An Empirical Study of Supply Factors Affecting on the Housing Price

Yan feng Han [1], Ying Chen [2]

Xi'an University of Architecture and Technology, School of Management, Xi'an, China, 710055

Email: cy19860304@126.com

Abstract: This article analysis the affecting of Supply factors on the housing price from the quantitative point of view, With the Trade 1998-2009 housing prices in Beijing as the research object, and select construction and installation costs, Land costs, The development area of completion and real estate investment as a representatives of the supply factors, At the same time, Through this article used mathematical analysis about multiple linear regression , And combined with SPSS software tools for regression analysis , Thus to explore the size of the impact for supply factors on housing prices.

Key words: Supply factors; House prices; Empirical study

Introduction

Although the research literature on housing prices is very abundant in recent years, The housing price is affected and restricted by the political, economic, social, administrative, natural and the other factors, However the current domestic research focused on qualitative analysis, And mere indicated that these factors have an impact on housing prices, While the lack of scientific quantitative analysis and empirical research result and reliable analytical tools and model of housing price , It is impossible to know which factors impact on the urban housing price and how is the degree of influence. This paper through the establishment of multiple linear regression model, With Beijing in 1998 -2009 commercial housing market data, And select the supply factors for example to analyze the relationship between the housing price and the supply factors.

1 The model of affecting factors for housing prices under the influence of supply factors

To construct the model of affecting factors for housing prices under the influence of supply factors**:**

$$y = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + \varepsilon$$

This model consists of five predictor variables, namely: x_1 --Construction and installation cost (/); x_2 --the land cost (/); x_3 -- Completed (million square meters); x_4 --real estate investment (billion); five regression parameters -- a_0、a_1、a_2、a_3、a_4, a random disturbance ε .

2 Linear regression model estimation and parameter handling

The collected data on the dependent variable and independent variables (Refer schedule 1) are inputted to the computer SPSS analysis software, Select the force into the (ENTER) method for multiple linear regression analysis, And analysis the degree of the impact of various factors on housing prices.

2.1 OLSE (least squares) estimates

According to SPSS software for data analysis of regression coefficients, the following table:

2.2 Regression statistical test

2.2.1 The significance tests of regression equation

In the study of practical issues, we can not determine indeed linear relation between explain the variables and explanatory variables in advance, Before conducting regression coefficient estimates, The use of multiple linear regression equation to fit linear relation between explain the variables and explanatory variables is just some of the assumptions according to some qualitative analysis. Therefore, after obtained linear

978-1-935068-18-1

regression equation, It is the need of significance tests for the regression equation, This article carried out significance test of the regression equation by goodness of fit to take this test method and F test significance.

Table 1: SPSS regression coefficients 1

Coefficient

Model		Non-standardized coefficient		Standard coefficient	t	Sig.
		B	Standard error	Trial		
1	(Constant)	-67.340	374.585		-.180	.862
	x_1	.964	.316	.252	3.045	.019
	x_2	1.730	.205	.758	8.447	.000
	x_3	.053	.140	.012	.377	.718
	x_4	-.038	.268	-.008	-.142	.891

a. Dependent variable: y

To be dependent variable y on the multiple linear regression equation of x_1, x_2, x_3, x_4:

$$\hat{y} = -67.34 + 0.964x_1 + 1.73x_2 + 0.053x_3 - 0.038x_4$$

2.2.1.1 Goodness of fit test

Goodness of fit test is to test the sample level of the regression equation fit the observed values. In the application of multiple linear regressions, we always use the correlation coefficients to represent the quality of the degree about the regression equations fitting the original data, And to measure the size of the linear relationship between the dependent variable and independent variable as a whole. The range of fitting is [0, 1] if the regression equation is fully fitted the sample observations, Then R^2=1. Entirely fitting is an extreme case, It is unlikely to occur in practical problems, that is not equal to 1, But R^2 is closer to 1; the fitting degree of the regression equation is the higher.

Table 2: SPSS regression analysis 2

Model Summary

Model	R	R^2	Adjustment R^2	Standard error of estimate
1	1.000[a]	.999	.999	118.13524

a. Predictor variable: (Constant), x_1, x_2, x_3, x_4

According to above the regression results, we can be seen, the model goodness of fitting is as much as 0.999, adjusted R^2 is as high as 0.999. It can be said that the model goodness of fit for the observed values of the sample is very high. But to judge the quality of regression analysis, the goodness of fitting is not the only criteria, and sometimes, you can sacrifice a bit of goodness of fit for the pursuit of economic significance of the model.

2.2.1.2 F test

The significance test of the regression equation is to test whether the establishment of linear relation is

significant in the overall. According to regression results shown in Table 3 as follows:

Table 3: SPSS regression analysis 3

Anova[b]

Model		Squares	df	Mean square	F	Sig.
1	Return	1.191E8	4	2.977E7	2133.178	.000[a]
	Residual	97691.542	7	13955.935		
	Total	1.192E8	11			

a. Predictor variable: (Constant), x_1, x_2, x_3, x_4

b. Dependent variable: y

F statistical value is 2133.178, In the n = 12, k = 4, a given significance level α = 0.05 case, the former is obtained critical value $F_{0.05}(4,12-4-1)$ =4.12, $F > F_{\partial}$, So the linear model in the 95% level of probability is established under significant.

2.2.2 The significant test of the variables

For the model of multiple linear regression, The general linear relationship of equation is significant, Which can not explain each explanatory variable on the explained variable effects are significant, So we must test the significance for each explanatory variables to determine whether is retained as a important explanatory variables in the model. If a variable impact on the explanatory variables were not significant, It should be removed to create a more simple model. This is the task of variable significant test. According to regression results shown in table 1 above: t_0=-0.18, t_1=3.045, t_2 =8.447, t_3 =0.377, t_4 =-0.142, In a given significance level α = 0.05 case, Check the critical value of 7 degrees of freedom in t distribution table, get $t_{0.025}$ =2.3646, obviously, $|t_1| > t_{0.025}$, $|t_2| > t_{0.025}$, $|t_3| < t_{0.025}$, $|t_4| < t_{0.025}$, Therefore, x_1, x_2 adopted the significance test of variable, But x_3, x_4 not adopted the significance test of variable, That is, completed area and real estate development and investment as explanatory variables are not significant, The effect of construction and installation cost and land price on the housing pricing is significant.

2.2.3 Autocorrelation Test

According to regression analysis, DW statistic is 1.514, at 5% and 1% significance level, The values of the lower threshold limit d_L and limit d_U were 0.624 and 1.963, $d_L \prec D.W \prec d_U$, can not determine whether there is autocorrelation. Using partial correlation coefficient test, It can be seen from the partial correlation coefficients (PAC), all $|\rho|$<0.5 (corresponding to the direct box are also visible within the dotted line), Therefore, the equation does not exist autocorrelation.

2.2.4 Testing for Heteroscedasticity

The white testing is non-cross terms, and to get the following table:

Table 4: SPSS regression analysis 4

White Heteroskedasticity Test

F-statistic	0.232625	Probability	0.865423
Obs*R-squared	4.158634	Probability	0.786535

a. Predictor variable: (Constant), x_1, x_2, x_3, x_4

In the case of non-cross terms, degrees of freedom is 7, the results from the above table

978-1-935068-18-1 © 2010 SciRes.

were $nR^2 = 4.158634$, take a significant level of ∂=0.05, $\chi^2(7) = 14.067 > nR^2$, Therefore there is no heteroscedasticity in the equation.

2.3 The parameters processing of linear regression model

According to the above regression models, it can be seen x_1 ---construction and installation costs and x_2 ---land prices can significantly affect the explained variable. Hence, they are remained in the model, other factors must be removed out from the model, under the influence of supply factors, the model becomes:

$$y = a_0 + a_1 x_1 + a_2 x_2 + \varepsilon$$

Three regression coefficients a_0、a_1、a_2, a random disturbance ε. OLS estimates again, get the output in the following table.

Table 5: Regression results of a re-analysis of SPSS

Model Summary

Model	R	R^2	Adjustment R^2	Standard error of estimate
1	1.000[a]	.999	.999	107.41976

a. Predictor variable: (Constant), x_1, x_2

Table 6: Regression results of a re-analysis of SPSS

Coefficient

Model		Non-standardized coefficient		Standard coefficient	t	Sig.
		B	Standard error	Trial		
1	(常量)	3.992	196.184		.020	.984
	VAR00002	1.026	.263	.269	3.905	.004
	VAR00003	1.672	.157	.733	10.652	.000

a. Dependent variable: VAR00001

Anova[b]

Model		Squares	df	Mean square	F	Sig.
1	Return	1.191E8	2	5.954E7	5159.710	.000[a]
	Residual	103851.051	9	11539.006		
	Total	1.192E8	11			

a. Predictor variable: (Constant), VAR00003, VAR00002。

b. Dependent variable: VAR00001

Based on the above analysis, In the case of only two variables---the construction and installation costs and the land prices, The regression equation still has a high degree of goodness of fit, The value of R^2 is 0.999. At the same time F value also passed the test, to reflect the establishment of the regression equation was significant, indicating that the regression equation is meaningful. In the 5% significance level, all variables passed the significance test of the variables, which means that regression coefficient is significant. We can see construction and installation costs and the land prices is still quite strong explanatory power on commercial

housing price.

The above analysis and testing can be drawn, under the influence of supply factors, regression equation is as follows:

$$y = 3.992 + 1.026x_1 + 1.672x_2$$

3 The analysis for linear regression results

First of all, It can be seen construction and installation costs and the land prices enter into the final model as the variable from the above analysis, The adjusted value of R^2 is also reached 0.999, which indicate that the independent variables on the dependent variable to explain the extent of 99.9%, That is, the model has good explanatory power. In addition, we can see construction and installation costs and the land prices are positive effect (coefficient sign is positive) from the model. These results and the actual situation are consistent, also consistent with our everyday experience and common sense of economics. A city's commercial housing price level, Always improved as the construction and installation costs and the land prices of the city's real estate continuously increased. This is the inevitable result of supply and demand. Although the factors develop in accordance with an unconventional development model in the years of the development process in the Beijing, This law is still to be reflected in Beijing. T-statistics from the value of two factors, we can see that the land cost has considerable influence on commercial housing prices, followed by the construction and installation cost.

Then, it can be seen from the specific values that the corresponding price of commodity houses will increase 1.026 dollar, when the construction and installation costs increase a dollar per square meter, In general, the increasing of the construction and installation costs will bring significantly increased price of commodity houses. The corresponding price of commodity houses will increase 1.672 dollar, when the land cost increase a dollar per square meter. Therefore, the increasing of the construction and installation costs and the land prices will bring the increase in the price of commodity house; they are major factors for the increase in commercial housing price.

Thanks

My thesis is completed under the guidance of my mentor --Professor Han Yanfeng from topics, research into the whole process of writing, Lots of ideas of the paper is benefited from the profound inspiration and teachings of my mentor. In this, I express my heartfelt thanks to Teacher Han. Secondly, I would like to thank the teachers of School of Management at the Xi'an University of Architecture and Technology, Their earnest and selfless urge the successful completion of my thesis. At the same time, I would like to express my sincere gratitude to the students helped me on the learning and life at the graduate level.

Finally, Thanks the experts, professors of approval group for paper in busy for the guidance of my thesis!

References

[1] Run Lizhen Analysis and Empirical Research of the affecting factors of the House prices in Xi'an [D]. Chang'an University, a master's degree thesis, 2007.

[2] Li Kangjing The research of the house price influencing factors [D]. Tianjin University, a master's degree thesis, 2008.

[3] [3] Ma Zheng Analysis of the causes of rising real estate prices from both supply and demand [J]. Business Economics, 2008.

[4] [4] Qin Yingxia The multiple linear regression model of real estate price in china [J]. Practice and Theory of the Mathematics, 2009.

[5] [5] Liu Xiao-jun Shang Peng linear regression equation in house prices [J]. Management perspective, 2009.