Whether there is Crowding Out Effect of High Housing Price on Innovation Level: A Panel Data Analysis Based on Prefecture-Level Cities

Wentao Wang[1],Shiming Zhao[*], Yiding Wang[2]

thyekc@stud.tjut.edu.cn[1], zhaoshiming@email.tjut.edu.cn[*], zxz081129@163.com[2]

School of Management, Tianjin University of Technology, Tianjin, 300384, China

Abstract. The rapid development of China's real estate industry has not only driven China's economic development, but also brought about a series of social problems. Based on the innovation capability index and the average transaction price of second-hand housing in 293 prefecture-level cities in China from 2015 to 2021, this paper constructs a panel data model to explore whether high housing prices have an inhibitory effect on urban innovation level, and analyzes its impact mechanism and heterogeneity. This paper analyzes the U-shaped curve relationship between housing price and innovation index, and the heterogeneity of the impact of housing price on innovation development index in cities of different sizes. Finally, according to the experimental results, targeted suggestions are put forward for the relevant government departments of different city sizes, aiming to promote the innovation and development of cities by adjusting the average transaction price of second-hand houses.

Keywords: Fixed effect model; Panel data; Innovation; House price

1 Introduction

Since the reform and opening up in 1978, China's economy has developed rapidly, surpassing Japan's in 2010. In 2020, due to the impact of COVID-19, although the economic growth rate was less than 3%, the total size of China's economy increased to 101.60 trillion yuan, 276 times that of 1978, ranking second in the world[1].However, the rapid development of the real estate industry since the housing reform has brought a series of social problems, such as rising housing prices[2].Since 2010, Chinese cities have implemented real estate regulation policies. From 2011, growth slowed to 6.4 percent in 2019. Faced with resource constraints and the disappearance of demographic dividend, it is necessary to change factors and investment drivers and seek new economic drivers to promote healthy economic development[3].As the value of the real estate owned by the firm increases, the firm increases its financing and investment, which can encourage the firm to innovate[4].The increase in the relative housing price of the city leads to a general outflow of labor and the concentration of skilled labor, which promotes the evolution of the industrial structure from labor intensive to human capital intensive industries[5].The impact of high housing price on industrial structure and the change of industrial structure will further affect corporate innovation[6].From the perspective of talent flow, it will also have a variety of impacts on regional innovation capacity[7-9].By testing the demand-side mechanism in the context of internal circulation, it can be found that the impact

MSIEID 2023, December 08-10, Guangzhou, People's Republic of China

DOI 10.4108/eai.8-12-2023.2344795

of regional housing prices on entrepreneurship and innovation has an inverted U-shaped characteristic of first promoting and then inhibiting[10].In addition, the fluctuation of housing prices will also have an impact on technological innovation[11].Therefore, it is particularly important to explore whether there is a crowding-out effect of high housing prices on the level of innovation. This paper aims to explore the relationship between urban housing price and corporate innovation. This paper analyzes the "U-shaped" relationship and explores its mechanism from the perspectives of enterprise investment, labor mobility and policy regulation.

2 Model setting and data

2.1 Model Setting

Model 1: The impact of housing price on regional innovation development level is U-shaped, that is, it first inhibits and then promotes.

In the analysis of panel data, the fixed effect model can fully reflect the differences in intercept terms among individuals. Based on Model 1, the benchmark regression model of regional innovation development level and housing price is set as the equation (1):

$$\text{Innovation}_{it} = \alpha + \beta\ \text{hopri}_{it} + \gamma\ \text{hopri}_{it}^2 + \delta\text{hopri}_{it}^3 + \psi X_{it} + \varepsilon_{it} \tag{1}$$

Where, Innovationit represents the innovation index of a certain region in a certain year; α represents the intercept term of the model, that is, when the influence of other variables does not exist, what is the base of the innovation index in a certain region; hopriit represents the average transaction price of second-hand houses in a certain area at a certain year time node; X_{it} represents each control variable; ε_{it} is the random disturbance term.

Model 2: The impact of housing price on the level of regional innovation development varies significantly with different regions.

It is assumed that the impact of housing price on regional innovation development level varies from region to region, and the grouped regression method is used for testing and analysis. The regression analysis equation for testing is as the equation (2)：

$$\text{Innovation}_{it} = \lambda + \omega\text{hopri}_{it} + \omega X_{it} + \sigma_{it} \tag{2}$$

The domestic cities are divided into three groups according to the number of population for regression analysis. Megacities and giant cities are defined as large cities, while large cities and medium cities are defined as medium cities, and small cities and micro cities are defined as small cities.

2.2 Data Description

The names, meanings and data sources of variables in specific experiments are shown in Table 1. The data samples in this paper are based on the relevant indicators of 293 cities in China from 2015 to 2021 to construct a panel data model.Descriptive statistical analysis is shown in Table 2.

Table 1. Name, meaning and source of experimental variables

Name of variable	Meaning of	Source of information
Innovation	Regional Innovation index	Index of Regional Innovation and Entrepreneurship in China
hopri	House price	Python crawled Anjuke second-hand house transaction records
GDP	Gross regional product	China Statistical Yearbook
GDP_growth	Growth rate of regional GDP	China Statistical Yearbook
inc	Per capita disposable income of all residents	China Statistical Yearbook
popu	Urban permanent population	China Statistical Yearbook
sch_num	Number of institutions of higher learning	China Statistical Yearbook
edu	Proportion of employed persons with bachelor degree	China Family Longitudinal Survey

Table 2. Descriptive statistical analysis

	Value of observation	Minimum value	Maximum value	Mean value	Standard deviation
year	2051	2015	2021	2018.06	1.990
Innovation	2051	0.6830	100.0000	51.8667	28.9888
hopri	2051	219	59695.5	8048.55	6153.17
GDP	2051	38.1900	40269.0000	2813.0953	1534.23
GDP_growth	2051	2.1	30.4	7.536	4.1025
inc	2051	0.977	6.943	2.415	1.113
popu	2051	0.008	0.887	0.262	0.179
sch_num	2051	0	93	13	7.85
edu	2051	2.761	30.39	8.270	5.461

3 Empirical analysis

3.1 Regression Analysis

The cross-section fixed effect model is used for estimation, and the impact result between the transaction price of second-hand housing and the regional innovation and development level is obtained. The number of colleges and universities owned by the region from 2015 to 2021 is set as the instrumental variable of the regional innovation and development level, and the two-stage least squares regression is conducted on the regional innovation and development index, which is used to solve the endogeneity problem in the model.

Table 3. Regression results of Model 1

Name of variable	Nationwide (1)	Nationwide (2)	Nationwide (3)
	OLS	**FE**	**2SLS**
hopri	-0.209*[1]	-0.045***	-0.108***
	(-1.43)	(-3.27)	(-3.25)
$hopri^2$	0.210*	0.328***	0.063**
	(1.32)	(5.26)	(2.21)
$hopri^3$	0.018*	-0.001	0.001
	(0.91)	(-0.17)	(0.20)
GDP	0.007	0.046	0.010
	(0.13)	(0.82)	(0.38)
GDP_growth	0.129	0.626***	0.328***
	(1.47)	(2.86)	(3.28)
inc	2.757***	4.671***	2.653***
	(2.89)	(5.51)	(2.77)
popu	0.269***	0.089*	0.186**
	(3.59)	(1.03)	(2.54)
sch_num	1.660**	6.464***	2.436***
	(2.41)	(3.05)	(3.56)
edu	2.982***	6.046***	3.185***
	(2.63)	(4.73)	(2.80)
Value of observation	2051	2051	2051
Region fixed effect	No	Yes	No
Adj-R^2	0.871	0.790	0.868
Statistic of F	252.4	26.22	
Hausman test (p value)	0.000	0.000	0.000

Note:1. The standard error based on heteroscedastic robustness standard error method is estimated in brackets.

2.*, **, ***represent the regression coefficients are significant at the level of 10%,5% and 1% respectively.

It can be seen that the regression results of the three models shown in Table 3 all pass the F test, indicating that the regression method of panel data model has scientific basis. The adjusted R square of the least squares estimation and the two-stage least squares estimation are both greater than 0.8, indicating that the model has a good fitting effect. The explanatory variables and control variables selected in this paper have a good level of explanation for the explained variables. Among them, the fixed effect model passes the Hausman test, indicating that the panel data model is suitable for the application of fixed effect model rather than random effect model. Under the cross-section fixed effect model, after controlling the regional effect, the first power term and the second power term of the housing price are both significant at the level of 1%, while the third power term is still insignificant at the level of 10%, which is consistent with the research results of some other scholars in China. Regarding the coefficient of the regression results, the first power term of the housing price is significantly different from zero, the second power term is significant at the level of 1%, and the regression coefficient is significantly greater than 0, and the third power term is not significant at the level of 10%, and the regression coefficient is not obvious. The above analysis shows that the regression results in Table 3 are consistent with the hypothesis of Model 1 in this paper.

Among the other control variables: the estimated coefficient of GDP is 0.046, which is positive at the significance level of 10%. The estimated coefficient of GDP growth rate is 0.626, which is positive at the significance level of 1%, so the regression result of this variable is significant at the level of 1%. The estimated coefficient of residents' per capita disposable income is 4.671, so the per capita disposable income of all residents is significantly correlated with the level of innovation development at the level of 1%. Cities with a large population base will also have more innovative talents, so the variable of the number of permanent residents in the city is significant at the level of 10%. The number of institutions of higher learning and the proportion of employees with bachelor's degree both represent the innovation ability of workers in a region. The regression curve is added to the scatter plot of the explanatory variable and the explained variable, as shown in Figure 1. It can be clearly seen from the figure that there is a U-shaped relationship between the explanatory variables and the explained variables.

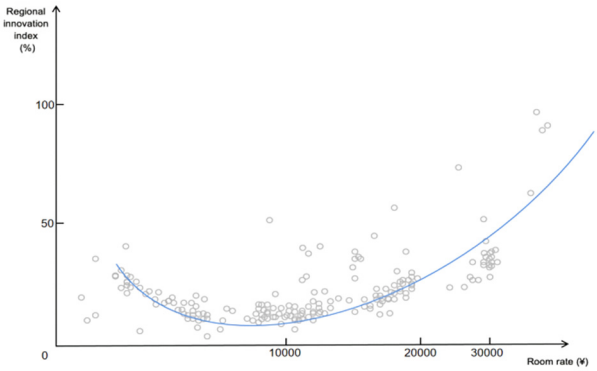

Fig. 1. Scatter plot of regression curves

$$l: hopriit（min） = -\beta/2\gamma \quad (3)$$

The corresponding hopriit=9788 of the line is put into equation (3) to calculate the lowest point in the "U-shaped" curve, and then all the sample points with hopriit less than 0.52 in the experimental data are screened out for linear regression analysis with all the sample points with hopriit greater than 9788. The results are shown in Table 4.

Table 4. Regression results with breakpoints

Name of variable	Grouping of groups （1）	Grouping of groups （2）
	OLS	**OLS**
hopri	-0.527*[1] (-1.43)	0.738*** (-3.27)
GDP	0.011 (0.13)	0.031 (0.82)
GDP_growth	0.301 (1.47)	0.524*** (2.86)
inc	2.831*** (2.89)	3.194*** (5.51)
popu	0.288***	0.070*

	(3.59)	(1.03)
sch_num	1.900**	5.236***
	(2.41)	(3.05)
edu	2.091***	5.847***
	(2.63)	(4.73)
Value of observation	181	112
R^2	0.920	0.899
Statistic of F	54.223	46.236

Note:1. The standard error based on heteroscedastic robustness standard error method is estimated in brackets.

2.*, **, ***represent the regression coefficients are significant at the level of 10%,5% and 1% respectively.

According to the results of regression breakpoint, the scatter plot of regression line with breakpoint is drawn as shown in Figure 2. It can be intuitively seen from the figure that the regression results of each sample point on the left side of the symmetry axis of the "U-shaped" curve show a negative regression line, and the regression results of each sample point on the right side of the symmetry axis of the "U-shaped" curve show a positive regression line. The U-shaped relationship between explanatory variables and explained variables has been confirmed.

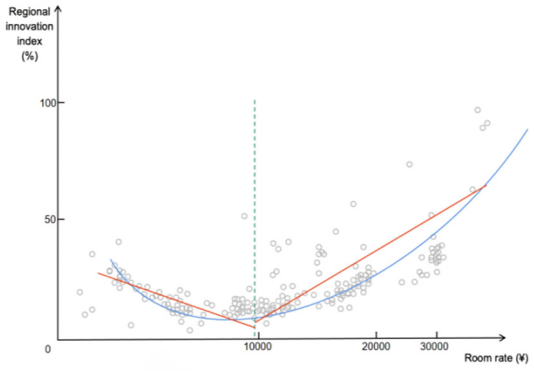

Fig. 2. Grouped regression diagram

The regression results of panel data are shown in Table 5 for the question of whether there is heterogeneity in the relationship between housing prices and innovation development levels.

Table 5. Regression results of Model 2

Name of variable	Large （1）	Middle （2）	Small （3）	Large （1）	Middle （2）	Small （3）
	FE	**FE**	**FE**	**FE**	**FE**	**FE**
hopri	0.652	-0.619***[1]	-5.399**	3.371***	-0.295**	-2.226
	(-0.72)	(-4.20)	(-2.82)	(3.05)	(-1.83)	(-0.86)
GDP	0.170**	-0.970	-0.903*	0.839	-4.113**	-0.936***
	(2.33)	(0.59)	(-2.01)	(1.62)	(-2.27)	(-2.72)

GDP_growth	0.611	2.938***	0.854***	3.426***	1.265	0.799***
	(0.74)	(3.49)	(3.72)	(3.18)	(0.66)	(4.16)
inc	14.568***	9.079*	0.044	7.307**	0.849	3.615*
	(3.63)	(2.04)	(0.02)	(2.08)	(0.18)	(1.81)
popu	-0.014	0.016	0.023**	-0.001	0.038	0.047*
	(-0.87)	(0.87)	(3.03)	(-0.02)	(0.70)	(1.82)
sch_num	0.065	0.563***	-0.058	-0.117	0.073	-0.027
	(0.84)	(3.43)	(-1.16)	(-1.63)	(0.41)	(-0.49)
edu	3.454**	2.451**	-0.002	3.147***	2.746	0.706*
	(3.09)	(2.39)	(-0.01)	(3.24)	(1.05)	(1.69)
Value of observation	462	784	805	462	784	805
Region fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Time fixed effect	No	No	No	No	No	No
Adj-R^2	0.676	0.684	0.774	0.787	0.855	0.829
Statistic of F	97.67	49.46	347.4	26.80	31.39	29.01
Hausman test (p value)	0.000	0.000	0.000	0.000	0.000	0.000

Note:1. The standard error based on heteroscedastic robustness standard error method is estimated in brackets.

2.*, **, ***represent the regression coefficients are significant at the level of 10%,5% and 1% respectively.

The 293 prefecture-level cities in China are divided into three groups of large, medium and small for regression analysis, and it is found that the regression result of housing price in large cities is positive but not significant, and the regression result of small and medium-sized cities is significantly negative. Under the dual fixed effect, the regression result of housing price in big cities is significantly positive, the regression result of medium cities is significantly negative, and the regression result of small cities is negative but fails the test.

3.2 Robustness Test

In order to ensure the robustness of the regression results, after replacing the experimental data with the cross-sectional data of 2021, the least square method is used for regression analysis to test the robustness of the model. After replacing with other data, it can be seen from Table 6 that the quadratic term of housing price is still positive and significant at 1% confidence level. Therefore, the regression analysis results in this paper have high robustness.

Table 6. Robustness test

Name of variable	Nationwide （1）
	OLS
hopri	-0.349*[1] (-1.43)
$hopri^2$	0.440*** (3.57)

$hopri^3$	0.002* (1.08)
GDP	0.013 (0.26)
GDP_growth	0.226* (1.17)
inc	2.903*** (2.78)
popu	0.374*** (3.61)
sch_num	1.087** (2.33)
edu	2.889*** (2.90)
Value of observation	293
Adj-R^2	0.921
Statistic of F	198.3
_cons	3.428*** (0.242)

Note:1. The standard error based on heteroscedastic robustness standard error method is estimated in brackets.

2.*, **, ***represent the regression coefficients are significant at the level of 10%,5% and 1% respectively.

4 Conclusion

This paper uses panel data model to test the two models, and finds that there is a U-shaped relationship between regional innovation and development index and the average transaction price of regional second-hand housing. When the average transaction price of second-hand housing is low, there is a significant negative correlation between the average transaction price of urban second-hand housing and the innovation and development index, which means that as the average transaction price of second-hand housing increases, the regional innovation and development index will gradually decrease. When the average transaction price of second-hand housing reaches a certain threshold, there is a significant positive correlation between the average transaction price of urban second-hand housing and the innovation and development index, which means that with the increase of the average transaction price of second-hand housing, the regional innovation and development index will gradually increase. This "U-shaped" curve relationship has passed the regression discontinuity verification and robustness test, proving its existence and stability. The average transaction price of second-hand housing in cities of different sizes has different impacts on regional innovation and development index. In small cities, an increase in the average transaction price of second-hand houses will lead to a decline in the innovation index, in medium-sized cities, this downward trend is relatively slow, and in large cities, an increase in the average transaction price of second-hand houses will lead to an increase in the innovation index. According to the analysis, this may be caused by the differences in residents' expectations of future urban development and housing price in cities of different sizes. For the optimization of housing price regulation policy, we should first maintain the stability of housing price, avoid the

inhibitory effect of too rapid rise in housing price on urban innovation, crack down on real estate speculation, and prevent irrational rise. We should give full play to the role of the government and the market, guide and coordinate the implementation of policies, reasonably guide the direction of market development, and avoid the rapid rise of housing prices. We will support urban innovation activities and provide financial support and policy preferences.

Reference

[1]Li J W. The overall forecast and outlook of China's economic and social development (1996-2020). Management World (4), 10 (1997).

[2]Huang X W.. Research on the role of real estate economy in promoting urban development: A review of "Real Estate and Urban Development: Problems and Strategies". Modern Urban Research (2022), 37(10), I0004.

[3]Peng Y F. Analysis of the current situation and influencing factors of China's innovation-driven development. Social Science Trends, 000(004), (2021).27-33.

[4]Chaney T, Sraer D, Thesmar D. The Collateral Channel: How Real Estate Shocks Affect Corporate Investment. American Economic Review, (2012), 102(6).

[5]Shao C D, Su D N, Deng H T. Housing price, land finance and urban agglomeration characteristics: The road to Chinese-style urban development [J]. Management World, (2016).

[6]Lai J M, Ouyang J, Xu B H. Research on the impact of stable urban housing price on industrial structure upgrading [J/OL]. Price Theory and Practice: 1-4 (2023).

[7]He Z H, Sui S M, Wang J L. The impact of housing price on enterprise innovation: Based on the perspective of innovation input and talent policy [J]. Studies in Science of Science, (2022),40(12):2281-2293.

[8]Amy Jing Lin, Jianming Cai, Douglas WEBSTER, & Larissa MULLER. (2018). International practice and enlightenment of humanistic construction of science and technology innovation zones. Geographical Research.

[9]Xu Wenting, Zhou Jianjun, Ju Fang. Does high housing price inhibit female entrepreneurship? - Empirical analysis based on CMDS data [J]. Science Decision, (2022).

[10]Yang Haodong, Wang Gaofeng. Why does housing price affect regional innovation and entrepreneurship? - Demand-side mechanism test under the background of internal circulation [J/OL]. Price Monthly, (2023).

[11]Zhong Guohui, Yao Jinhai. The spatial effect of housing price fluctuations on technological innovation: Based on the provincial panel spatial lag model [J]. Price Monthly, (2022).