A Study on the Influence Factors of Real Estate Prices Based on Econometric Model: A Case of Wuhan

Yi-hong XU\textsuperscript{1,a}, Ke-xin XU\textsuperscript{2,*} and You-quan CHEN\textsuperscript{3}

\textsuperscript{1}Jianghan University, Wuhan, 430056, China

\textsuperscript{2}Department of mathematics and statistics the University of Melbourne, VIC, 3010, Australia

\textsuperscript{3}Jianghan University, Wuhan, 430056, China.

\textsuperscript{a}Email: xuyihong2208@126.com

\textsuperscript{*}Corresponding author

Keywords: Econometric model, Real estate prices, Influence factors, Error correction model.

Abstract. Based on Econometric Model, this paper analyzes the factors influencing real estate prices, aiming at grasping the mechanism of real estate price changing and then providing a reference for government policy. According to the statistical data of the real estate in Wuhan from 2000 to 2012, the error correction model of the real estate price is established to analyze the factors. The result shows that the factors influencing real estate prices are diversified and complex, and there exists a long-term equilibrium relationship among the loan interest rates, the population and the price of the commercial buildings. Therefore, in the long run, the price of the commercial building will be supported by the supply and the demand, and the effect of interest rate is relatively more notable.

Introduction

Due to the characters such as large market demand, long industrial chain and integrated with knowledge-intensive, technology-intensive and labor-intensity, the real estate industry has been the basic industry promoting the national economy to grow. Real estate is the leading industry in national economy. It has long industry chain and high relevancy with other industries, which determines that the sensitivity coefficient and influence coefficient of real estate market are above the average of various industrial sectors in national economy. However, the real estate market is not developed in China and the industry structure is weak, which caused the real estate prices have been constantly climbing in recent decades and a lot of problems and risks exist in the real estate market.

The factors influencing real estate prices are diversified and complex. We need to analyze from the angle of supply, demand and supply-demand, and establish error correction model of real estate prices factors, with the aim of discussing which factors influence real estate prices, how they influence, and whether the sensitivity coefficient is high on the basis of empirical research model. We hope to understand the mechanism of real estate prices changing through these analyses. According to the real estate empirical analysis from 2000 to 2012 in Wuhan, the paper establishes error correction model of the real estate prices factors in Wuhan, discusses statistical model of long-term trend and short-term fluctuation, and then proposes some countermeasures to reasonably control house prices in Wuhan.
Reviews of Domestic and Foreign Research Status

In terms of the factors affecting real estate prices, domestic and foreign scholars have done quite a few researches around relations between economic fundamentals factors: Case and Shiller (1990) performed a time series analysis by using cross-section regression on quarterly data of four American metropolis, which proved that housing costs were positively correlated with percentage of prices, population, personal income and excessive profits or price changes in real estate market several years later. It also explained single-family-home market was not effective. Poterba chose four influencing factors, median housing price, building cost, population and income to analyze and drew a conclusion that the rising urban housing price could be explicated by changes in true income and building cost. Clapp and Gaccoto performed a simple OLS regression analysis by using repeat sale price index, assessment of residential value price index, expected inflation rate, unexpected inflation rates, employment change, which proved economic change (population, employment, economy) is capable to predict changes in housing price. This doesn’t accord with efficient markets hypothesis. Case and Mayer (1996) performed an empirical analysis by using Two Stage Least Square Method on annual data in Boston, and found the median household income had little effect on housing price. Malpezzi (1999) expounded housing price was not random and could be partly predicted successfully. Quigley (1999) emphasized that relevant indexes of economic fundamental could explain the trend of real estate price; though in the short term they couldn't explain too much price fluctuation of the real estate.

At home, Zeng Xiang added dummy variable D to represent political and economic factors (Midwest 0, Eastern 1), and described the land acquisition cost (X2), per capita disposable income (X3), affordable housing sale area (X4) are positively correlated with house price changes, then he obtained a simple regression model as is shown in equation 1:

\[
Y_t = -85.32182 + 0.681462X_{2t} + 0.176461X_{3t} - 4.799472X_{4t} + 1272.445D_t
\] (1)

Zhou Xiaoyan analyzed the reasons of housing price rise from the supply-demand angel, and then she put forward the suggestion that we can promote the construction and consumption of low-cost housing through property taxes. Hu Xiufu (2002) proved that such factors like stopping welfare housing and land transaction prices affected real estate price in China; Wang Shuyun (2005) explained that it was bank’s extreme expansion in the real estate market that caused real estate price to rise; Yan Jinhai (2006) indicated that land price fluctuation caused housing price fluctuation in the short-term, and land price and real estate price influenced reciprocally in the long-term trend, Zhang Mu (2006) emphasized that the inflow of foreign fixed investment continued to promote the rise of real estate price in China; Cao Yulong and Xu Fei indicated that the main housing price factors in Shanghai are land transaction price, completed construction area that very year, provident fund loan interest rate of personal housing and loan interest rate of individual housing; Qi Qian (2011) proved that house price was mainly influenced by factors such as interest rate, per capita income, price of land and real estate development investment through empirical research. Qiao Lin and Chu Chunli (2012) analyzed 33 cities in our country through factor of contribution rate method and drew a conclusion that population and income growth are the most important factors affecting housing price. In addition, investment behavior existing in a certain extent was also a reason causing the growth of housing price.
Empirical Analysis of Influence Factors of Real Estate Prices

Real estate price including price of land and cost of building is a diversified economic category. The value of building, natural land resource, labor input in land combine the value of real estate.

Real estate price is affected by many factors such as economy, politics, society, politics, regulations, internal environment and surroundings.

Error Correction Model

Error correction model was first put forward by Sarger. In 1978, Davison, Hendry, Srba and Yeo put forward the basic framework of error correction model, so ECM can also be called DHSY model. ECM is a control measure to the long-term equilibrium relationship—deviation of short-term fluctuation between variables of cointegration.

The Concept of Cointegration. The concept of cointegration was first put forward by Engle and Granger in 1987, and it made the regression of nonzero single integer become meaningful.

There is long-term dependency relation between some variables in economic system, which is called equilibrium relationship in economics and is the basis of econometric modeling.

The equilibrium relationship shows that mechanism forming in economic system is stable. When season factors or random disturbance make it departure from its balance point, the balance mechanism will adjust to make it return to equilibrium state in the next installment.

But if the deviation is enduring, the balance mechanism between variables will be unstable and the balance relationship will be destroyed. Cointegration is the statistic expression of this balance relationship.

Time series \( \{X_t\}, \{Y_t\}\) is two course of \( I(1)\), if there exists \( \beta\) making \( Y_t - \beta X_t\) become \( I(0)\), we call \( \{X_t\}\) and \( \{Y_t\}\) are cointegration. Actually, cointegration means that some kind of linear combination of multiple non-stationary time series is stationary.

Generally, if sequences \( Y_1, Y_2, \ldots, Y_{k_t}\) are all \( d\) integration, then there exists nonzero vector \( \alpha = (\alpha_1, \alpha_2, \ldots, \alpha_k)\),

which make \( \alpha Y_t \sim I(d-b)\),

when \( b > 0\), \( Y_t = (Y_{1_t}, Y_{2_t}, \ldots, Y_{k_t})'\),

we call \( Y_{1_t}, Y_{2_t}, \ldots, Y_{k_t}\) is \( (d,b)\) cointegration, denoted as \( Y_t \sim CI(d,b)\). \( \alpha\) is co integrated vector.

Error Correction Model (ECM). The co-integration relationship between variables shows that there exists long-term and stable equilibrium relationship, which forms under constant fluctuations of the short-term dynamic process. The long-term equilibrium relationship between variables is due to an adjustment mechanism-error correction mechanism which deviation of long-term relationship is controlled within limits. There exists error correction mechanism in every set of cointegration sequence which reflects short-term regulation behavior.

For sequence \( Y_t, X_t\), error correction model is shown in formula 2

\[
\Delta Y_t = \beta_0 + \beta_1 \Delta X_t - \lambda \Delta \text{ecm}_{t-1} + \epsilon_t
\]
$ecm$ represents error correction term. In general case, $0 < \lambda < 1$.

For error correction model, Engle and Granger proposed the famous Granger representation theorem in 1987. If variable $X$ and $Y$ are co-integrated, the short-term non-equilibrium relationship between them can always be formulated by an error correction model. To establish an error correction model, we need do co-integration test on variables. When the variables have long-term equilibrium relationship, they can be error correction terms, and then we can establish short-term model by regarding error correction term as an explanatory variable together with other explanatory variables reflecting short-term fluctuation.

Empirical Analysis of Influence Factors of Real Estate Prices in Wuhan

In the real estate market, due to the complexity and diversity of factors affecting the market price, we can’t fully select all factors to do empirical analysis on real estate prices in Wuhan, so we just select possible factors affecting economic fundamental. Therefore, we select real estate data of Wuhan in 2000 to 2012, a total of 13 years, from three angles, supply, demand, and supply-demand comparison to analyze and study.

The second and following paragraph under the first and second level titles (the Introduction part is excluded) need to indented for one character. Please do not forget.

Supply Model of Real Estate Prices Factors in Wuhan. From the microscopic point of view, the more the real estate development investment is, the larger the investment of government or developers is, and the greater the size of the building is. Taking no account of the housing vacancy rate, the real estate development investment is positively correlated with the average price of real estate.

Land price index is based on Wuhan base land prices in 2000 as a base period. As one of the focus of all sectors of society, whether the land price rise really promotes the housing prices has been a troubling problem. According to data from Ministry of Land, China’s land transfer amount in 2009 increased by 63.40% compared to 2008. Though on the current situation, the causal relationship between growth of land and surge of housing prices don't reach a consensus, but in the course of China's real estate price fluctuations in recent years, the rise of land prices has a very important impact on the fluctuation of real estate prices.

It is almost the first choice for contemporary citizen to purchase residence by mortgage loans. So real estate loan interest rate is bound to be a factor affecting housing price fluctuations.

Completed residential area represents the total area of houses that has been completed and can be sold to citizen. The greater the area of completed residential are, the more the house is, and then the housing prices may fall.

Therefore, we introduce the following four variables to describe the supply model of real estate prices factors in Wuhan:

- Real estate development investment (X1): unit/ yuan;
- Land price index (X2);
- RMB loan interest rates rate (X3);
- Completed residential area (X4): unit / ten thousand square meters;

\[
Y_t = 6240.000657 + 0.2698043582X_{3t} - 8.687216467X_{4t} + \epsilon\tag{3}
\]

Then we build a regression model for X5, X6 and X7 to verify the factors that the demand factors have on the real estate prices in Wuhan. However, we find that variable X3 is highly related to X4, without passing the significance testing, so we use Stepwise Regression to obtain the supply model of real estate prices factors in Wuhan, as shown in formula 3 and table1 of the regression result.
Table 1. Regression Model of Supply of the Real Estate Prices Factors in Wuhan.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
<th>F-statistic</th>
<th>Prob (F-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>6240.001</td>
<td>1348.65</td>
<td>4.626848</td>
<td>0.0009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X3</td>
<td>0.269804</td>
<td>0.021491</td>
<td>12.55409</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X4</td>
<td>-8.687216</td>
<td>1.753635</td>
<td>-4.953834</td>
<td>0.0006</td>
<td>99.92798</td>
<td>0</td>
</tr>
</tbody>
</table>

**Demand Model of the Real Estate Prices Factors in Wuhan.** From the microcosmic point of view, the greater the population is, the greater the demand for housing. The population growth is positively correlated with average price of commercial housing.

GDP is an objective indicator reflecting citizen’s payment capacity. Generally speaking, the larger the variable is, the stronger the residents' purchasing ability. GDP had grown eightfold from 2000 to 2012, which was the main reason that stimulated housing consumption to increase rapidly. GDP is positively correlated with housing price. Similarly, the higher the average disposable income is, the stronger the purchasing ability is. The average disposable income is positively correlated with housing price.

Therefore, we introduce three variable indicators to describe model of demand of real estate prices factors in Wuhan:

- Population (X5): unit / ten thousand people
- GDP (X6): unit / one hundred million yuan
- Average disposable income (X7): unit / one yuan

Then we build a regression model for X5, X6 and X7 to verify the factors that the demand factors have on the real estate prices factors in Wuhan. However, we find x7 is insignificant and isn't tested by t, so we use Stepwise Regression to obtain the supply model of real estate prices factors in Wuhan, as shown in formula 4 and table2 of the regression result.

\[
Y_t = 932.2333644 + 1.280298306\times X5_t - 0.3546199108\times X6_t + \epsilon \tag{4}
\]

Table 2. The Regression Result of Demand Model of Real Estate Prices Factors in Wuhan.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
<th>F-statistic</th>
<th>Prob (F-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>932.2334</td>
<td>239.6492</td>
<td>3.889992</td>
<td>0.003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X5</td>
<td>1.280298</td>
<td>0.179799</td>
<td>7.120715</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X6</td>
<td>-0.35462</td>
<td>0.059842</td>
<td>-5.925896</td>
<td>0.0001</td>
<td>124.6097</td>
<td>0</td>
</tr>
</tbody>
</table>

**Supply-demand Model of Real Estate Prices Factors in Wuhan.**

(1) Co-integration Test

We have built two fitting models-supply model and demand model on the real estate prices factors in Wuhan. However, because we need to study the real estate prices factors under supply and demand and as the main factors they have long-term equilibrium relationship. So we did unit root test on all variable, after confirming they are the same order integration, we did co-integration test.
Table 3. The Test Result of the Real Estate Prices Factors in Wuhan(1).  

<table>
<thead>
<tr>
<th>variable</th>
<th>The Test Result</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>critical value</td>
<td>p-value</td>
<td>Whether stationary or not</td>
<td></td>
</tr>
<tr>
<td>real estate development investment</td>
<td>-1.9755</td>
<td>6.053093</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>land price index</td>
<td>-1.9755</td>
<td>0.260499</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>RMB loan interest rates</td>
<td>-1.9755</td>
<td>2.315134</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>Completed area of residential population</td>
<td>-1.9755</td>
<td>0.04843</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-1.9755</td>
<td>3.820809</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>average disposable income</td>
<td>-1.9755</td>
<td>3.068736</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>average price of commercial buildings</td>
<td>-1.9755</td>
<td>-0.97263</td>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>

Table 4. The Test Result of the Real Estate Prices Factors in Wuhan(2).  

<table>
<thead>
<tr>
<th>variable</th>
<th>The Test Result</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>critical value</td>
<td>p-value</td>
<td>original sequence integration order</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>real estate development investment</td>
<td>-1.9835</td>
<td>3.02704</td>
<td>I(2)</td>
<td></td>
</tr>
<tr>
<td>land price index</td>
<td>-1.9835</td>
<td>3.08916</td>
<td>I(2)</td>
<td></td>
</tr>
<tr>
<td>RMB loan interest rates</td>
<td>-1.9835</td>
<td>3.33204</td>
<td>I(2)</td>
<td></td>
</tr>
<tr>
<td>Completed area of residential population</td>
<td>-1.9835</td>
<td>-2.28956</td>
<td>I(2)</td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>-1.9835</td>
<td>-1.34904</td>
<td>I(2), d&gt;2</td>
<td></td>
</tr>
<tr>
<td>average disposable income</td>
<td>-1.9835</td>
<td>-2.10536</td>
<td>I(2)</td>
<td></td>
</tr>
<tr>
<td>average price of commercial buildings</td>
<td>-1.9835</td>
<td>-3.07811</td>
<td>I(2)</td>
<td></td>
</tr>
</tbody>
</table>

According to table 3,4, we can know the original series are all non-stationary series except GDP, and other variables are all integrated of order 2. Therefore, we get rid of variable X6(GDP) and do co-integration test on other variables.

\[ Y_t = 5.169608791 - 0.06264861167 \times X_{3t} + 0.175995645 \times X_{5t} + \epsilon \]  

(5)

Firstly, we adopt OLS least square estimation on these variables, and then use stepwise regression to obtain model formula like formula 5 and output result table 5.

Table 5. Test Result of Stepwise Regression.  

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
<th>F-statistic</th>
<th>Prob (F-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>5.169609</td>
<td>59.77823</td>
<td>0.08648</td>
<td>0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X3</td>
<td>-0.062649</td>
<td>0.066128</td>
<td>0.947384</td>
<td>0.0012</td>
<td>98.86131</td>
<td>0</td>
</tr>
<tr>
<td>X5</td>
<td>0.175996</td>
<td>0.263554</td>
<td>0.667778</td>
<td>0.0231</td>
<td></td>
<td>98.86131</td>
</tr>
</tbody>
</table>

Then we test whether the residual sequence \( \epsilon_t \) is stationary or not. The value of ADF is -3.661137 and is less than critical value -3.2695, which can show that variable indexes reserved by stepwise regression, included RMB loan interest rates, population has long-term equilibrium relationship(co-integration relationship) with price of commercial buildings.
As we can see from long-term equilibrium model (formula 5), on the demand side, population has long-term equilibrium relationship with price of commercial building. When the population adds one unit, the prices of commercial buildings add one unit. The larger population is, the more the housing in need is, thereupon, the higher the price of commercial building is.

On the demand side, RMB loan interest rate has long-term equilibrium relationship with price of commercial building. When the RMB loan interest rate adds one unit, the price of commercial building decreases about 0.06 unit. The higher the RMB loan interest rate, the larger the house-purchase burden for residents, the more insufficient the supply is.

Therefore in the long term, price of commercial building in Wuhan is supported by supply and demand, and the effect of interest rate is relatively more notable. In order to curb the growth in house price, we can add supply properly to make the price of commercial building fall a little automatically under the interaction of supply and demand.

(2) Build ECM

In order to reduce multicollinearity caused by distributed-lag model, we continue to build ECM to reflect short-term regulation behavior and long-term equilibrium relationship between variables. According EG method of Engle-Granger, we can obtain the ECM formula formula 6, and output result table6.

\[
Y_t = 2.0265343 - 0.133207X_{3t} + 0.2973095247X_{5t} - 0.6122581192ECM_{t-1} + \epsilon
\]  

Table 6. Test Result of ECM.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Std. Error</th>
<th>t-Statistic</th>
<th>Prob.</th>
<th>F-statistic</th>
<th>Prob (F-statistic)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>2.026534</td>
<td>99.94315</td>
<td>0.020277</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>X3</td>
<td>-0.133207</td>
<td>0.122567</td>
<td>1.086815</td>
<td>0.0382</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>X5</td>
<td>0.29731</td>
<td>0.43051</td>
<td>0.690598</td>
<td>0.0078</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ECM(-1)</td>
<td>-0.612258</td>
<td>1.038253</td>
<td>-0.5897</td>
<td>0.0001</td>
<td>99.92894</td>
<td>0</td>
</tr>
</tbody>
</table>

According to short-term regulation relation, we can know ECM represents short-term deviation of population and RMB loan interest rate, and its absolute value of coefficient represents the speed that the price departure.

When short-term population changes one unit, the price of commercial building in Wuhan will adjust about 0.13 unit; when the short-term RMB loan interest rate changes one unit, the price of commercial building in Wuhan will adjust about 0.3 unit; when the previous price of commercial building departures one unit, the current price will fluctuate about 0.61 unit.

References


