11.65)    | 2.777***<br>(11.902)   | 0.563***<br>(3.23)      | 0.560***<br>(2.7)       |
| lnydr           | -5.111***<br>(-20.971) | -4.320***<br>(-18.669) | -1.919***<br>(-10.758)  | -0.606***<br>(-2.985)   |
| lngrgdp         |                        |                        | 0.042*<br>(1.732)       | -0.048<br>(-1.434)      |
| lnstructure     |                        |                        | -0.202<br>(-0.947)      | 0.157<br>(0.633)        |
| lnurbanrate     |                        |                        | 3.574***<br>(10.319)    | 1.043***<br>(13.054)    |
| lnscale         |                        |                        | 4.649***<br>(25.374)    | 4.213***<br>(18.808)    |
| constant        | 16.093***<br>(15.093)  | 13.336***<br>(12.614)  | -34.950***<br>(-12.240) | -18.072***<br>(-11.317) |
| N               | 527                    | 527                    | 517                     | 517                     |
| R <sup>2</sup>  | 0.596                  |                        | 0.871                   |                         |
| Inspection type | FE                     | RE                     | FE                      | RE                      |

The model regression consequences reflect that the ratio of the juvenile dependency yields an essential passive relationship over the real estate development investing. For every 1% growth in the ratio of the juvenile dependency, the real estate advancement investing decreases by nearly 2%. On the one hand, the growth in the ratio of the juvenile dependency might bring about a relative decrease in labor force, thus affecting the social demand for housing. For another, the increase in the ratio of the juvenile dependency makes the proportion of children and adolescents larger, and the corresponding parents of the children and adolescents may be positive. In the age group of wealth accumulation, the purchasing power of housing is limited, the demand decreases, and the investing in the real estate advancement decreases accordingly. The relationship between the real estate development investing and old-age dependency rate has diverged, showing an essential positive correlation. For

every 1% growth in the ratio of the old-age dependency, real estate investment increases by about 0.56%. Judging from estimated coefficients, the sensitivity of juvenile dependency rate to the real estate development investing is far higher compared with that of the old age dependency rate. Thus, these ratios of the juvenile dependency and the whole population dependency have the same influence upon the real estate investing. In addition, the level of economic development and industrial structure are not significant for the investment of the real estate. Level of urbanization rate and population size both significantly affect the investment of the real estate at the level of 1 percent.

### 3.3. Estimated Results and Analysis by Region

Considering the uneven level of economic development in the eastern, central and western regions of China and the differences in real estate investment, it is necessary to divide the area to study the impact of population structure on real estate investment in different regions. Table 3 presents the estimation results based on the regional level.

**Table 3.** Regional Impact of Demographic Structure on Real Estate Development Investment.

| Variable       | East                    | Central               | West                   |
|----------------|-------------------------|-----------------------|------------------------|
| lnodr          | 0.089<br>(0.412)        | 0.871***<br>(3.206)   | 0.465<br>(1.226)       |
| lnydr          | -1.290***<br>(-5.031)   | -1.905***<br>(-8.493) | -2.494***<br>(-6.710)  |
| lngrgdp        | 0.142***<br>(2.983)     | -0.009<br>(-0.304)    | 0.035<br>(0.922)       |
| lnstructure    | 0.815*<br>(1.911)       | -0.867***<br>(-3.387) | -0.436<br>(-1.153)     |
| lnurbanrate    | 6.391***<br>(15.197)    | 5.374***<br>(22.689)  | 3.657***<br>(11.863)   |
| lnscale        | 2.239***<br>(5.737)     | -0.205<br>(-0.166)    | 5.473***<br>(6.608)    |
| constant       | -37.143***<br>(-11.539) | -5.024<br>(-0.490)    | -40.867***<br>(-6.046) |
| N              | 183                     | 133                   | 201                    |
| R <sup>2</sup> | 0.893                   | 0.946                 | 0.858                  |

The estimation results exhibit there exists evident regional distinctions in the influence of the rate concerning the old-age dependency upon real estate development investment. The ratio of the elderly dependency exerts an essential active effect upon the real estate development and investment in the central area, while it exerts no pivotal influence within eastern and western areas. The impact of the juvenile dependency ratio upon the spatial distribution of real estate development investing among disparate areas is essential at the level of 1%, whereas the extent of influence is disparate. The estimated coefficient shows that the juvenile dependency ratio has the most prominent impact on real estate investment in western areas, followed by the central area, together with the eastern area. Due to the overall relative lag in the economic advancement level, public infrastructure, along with the employment opportunities inside the western area, numerous young and middle-aged labor force outflows, especially the total number of children, which affects the juvenile dependency ratio. Children's parents go out to work

to obtain higher wages, so the demand for housing in the outflow areas decreases, which further reduces the investing of the real estate advancement. For the sake of control variables, the level of the economic advancement yields an essential significant impact over the real estate advancement investing just within eastern areas. The structure of the industry has a significant impact on real estate development investing in central and eastern regions. Besides, the population size yields an essential impact over the real estate investing within eastern and western areas, along with in the central area. The impact of urbanization rates in different regions on the real estate advancement investing is significant at the level of 1 percent, and the sensitivity is the strongest in the eastern region and the weakest in the western region.

### 3.4. Impact of Population Structure on Real Estate Development Investment Under Different Urbanization Stages

According to the international urbanization stage division standard, the national sample data is divided into two parts with 60% urbanization as the cut-off point and estimated separately. Table 4 presents the estimated results under different urbanization rates.

**Table 4.** Estimation Results of Different Urbanization Rates.

| Variable       | urbanbrate>60%        | urbanbrate≤60%        |
|----------------|-----------------------|-----------------------|
| lnodr          | 0.191<br>(0.94)       | 0.327<br>(1.41)       |
| lnydr          | -0.711**<br>(-2.03)   | -2.142***<br>(-9.76)  |
| lnprgdp        | 0.109***<br>(2.69)    | 0.006<br>(0.18)       |
| lnstructure    | 0.668<br>(1.28)       | -0.121<br>(-0.49)     |
| lnurbanrate    | 3.724***<br>(3.65)    | 4.371***<br>(22.07)   |
| lnscale        | 2.976***<br>(8.13)    | 5.685***<br>(8.65)    |
| constant       | -33.567***<br>(-8.62) | -49.450***<br>(-9.23) |
| N              | 107                   | 406                   |
| R <sup>2</sup> | 0.826                 | 0.873                 |

The estimation results show that the impact of the juvenile dependency ratio on real estate development investment has significant differences under different urbanization rates. The coefficient is only -0.71. In the process of low urbanization, the juvenile dependency ratio significantly affects real estate development investment at the level of 1%, and the impact coefficient is -2.14. Under the high urbanization rate, the per capita income, total social wealth, and employment opportunities are far better than those in the low urbanization development stage, and the population outflow is relatively small. Therefore, the change of the juvenile dependency ratio has relatively little impact on the housing demand, and has a relatively small impact on real estate development. Investment constraints are relatively weak. Under the low urbanization rate, the increase of the juvenile dependency

ratio means that the labor force bears a heavy responsibility. In order to obtain more employment opportunities and higher wages, the proportion of population outflow increases, which has a greater impact on the real estate market. This has led developers to significantly reduce investment in real estate development in areas with low urbanization rates.

## 4. Conclusions and Recommendations

A theoretical model of real estate development and investment data analysis support system is constructed to obtain relevant information on population structure and the real estate market through data mining. Empirical research is used to demonstrate these consequences. At the level of the nation, ratios of the total population dependency and youth dependency are significantly passively linked to the real estate development investment. The ratio of the old-age dependency is apparently actively connected to the real estate development investment. From a sub-regional level, the ratio of the old-age dependency yields an essential active influence over the real estate advancement investing just inside the central region. The juvenile dependency ratio in different regions has a significant influence upon spatial distribution of the real estate advancement investing, but the degree of influence varies. If the level of economic advancement is higher, the juvenile dependency ratio on the real estate development investing will be less binding. This shows the labor force in western regions goes to eastern and central regions in pursuit of a higher standard of living, which directly affects the housing demand and investment scale in the western region. In different stages of urbanization, the influence of population structure on real estate development investment is different. Under high urbanization rate, the negative impact of juvenile dependency ratio on real estate development investment is weak, indicating that the level of urbanization has improved and the intensity of population mobility has weakened. Real estate development investment does not fluctuate much but at a lower urbanization rate, population mobility increases to have a greater impact on real estate development investment.

Based on the research conclusions, we put forward suggestions on the supply-side structural reform of the real estate industry. First, the government and developers need to pay attention to the use of big data technology in the real estate market and increase the reliability and validity of decision-making through the analysis of the relationship between population structure and real estate development investment. It is also necessary to promote the sustainable development of the real estate market. Secondly, the construction of the western region needs to be strengthened to promote the balanced and coordinated development of various regions and reduce the intensity of population flow. Finally, differentiated policies must be formulated according to the urbanization construction process to introduce agricultural transfer population according to local conditions.

---

## References

- [1] Ma Muyuan, Xu Risheng, Xu Yafei. Housing management system based on big data technology [J]. Electronic Testing, 2022, 36 (08): 85-87. (In Chinese).
- [2] Luo Yuxia. The application of big data in real estate marketing [J]. Modern Marketing, 2021 (11): 70-71. (In Chinese).
- [3] Shen Zhigang, Chen Hui, Yue Qianqian. Design and implementation of real estate market monitoring system based on big data [J]. China Real Estate, 2022 (01): 50-57. (In Chinese).
- [4] Lou Liming, Lu Shachuan, Yu Jianzhong. Whole process monitoring of real estate land projects based on big data [J]. China Land, 2021 (03): 46-48. (In Chinese).
- [5] Jin Jianfeng, Yang Tuquan, Zhang Yunqiao. Research on the impact of population aging on investment demand in the real estate market [J]. Jiangsu Science and Technology Information, 2013 (13): 7-8. (In Chinese).
- [6] Zhang Hong, Yu Bo, Ju Lixin. Research on the spatial aggregation changes and influencing factors of housing development investment in China [J]. Operations Research and Management, 2016, 25 (02): 197-202. (In Chinese).
- [7] Yang Yingying, Lin Wenhui, Wei Dongying. Research on urbanization, population structure and housing supply structure transformation [J]. Modern Business, 2019 (28): 29-30. (In Chinese).
- [8] Zhu Linlin, Li Xiuting, Dong Jichang. Research on influencing factors of housing demand under the background of population aging: Based on the analysis of population structure, housing price and other factors [J]. Price Theory & Practice, 2019 (06): 95-98. (In Chinese).
- [9] Long Teng, Wan Yong. Research on the influence of population structure on housing price index in large and medium-sized cities [J]. Construction Economics, 2022, 43 (06): 83-90. (In Chinese).
- [10] Li Zhou, Wei Kai, SONG Yuchen, Zhang Junyu. Research on the influence of household population structure on housing demand in Liuzhou City [J]. Business Watch, 2022 (35): 55-61. (In Chinese).
- [11] Deng Hongqian, Zhang Xue. The influence of population structure on housing price: Based on the two dimensions of dependency burden and floating population [J]. Jiangnan Forum, 2021 (02): 12-20. (In Chinese).
- [12] Yin Luyang, MENG Fanyu. Research on the impact of demographic change on housing demand: A case study of Hunan Province [J]. Shanghai Real Estate, 2022 (01): 29-34. (In Chinese).
- [13] Gao Yuan, Tian Guofeng, Guo Yanfang. Research on the influence of demographic factors on urban commodity housing prices: A case study of Cangzhou City [J]. Business Economics, 2021 (04): 34-35+166. (In Chinese).
- [14] Jin Wei. Research on the impact of China's population structure on urban housing prices——Based on the analysis under the background of high-quality economic development in China [J]. Price Theory & Practice, 2020 (04): 72-75. (In Chinese).
- [15] Shi Xiaohao, Sun Lihua, Guo Bing, Wang Xianzhi. The impact of population structure change on urban housing demand and planning countermeasures——Case study of Shandong Province [J]. Urban Architecture, 2019, 16 (09): 30-31. (In Chinese).