An analysis of the impact of marriage on housing prices

Zhongfu Qin[1, a], Biao Gao[2, b]

[1]School of Architectural Engineering, Zhejiang University, Hangzhou 310000, China;

[2]School of Architectural Engineering, Zhejiang University, Hangzhou 310000, China.

[a]qinzhongfu@zju.edu.com, [b]gaobiao0913@126.com

Keywords: House price; Marriage; Panel data; Regression analysis

Abstract: Based on the empirical research of multiple linear regression model, this paper analyzes the effect of marriage on the housing price level. The results show that in underdeveloped cities, the impact of marriage on housing prices is more obvious; while in developed cities, the impact of marriage on housing prices is not significant.

1. Introduction

China’s real estate industry implemented market-oriented reforms in 1998, which played a huge role in releasing residents’ housing needs and stimulating the development of the real estate industry. From 1999 to 2020, the average annual growth rate of house prices in China's 31 administrative regions reached 8.0%. Especially in developed provinces and cities along the eastern coast, the rise in house prices was even more dramatic. The soaring housing prices have brought many economic and social problems, including a large number of people's dissatisfaction with the payment difficulties caused by the high housing prices, and the social concerns about China's economic bubble and financial system risks, etc., which makes the housing price level become One of the hottest issues in China.

With the rapid growth of housing prices, China is experiencing huge demographic changes. In the past ten years or so, the residents at the peak of births from 1981 to 1990 have entered the first marriage period, but the idea of "putting men before women" left over from ancient times has made the number of men much higher than women. Wang Meixia [1] pointed out that the serious imbalance between men and women has caused serious marriage squeeze in the marriage market. The competition in the marriage market is extremely fierce. For women’s families to have a life after marriage, they usually hope that the man will have a house before marriage, prompting many single men. In order to enhance competitiveness, go to buy real estate. Xiao Wu [2] based on a nationwide survey of 4,739 samples found that young people’s views on marriage show differences, and women’s families pay more attention to each other’s economic situation. More than 40% of people require a room before they can get married, and the proportion is The performance among urban youth is particularly prominent. Lian Si and Zhao Jinyan [3] pointed out that while housing affects the economic and social status of individuals, it objectively determines the position of young people in the marriage market and restricts young people’s marriage decisions. In marriage and love, housing has become an important indicator to measure each other's economic conditions and family status. The same stratum enhances the cohesion and identity exclusion of their strata through the "door-to-household" marriage marked by "housing". The study of Liu Xueliang [4] found that owning property rights housing will significantly improve the marriage happiness of residents, especially women. Putting housing under the framework of marriage competition will make its role more prominent. In the context of asymmetric information in the marriage market, being able to afford decent housing is an important signal for people’s financial resources, making them dominant in the competition in the marriage market. . In terms of foreign research, C. Brown and J. Madge [5] believe that housing is an important factor influencing residents' marriage time. For couples who are about to buy a house and get married, the increase in house prices will increase the cost of marriage,

thereby reducing their willingness to get married. The Australian Bureau of Statistics [6] published a 1999 statistics on the first time Australian young people bought a house. Australian young people tended to marry later and execute house purchase decisions later. According to the survey, in addition to the prolonged education of young people, it also includes the need for young people to work to accumulate housing purchase funds.

The above research shows that housing is an important consideration for many people to make a marriage decision. In other words, marriage decision will increase the purchase demand of housing to a certain extent, and then affect the housing price level through supply and demand. Although previous studies have considered the relationship between marriage and housing prices, few have quantified the factors of marriage to analyze their specific impact on housing prices. Therefore, this article attempts to use the multiple linear regression model based on the city's annual panel data to verify the correlation between marriage and housing prices, and to explore the specific degree and law of its impact.

2. Model Construction of the Effect of Marriage on Housing Price Level

2.1 Overview of Multiple Linear Regression Models.

The multiple regression analysis method is usually used to study the relationship between a dependent variable dependent on multiple independent variables. If the relationship between the two can be described in a linear form, a multiple linear model can be established for analysis to reveal the relationship between these parameters and output Quantity relationship. The formula of the multiple regression model is shown in Equation 1:

$$\hat{y} = a_0 + a_1X_1 + a_2X_2 + a_3X_3 + \cdots + a_kX_k + u \quad (1)$$

In the formula, $\hat{y}$ represents a random dependent variable, $X_i(i = 1,2,\dots,k)$ represents a non-random independent variable, $a_i(i = 1,2,\dots,k)$ represents the regression coefficient of the regression model; u Represents the random interference term of the equation.

During the establishment of multiple regression models, the selection of indicators is an important factor related to the precision of the model and the credibility of the predicted trend. Usually, the relevant indicators can be screened before and during regression. The screening before regression can be performed by qualitative analysis. Discuss.

2.2 Choice of Influencing Factors

At this stage, there have been many studies on the influencing factors of housing price levels. This article has collected and sorted out several influencing factors of housing prices by reading a lot of literature, as shown in Table 1:

Table 1 Influencing Factors of Housing Price Level

Main Target	Influencing Factors	Source
Influencing Factors of Housing Price Level	GDP	Pengfei Wang[7], Ke Chen[8]
	Floor Space Sold	Pengfei Wang[7], Zhongxuan Yang[9]
	GDP Per Capita	Yonghong Xu[10]
	Wage	Ke Chen[8]], Zhuohang Li[11], Yonggang Li[12]
	Residents' Disposable Income	Zhongxuan Yang[9]
	Residence Income	Xiaohan Zhou[13]
	Annual Registered Population	Yuankun Li[14]
	Population	Jichao Zheng[15]

The above-mentioned influencing factors are from different literatures, and it is inevitable that

there will be duplication or crossover. Based on this, this paper adjusts the above influencing factors based on the system analysis method, as shown in Table 2.

Table 2 Screening Table of Influencing Factors of Housing Price Level

Main Target	Influencing Factors before Adjustment	Influencing Factors after Adjustment	Adjustment Method	Brief Explanation of Adjustment Reasons
Influencing Factors of Housing Price Level	GDP	GDP	Merge	The per capita GDP is determined by the regional GDP and the number of population. This study divides it into total population and GDP at the end of the year.
	GDP Per Capita			
	Housing Sales Area	Housing Sales Area	Keep	Important factors
	Population	Total Population at the End of the Year	Merge	The three influencing factors are all related to the population, and the total population at the end of the year can best reflect the main activities and consumer groups of the city throughout the year
	Annual Registered Population			
	Wage	Average Wage Level	Merge	The income of residents is largely related to wage levels, and the use of average wage levels can effectively characterize this part of the influencing factors
	Residents' disposable income			
	residence income			
	GDP Per Capita			
	Real Estate Investmen	Real Estate Investment	Keep	Important factors

2.3 Construction of Regression Model

This paper adopts the form of linear regression model, and the established equation selects 1 dependent variable, 1 independent variable and 5 control variables. Five of the control variables are derived from Table 2. The independent variables are the focus of this study, and the number of annual marriage registrations that can guarantee the marriage intention is selected. The basic model is shown in Equation 2:

$$\mathrm{HP} = a_0 + a_1\mathrm{MR} + a_2\mathrm{GDP} + a_3\mathrm{SA} + a_4\mathrm{PO} + a_5\mathrm{IN} + a_6\mathrm{MA} + u \quad (2)$$

Among them, HP represents the housing price level, MR is the logarithm of annual marriage registration, GDP is the gross domestic product, SA represents the average wage level, PO is the total population at the end of the year, IN is the real estate investment amount, MA is the housing sales area, all are cities Level data. u is the random interference term of the equation, α_0 is the effect of omitting variables.

3. An Empirical Study on the Impact of Marriage on Housing Price Level

3.1 Overview of Multiple Linear Regression Models.

If the national or provincial-level real estate price fluctuation levels are used, local differences cannot be shown, so this paper uses city-level data. The selected cities are mainly municipalities

directly under the Central Government or the capital cities of various provinces. The real estate market development level of these cities is relatively representative in the provinces and regions where they are located. This choice not only takes into account the influence of the city, but also takes into account the representativeness of the selected city sample. Based on the above considerations, this paper selects the 2011-2016 annual data of 4 municipalities, 16 provincial capital cities and 2 key cities for analysis. The regional division of each city is shown in Table 3, and the division refers to the regional division standards in the "Seventh Five-Year Plan" adopted by the Fourth Session of the Sixth National People's Congress.

Table 3 24 selected cities and their regional divisions

Area	City	Number
East Area	Beijing, Tianjin, Shanghai, Nanjing, Suzhou, Hangzhou, Ningbo, Xiamen, Jinan, Qingdao, Guangzhou, Shenzhen, Shijiazhuang	13
Central Region	Hefei, Nanchang, Zhengzhou, Wuhan, Changsha, Shijiazhuang, Hohhot	7
Western Region	Chengdu, Chongqing, Xi'an, Yinchuan	4
Total	4 municipalities, 15 provincial capital cities, 5 key cities	24

In order to investigate how the impact of marriage on housing prices differs in different city development levels, in addition to the 24 cities as a whole, this article divides them into several categories for in-depth research. The standard of division is based on the development status of the real estate market in each city, and it is based on the city’s annual housing price level in 2011, in order to divide the city into three categories, as shown in Table 4.

Table 4 Classification of selected cities

Category	Criteria for the classification	City	Statistical Classification
1	15000≤House price level	Beijing, Shanghai, Xiamen, Hangzhou, Ningbo, Shenzhen	Developed City
2	10000≤House price level<15000	Tianjin, Suzhou, Qingdao, Nanjing, Guangzhou	
3	Housing price level ≤10000	Chongqing, Zhengzhou, Changsha, Yinchuan, Xi'an, Wuhan, Nanchang, Jinan, Hefei Chengdu, Shijiazhuang, Harbin, Hohhot	Underdeveloped City

3.2 Data stationarity test

Panel data requires unit root test to determine whether it is a stationary series. In this paper, we first conduct a summary type unit root test on HP (housing price level). The test results are shown in Figure 1. According to the LLC test result, the test statistic obtained is 2.69446, and the corresponding p value is 0.9965, so the null hypothesis that each cross-section data has the same unit root process cannot be rejected. The value of the IPS statistic is 3.91468, and the corresponding p value is 1, indicating that the assumption that "all interface member sequences have a unit root" cannot be rejected. Therefore, it can be considered that the HP sequence has a unit root. In the same way, unit root tests were conducted on MR, GDP, SA, PO, IN, and MA, respectively, and it was found that their horizontal levels are all unstable series. Therefore, the unit root test of the first-order difference is continued.

Figure 2 shows the results of the LLC unit root test for the first-order difference of HP (housing price level). The LLC statistic value is -1.885586, and the p-value is 0.0317. At a 5% significance level, it can be considered that the first-order difference that refuses to have data has a unit root. Finally, we can conclude that HP's data is non-stationary, but its first-order difference sequence is stationary. In the same way, the unit root test of the first-order difference is performed on MR, GDP,

SA, PO, IN and MA, respectively, and the first-order difference sequence is found to be stationary. It is concluded that the variable data selected by the model in this paper are single-integer of the same order, so the co-integration test can be continued.

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	2.69446	0.9965	24	120
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	3.91468	1.0000	24	120
ADF - Fisher Chi-square	24.7224	0.9979	24	120
PP - Fisher Chi-square	24.8420	0.9977	24	120

Fig. 1 HP unit root test

Method	Statistic	Prob.**
Levin, Lin & Chu t*	-1.85586	0.0317

** Probabilities are computed assuming asympotic normality

Intermediate results on D(HP?)

Fig. 2 Unit root test of HP's first-order difference

3.3 Data co-integration test

After performing mixed regression analysis on the model, the residual sequence is obtained as shown in Figure 3, and then the residual sequence is tested for unit root. FIG. 4 shows the unit root test results of the sequence group consisting of the residual sequence of each cross-sectional data. The LLC test statistic is equal to -7.51275, and the probability value P is equal to 0, so that the null hypothesis that the residual sequence of each cross-section regression equation has the same unit root can be rejected at the 5% significance level. The Fisher-PP test statistic is also very significant, and its probability value is small, so it can reject the null hypothesis that all residual sequences of cross-section regression equations have unit roots. According to the above unit root test, it can be considered that the residual sequence of each cross section of the data in the model does not have a unit root, that is, these residual sequences are stationary, indicating that the panel data series HP and MR, GDP, SA, PO, IN, MA There is a cointegration relationship. Passing the cointegration test shows that there is a long-term stable equilibrium relationship between the variables in this paper, and the original equation can be directly regressed on this basis. The regression results at this time are more accurate.

References are cited in the text just by square brackets [1]. (If square brackets are not available, slashes may be used instead, e.g. /2/.) Two or more references at a time may be put in one set of brackets [3, 4]. The references are to be numbered in the order in which they are cited in the text and are to be listed at the end of the contribution under a heading References, see our example below.

View | Proc | Object | Print | Name | Freeze | Default | Sort | Edit+/- | Smpl+/- | Compare+/- | Transpose+/- | Title | Sample

	RESIDCH...	RESIDZHE...	RESIDCHA...	RESIDYIN...	RESIDXIAN	RESIDWU...	RESIDTIAN...
2011	8450.028	1507.644	447.1775	-952.3730	3975.836	-2398.094	-876.2930
2012	5633.933	-521.4066	-1667.569	-2231.716	-192.3330	-3108.204	-3021.578
2013	4200.346	-1001.957	-1261.660	-1699.441	-1713.226	-2619.885	-2294.352
2014	650.1145	-2753.126	-5536.339	-1835.818	-2354.480	-3854.035	-5719.085
2015	-1315.266	-1268.018	-3228.055	-5217.807	-3947.718	-1989.126	-6193.659
2016	2266.201	4156.863	-865.6135	-5324.468	-3558.770	4138.536	4079.079

Fig.3 Data cross-section residual sequence

Method	Statistic	Prob.**	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	-7.51275	0.0000	24	120
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-0.29391	0.3844	24	120
ADF - Fisher Chi-square	52.7271	0.2963	24	120
PP - Fisher Chi-square	66.4341	0.0401	24	120

Fig.4 Unit root test of residual sequence of cross-section data

3.4 Test results and analysis of the model

The random effect model is established first, and then the Hausman test is performed. The test results are shown in Figure 5. The Hausman Test statistic from which the random effect model can be obtained is 20.26, the p-value is 0.0025, and the null hypothesis is rejected at the 5% significance level. A random effect model should not be established. This article selects the mixed effect model for

testing. Use the mixed effect model to test the overall statistical data in this paper, as shown in Figure 6. The goodness of fit of the model is 0.7667, and the fitting effect is better. The p-value of PO is 0.6237, which is not significant at the 5% significance level. While IN is significant at the 5% significance level, it finally enters the equation. However, the sign is positive, which is different from the expected sign. The reason is that real estate investment generally has market foresight. When making investment decisions, it will make an in-depth analysis of changes in housing prices in the future market. Real estate companies will increase their investment if they anticipate the trend of rising housing prices. Possibility, and thus it has a positive relationship with fluctuations in housing prices. In addition, judging from the p-values of each variable, the variables that finally enter the equation are GDP, SA, IN, and MA, which have a significant impact on housing prices at a 5% significance level. Nationwide, the p value of the key MR (logarithm of marriage registration) studied in this paper is 0.3846, which does not enter the equation.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MR?	0.033927	0.054435	0.623266	0.5355
GDP?	0.539974	0.211634	2.551459	0.0133
SA?	0.333286	0.033871	9.839944	0.0000
PO?	-8.905672	6.713409	-1.326550	0.1897
IN?	3.307442	1.468342	2.252501	0.0280
MA?	-9.946583	2.044460	-4.865139	0.0000

R-squared	0.774367	Mean dependent var	21287.52
Adjusted R-squared	0.755564	S.D. dependent var	10292.53
S.E. of regression	5088.676	Akaike info criterion	19.99393
Sum squared resid	1.55E+09	Schwarz criterion	20.19299
Log likelihood	-653.7997	Hannan-Quinn criter.	20.07259
Durbin-Watson stat	1.010563		

Fig. 5 Hausmann test results of the model

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MR?	0.058671	0.013477	4.353542	0.0000
GDP?	0.452959	0.185971	2.435649	0.0173
SA?	0.107732	0.009535	11.29844	0.0000
PO?	-4.003224	1.318089	-3.037143	0.0033
IN?	0.282747	0.751549	0.376219	0.7079
MA?	-2.152248	0.668598	-3.219044	0.0019

R-squared	0.196100	Mean dependent var	7908.026
Adjusted R-squared	0.140274	S.D. dependent var	1926.164
S.E. of regression	1785.967	Akaike info criterion	17.88711
Sum squared resid	2.30E+08	Schwarz criterion	18.06840
Log likelihood	-691.5973	Hannan-Quinn criter.	17.95968
Durbin-Watson stat	1.052715		

Fig. 6 Mixed regression test results

Similarly, the mixed effect model is used to analyze 11 cities with developed real estate markets, as shown in Figure 7. The goodness of fit of the model is 0.7744, and the fitting effect is better. Judging from the sign of the variable coefficient, its positive and negative values are consistent with the statistical analysis results nationwide, and no additional explanation is given here. From the p-value of each variable, the variables that finally enter the equation at this time are GDP, SA, IN, and MA, which have a significant impact on housing prices at a 5% significance level. Nationwide (the regional distribution of the selected samples can roughly represent the national level), the p value of the key MR (logarithm of marriage registration) studied in this paper is 0.5355, and it has not entered the equation.

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MR?	0.033927	0.054435	0.623266	0.5355
GDP?	0.539974	0.211634	2.551459	0.0133
SA?	0.333286	0.033871	9.839944	0.0000
PO?	-8.905672	6.713409	-1.326550	0.1897
IN?	3.307442	1.468342	2.252501	0.0280
MA?	-9.946583	2.044460	-4.865139	0.0000

R-squared	0.774367	Mean dependent var	21287.52
Adjusted R-squared	0.755564	S.D. dependent var	10292.53
S.E. of regression	5088.676	Akaike info criterion	19.99393
Sum squared resid	1.55E+09	Schwarz criterion	20.19299
Log likelihood	-653.7997	Hannan-Quinn criter.	20.07259
Durbin-Watson stat	1.010563		

Fig. 7 Test results of developed cities

Variable	Coefficient	Std. Error	t-Statistic	Prob.
MR?	0.058671	0.013477	4.353542	0.0000
GDP?	0.452959	0.185971	2.435649	0.0173
SA?	0.107732	0.009535	11.29844	0.0000
PO?	-4.003224	1.318089	-3.037143	0.0033
IN?	0.282747	0.751549	0.376219	0.7079
MA?	-2.152248	0.668598	-3.219044	0.0019

R-squared	0.196100	Mean dependent var	7908.026
Adjusted R-squared	0.140274	S.D. dependent var	1926.164
S.E. of regression	1785.967	Akaike info criterion	17.88711
Sum squared resid	2.30E+08	Schwarz criterion	18.06840
Log likelihood	-691.5973	Hannan-Quinn criter.	17.95968
Durbin-Watson stat	1.052715		

Fig. 8 Test results of underdeveloped cities

Next, analyze the 13 cities with underdeveloped real estate markets. The statistical results are shown in Figure 8. From the p-value of each variable, the variables that pass the significance test also become MR, GDP, SA, PO, and MA. PO (population number) does not enter the equation in developed cities, but enters the equation in undeveloped cities. The reason is that the economic development of developed cities is relatively high, which is more attractive to the population, but the purchasing power of housing for this part of the newly increased population is affected Limitations of many factors, the most important of which is the high level of housing prices. Look at the logarithm of marriage registration that we focus on in this article. In the statistical results of underdeveloped cities, the coefficient is 0.058671, and the p-value is 0, indicating that it has a significant impact on

the housing price level.

The statistical results of MR between developed and undeveloped cities are quite different. This article believes that there are two main reasons:

First, from the perspective of marriage, most of the cities with developed real estate markets are concentrated in coastal provinces, which have a higher degree of openness and a more thorough transformation of traditional Chinese concepts. Families in this part of the city can already accept the situation of not buying a house before marriage, which limits the pull of marriage willingness to the demand for buying a house, so that the effect of the marriage registration logarithm (MR) on the housing price level is not significant enough. In contrast, underdeveloped cities are more conservative in their ideology and concepts, and the idea of "having a house before getting married" is more common, making the act of marriage significantly affecting the demand for housing, which in turn affects the price level of housing.

Second, developed cities tend to have a strong attraction for foreign youths, and they often have a good educational background. This group of people has limited economic capacity at this stage, but their thinking is relatively independent. When making marriage decisions, they often tolerate the lack of housing before marriage and choose to rent a house for a certain period after marriage. To a certain extent, it also limits the pull of marriage demand in developed cities on housing demand, and weakens the impact of the number of marriage registrations on housing price levels.

4. Conclusion

This paper finds through theoretical research and empirical analysis:

(1) From a nationwide perspective, the effect of marriage on housing price levels is less obvious, and the idea of considering the fluctuation of housing prices by using willingness to marry is not widely used.

(2) The effect of marriage on the housing price level is affected by the degree of regional development: in regions with less developed economic development, due to the deep-rooted traditional marriage concept, the impact of marriage on housing demand is more obvious, and its impact on housing price levels It is also more significant. Conversely, in regions with a relatively high degree of economic development, marriage does not have an obvious explanation for housing price levels, and it does not provide guidance for predicting housing prices.

References

[1] Wang Meixia. Analysis of the impact of population structure on China's housing prices [D]. Shandong University, 2016.

[2] Xiao Wu. A survey of the Chinese youth's outlook on marriage[J]. Contemporary Youth Research, 2016(05): 79-85.

[3] Lian Si, Zhao Jinyan. Do I have to buy a house when I get married?——Research on the impact of youth housing on marriage[J]. China Youth Research, 2017(07):61-67.

[4] Liu Xueliang, Wu Jing, Deng Yongheng. Demographic shock, marriage and housing market [J]. Nankai Economic Research, 2016 (01): 58-76.

[5] Madge J, Brown C, Policy Studies Institute. First homes: a survey of the housing circumstances of young married couples[M]. London: Policy Studies Institute, 1981.

[6] 1999 Australian Bureau of Statistics survey of first home buyers [EB/OL]. http://www.abs.gov.au/ausstats/abs@.nsf/2f762f95845417aeca25706c00834efa/6b97231bd938c12eca2570ec00114199! OpenDocument.

[7] Wang Pengfei, Yu Kaichao. Analysis of factors influencing house prices in Kunming based on multiple linear regression[J]. Software, 2018, 39(09): 152-157.

[8] Chen Ke, Chen Wei. An empirical study on the relationship between GDP, wages and housing prices——Taking Shenzhen as an example[J]. Special Zone Economy, 2017(10): 37-40.

[9] Yang Zhongxuan, Yang Yangyang. Research on the influencing factors of housing prices in Henan Province and their prediction[J]. Science and Technology Innovation and Productivity, 2017(07): 33-36+39.

[10] Xu Yonghong, Wu Linying. The dynamic relationship between population characteristics and housing price fluctuations in various regions of China [J/OL]. Statistical research: 1-11[2019-01-10].

[11] Li Zhuohang, Qin Shuning. A study on the empirical relationship between housing prices in Shanghai, average wages of employees and personal housing loan balances [J]. Contemporary Economy, 2018(19): 146-148.

[12] Li Yonggang, Yu Yonghu, He Binbin. Urbanization, income gap and housing price fluctuations——Research based on dynamic panel model[J]. Journal of Nanjing Audit University, 2013, 10(04): 9-17.

[13] Zhou Xiaohan. The Impact of Urban Residents' Income Gap on Housing Prices——Based on the Dynamic Panel Model of 38 Large and Medium Cities[J]. Construction Economics, 2018, 39(12): 75-80.

[14] Li Yuankun. On the analysis of the impact of population structure on China's housing prices [J]. Modern Business, 2018(29): 185-186.

[15] Zheng Jichao, Ni Zeqiang, Sun Wei. The impact of population structure on housing prices——an analysis based on panel data[J]. Jianghuai Forum, 2016(05):22-27+52.

[16] Liu Xiaoli. The Effect of House Price Fluctuation on Macroeconomics under the Purchase Restriction Policy——Taking Shanghai Time Series Data from 2011 to 2017 as an Example[J]. Finance and Accounting Monthly, 2018(21): 170-176.