Advances in Engineering Research, volume 152
6th International Conference on Machinery, Materials and Computing Technology (ICMMCT 2018)

Investment Trend and Influential Factors Research in Xi'an Real Estate Development - Based on BACE

Shao Zhiwei
College of General Studies, Eurasia University, Xi'an, ShaanXi Province, 710065, China

Keywords: Real estate investment(REINV), Classic Bayes estimate, Error correction model

Abstract: According to the macroeconomic data in Xi'an real estate industry from 1988 to 2016, the long-term development and short-term fluctuations of real estate investment (REINV) conform to the following rules: firstly, continue their own developmental inertias; secondly, the impact on Xi'an REINV mainly comes from the comprehensive effect of social fixed investment, real estate stock and per capita consumption expenditure. In the long-term equilibrium model, per capita consumption expenditure and social fixed investment have a greater impact on the development investment, and are positively correlated - 0.67 units changes in per capita consumption expiture will cause an unit change in REINV. Meanwhile, real estate stock is negatively correlated to REINV. Viewed from the short term, the simultaneous fluctuations in per capita consumption expenditure, social fixed investment and real estate stock these three factors will certainly bring about big changes in the current investment on Xi'an real estate development, and will be balanced within two years.

1. Introduction

Nowadays China economic development is in a take-off stage. The proportion of real estate investment (REINV) in the total investment in fixed assets is gradually increasing, and the industrial added value is increasing as well. The real estate industry has played an important role in stimulating economic growth. The investment in real estate development has increased so fast, from 317.84 billion yuan in 1997 to 10,979.9 billion yuan in 2017. Therefore, how to grasp the REINV development trend, how to make macro-control and how to ensure a healthy and stable development of real estate industry are the main problems faced by real estate developers and government functional departments.

2. Research Status and Indicator Selection on REINV

The driving factors at home and abroad related to REINV or housing investment mainly focus on economic factors, social factors, industrial development and residential consumption and income. The researches studied by domestic scholars Han Jia[1], Liu Hong[2], M. Ball、T. Morrison [3] and others mainly focus on social fixed assets investment and per capita GDP. In domestic and foreign researches, the influence of social factors on REINV is basically the same, and the selection of social factors is urbanization rate [8]. The indicator commonly used by domestic scholars to study industrial development is real estate stocks, while the indicators of residential consumption and income are per capita consumption expenditure and disposable income of urban residents. The main policy factors affecting REINV are real estate industry policy, financial policy and taxation policy. Legal factors mainly refer to the laws and regulations related to real estate industry[5]. Since policy factors and legal factors cannot be measured in quantitative analysis, they are not included in the explanatory variable.

3. The Theoretical Method of REINV

3.1 Causality test

For two variables Y_t and X_t, make a regression as follows:

$Y_t = \sum_{j=1}^{m} \alpha_j X_{t-j} + \sum_{j=1}^{m} \beta_j Y_{t-j} + \mu_{1t}$ and $X_t = \sum_{j=1}^{m} \lambda_j X_{t-j} + \sum_{j=1}^{m} \delta_j X_{t-j} + \mu_{2t}$

Statistics:

$$G = \frac{(RSS_R - RSS_U)/m}{RSS_U/(n-k)} \tag{1}$$

If $G \geq F_\alpha(m, n-k)$, reject the null hypothesis $H_0: \beta_j = 0\ (\forall j)$ and accept the alternative hypothesis $H_1: \beta_j \neq 0\ (\exists j)$. That is, Y_t is the Granger cause for X_t.[7][8]

3.2 Classic Bayesian average estimation model

Assume that the number of independent variables is n, so the total number of regression models is $2^n - 1$. BACE is based on the model: $2^n - 1$[6].

Posterior probability in model K_j:

$$p(K_j/Y) = \frac{p(K_j)T^{-M_j/2}RSS_j^{-T/2}}{\sum_{i=1}^{2^n-1} p(K_i)T^{-M_i/2}RSS_i^{-T/2}} \tag{2}$$

Where $P(K_j) = C_n^{M_j} p^{M_j}(1-p)^{n-M_j}$, M_j is the number of explanatory variables selected in the model, T represents sample size. The mean of regression coefficient corresponding to explanatory variables is

$$E(\beta/Y) = \sum_{j=1}^{2^n-1} P(K_j/Y)\hat{\beta}_j \tag{3}$$

Posterior standard deviation:

$$Var(\beta/Y) = \sum_{j=1}^{2^n-1} P(K_j/Y)Var(\beta/Y,K_j) + \sum_{j=1}^{2^n-1} P(K_j/Y)(\hat{\beta}_j - E(\beta/Y))^2 \tag{4}$$

3.3 Stepwise regression method on testing multicollinearity

Stepwise regression is to introduce the variables one by one to the model. For each explanatory variable, performing an F test and make t-test for each selected explanatory variable. When the previously introduced explanatory variable becomes no longer significant due to the introduction of the later one, it indicates that there is a variable causing severe collinearity. At this point, make multifaceted comparison and retain the optimal variable. At the same time, culling other variables which cause severe collinearity to ensure that regression equation keeps the significant variables before new variable is introduced[4].

3.4 Cointegration test

If a time series to undergo d times differences before being a stable sequence, this series is called d- order single integer, recorded as $I(d)$. If a group of series $Y(t), X_1(t), X_2(t), \cdots\cdots, X_k(t)$ are all d-order single integer, $\alpha X'(t) \sim I(a-b)$, where $b > 0$, $\alpha = (a_0, a_1, a_2, \cdots\cdots, a_k)$, $X(t) = (Y(t), X_1(t), X_2(t), \cdots, X_k(t))'$, then this group is called (d,b)-order cointegration[6].

3.5 Autoregressive distribution lag model

$$Y(t) = \beta_0 + \sum_{i=1}^{m} \delta_i Z(t-i) + \sum_{i=1}^{n} \gamma_i Y(t-i) + \varepsilon_t \tag{5}$$

It reflects the long-term trend of explanatory variables, and the influential intensity of it on the variable being explained.

3.6 Error correction model

$$\Delta Y(t)=\tilde{\beta}_0+\sum_{i=1}^{m}\tilde{\delta}_i\Delta Z(t-i)+\sum_{i=1}^{n}\tilde{\gamma}_i\Delta Y(t-i)+k(\hat{Y}(t)-b_0-\sum_{l=1}^{k}b_l X_l(t))+\varepsilon_t \quad (6)$$

where $\varepsilon_t \sim N(0,\sigma^2)$.

Explained variables' fluctuations in the model can be divided into two parts: long-term equilibrium and short-term fluctuation. The difference items in each section represent the effect of explanatory variables' short-term fluctuations.

4. An Empirical Analysis of the Dynamic Factors in REINV——Taking Xi'an as an Example

4.1 Analysis of the Dynamic Factors Affecting REINV

4.1.1 ADF test

When the second-order difference of Xi'an REINV is in the first-order lag, $\text{ADF}=-8.1560$.

When residual sequence is in the first-order lag, $\text{ADF}=-8.1560$<-2.6924. It is a stationary sequence at a confidence level of 0.01.

4.1.2 Causality test between REINV and urbanization rate in Xi'an

$$G=\frac{(RSS_R-RSS_U)/m}{RSS_U/(n-k)}=8.3108>F_{0.05}(2,31)=2.58$$

Therefore, urbanization rate is a Granger cause of *LNY*[5]. The causality test method for relationship of *LNY* and other factors is the same, which will not be described here.

Table 1 The Causality Test between Xi'an REINV and dynamic factors

	ADF	Confidence Level	Stability	*G*	*F*	Results
Xi'an REINV(*LNY*)	-8.3470	0.01	Stable	/	/	/
Urbanization Rate*100(*X*1)	-6.8437	0.01	Stable	8.3108	$F(2,31)$	cause
Real Estate Stock (*LNX*2)	-5.4361	0.01	Stable	5.6161	$F(1,31)$	cause
Per Capita GDP(*LNX*3)	-4.6572	0.05	Stable	7.4804	$F(2,31)$	cause
Per Capita Disposable Income(*LNX*4)	-4.0768	0.05	Stable	6.1376	$F(1,31)$	cause
Per Capita Deposit Balance(*LNX*5)	-6.3214	0.01	Stable	4.6573	$F(2,31)$	cause
Per Capita Consumption Expenditure(*LNX*6)	-5.3216	0.01	Stable	5.4313	$F(2,31)$	cause
Social Fixed Assets Investment(*LNX*7)	-5.0057	0.01	Stable	6.9223	$F(2,31)$	cause

4.2 Apply Classic Bayesian average estimation model to LNY

According to BACE calculation formula, the following results are obtained:

Table 2 Analyzing results based on BACE method

Variable	Coverage Probability	Conditional Mean	Conditional Standard Deviation	Symbol determination rate	t-test significant
*LNX*7	0.99773	1.87688	1.13769	1.00000	1.00000
*LNX*2	0.96435	-2.33442	0.80293	0.88754	0.78824
*LNX*6	0.47522	1.37646	0.86224	0.52350	0.49738
*LNX*4	0.45126	-0.53021	0.33216	0.674532	0.17683
*LNX*5	0.27563	0.02765	0.67497	0.70546	0.22290
*LNX*3	0.25332	0.15034	0.25673	0.72353	0.58845
*X*1	0.27654	0.01652	0.03164	0.37497	0.25425

In table 2, coverage probability is the sum of posterior probabilities in all 2^7-1 models containing the variable, calculated by formula 2 and reflecting the strength of explanatory variables. Conditional mean equals to the formula 3 divided by coverage probability of explanatory variable. The robustness of explanatory variables is determined by symbol determination rate and t-test significance. The first type of explanatory variables include social fixed investment and real estate stock, of which the coverage probabilities are 0.99773 and 0.96435 respectively. Their posterior probabilities which should be included in the model are greater than their prior probabilities. The information contained in the data further reinforces the belief that the model should contain these two explanatory variables. In addition, the symbol determination rates of them are 1 and 0.88754, and t-tests significance are 1 and 0.78824, which indicates a better robustness. The second type of explanatory variables include per capita consumption expenditure and per capita disposable income, presenting that these two variables also have a strong explanatory power for REINV. While among the seven explanatory variables, the t-test significance of per capita disposable income is the lowest, which is only 0.17683. The multicollinearity among explanatory variables is very strong, thus affects t-test significance certainly. The third type of explanatory variables include the per capita deposit balance, per capita GDP, and urbanization rate. Their coverage probability and robustness are lower than the first four explanatory variables no matter from which point of view, so prudent consideration should be given. Therefore, obtained by BACE method, *LNX*7 and *LNX*2 are the main driving forces of *LNY*, *LNX6* and *LNX4* are the secondary influencing factors, and *LNX5*, *LNX3* and *X1* belong to weak factors.

4.3 Stepwise regression method to test multicollinearity

Stepwise regression is to introduce other explanatory variables into the above regression model so as to find the best regression equation[3].

Table 3 Regression form for LNY, X1, LnX2, LNX3, LNX4, LNX5, LNX6 and LNX7

	C	*LNX*7	*LNX*2	*LNX*6	*LNX*5	*LNX*4	*LNX*3	*X*1	$\overline{R}^2$	D.W
LNY=f(*LNX*7)	-8.75	1.35							0.96	0.93
t value	-7.93	22.4								
LNY=f(*LNX*7，*LNX*2)	6.83	2.07	-2.59						0.97	1.78
t value	1.73	9.78	-3.49							
LNY=f(*LNX*7，*LNX*2，*LNX*6)	6.86	1.81	-2.60	0.44					0.98	1.94
t value	1.81	7.06	-3.65	1.59						
LNY=f(*LNX*7，*LNX*2，*LNX*4)	9.11	1.88	-2.77		0.23				0.98	1.94
t value	2.09	7.20	-3.69		1.17					
LNY=f(*LNX*7，*LNX*2，*LNX*6，*LNX*4)	5.84	1.84	-2.53	2.33		-1.96			0.98	2.15
t value	1.53	7.30	-3.61	1.64		-1.35				
LNY=f(*LNX*7，*LNX*2，*LNX*6，*LNX*4，*LNX*3)	4.77	1.92	-2.42	7.70		-1.74	-0.60		0.98	2.18
t value	1.04	6.01	-3.18	1.60		-1.20	-0.44			
LNY=f(*LNX*7，*LNX*2，*LNX*6，*LNX*4，*LNX*3，*LNX*5)	4.71	1.92	-2.41	2.70	-0.01	-1.73	-0.59		0.98	2.18
t value	0.77	5.63	-2.94	1.54	-0.01	-1.14	-0.37			
LNY=f(*LNX*7，*LNX*2，*LNX*6，*LNX*4，*LNX*3，*LNX*5, *X1*)	4.87	1.96	-2.44	2.72	-0.02	-1.84	-0.58	0.004	0.97	2.15
t value	0.78	5.24	-2.87	1.50	-0.03	-1.15	-0.35	0.33		

Introducing *LNX2* to the initial model, the model goodness-of-fit is improved, the parameter symbols are reasonable,and the explanatory variables pass the t-test as well. Then introducing *LNX4* and *LNX6* separately, the fitting degree is improved once again, at the same time its symbol is reasonable. Only D.W value falls into an undeterminable area, but it is necessary to determine whether there is autocorrelation through LM test. There is no obvious changes in the fitting degree when introducing other explanatory variables, and each step happens that some of the explanatory variables cannot pass t-test. Similarly, for any combination group of three or more explanatory variables, there are always variables that cannot pass t-test. Therefore, based on BACE analysis and in order to eliminate the collinearity, only three explanatory variables can be selected as the optimal, which are social fixed investment*(LNX7*), real estate stock (*LNX2)* and per capita consumption expenditure (*LNX4*).

4.4 Cointegration Analysis of Xi'an REINV and Various Dynamic Factors

According to the cointegration theory, Johansen test is applied to get the results: The maximum eigenvalue test shows that there are four cointegration equations at the 0.05 significance level between Xi'an REINV and its Granger reasons.

Table 4 Cointegration test among LNY, LnX2, LNX4 and LNX7

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.825890	88.39335	47.85613	0.0000
At most 1 *	0.639809	41.19546	29.79707	0.0016
At most 2	0.360099	13.62517	15.49471	0.0938
At most 3	0.056533	1.571229	3.841466	0.2100

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None *	0.825890	47.19790	27.58434	0.0001
At most 1 *	0.639809	27.57029	21.13162	0.0054
At most 2	0.360099	12.05394	14.26460	0.1086
At most 3	0.056533	1.571229	3.841466	0.2100

The cointegration test includes two parts: trace statistics test and maximum eigenvalue test. Judging by the test level of 0.05, the trace statistics test has the results of 88.39>47.86, 41.95>29.76, 13.63<15.49, and the maximum eigenvalue test has the results of 47.20>27.58, 27.57<21.13, 12.05<14.26. Hence, therefore are two long-term cointegration equations between four variables of *LNY*, *LNX4*, *LNX7* and *LNX3* in the case of intercept and trend terms existence.

4.5 Autoregressive distributed lag model (ARDL) in Xi'an REINV and dynamic factors.

$$
\begin{aligned}
&LNY_t = -6.07 + 0.12 * LNY_{t-2} - 0.06 * LNX2_t + 1.20LNX4_t + 0.52LNX7_t \\
&+[MA(1) = 0.9627, BACKCAST = 1990] \qquad (7) \\
&T = (-8.5230, 1.6895, -1.6738, 5.0058, 3.6489, 105.9486) \\
&R^2 = 0.99974, F = 1632.755, D.W. = 1.9232, AIC = -1.1109, SC = -0.8230
\end{aligned}
$$

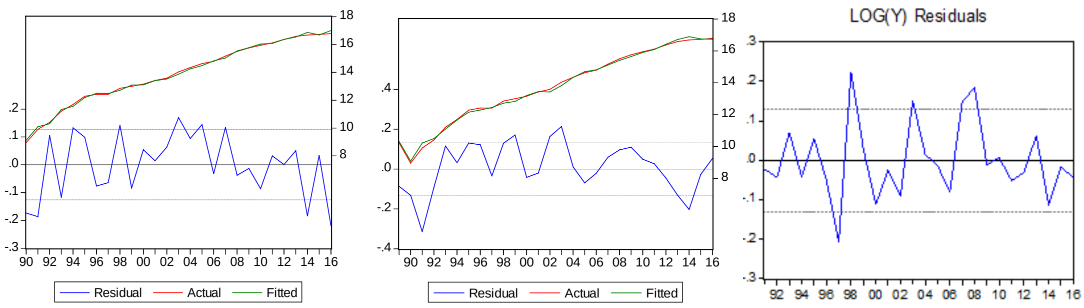
4.6 Long-term equilibrium model of Xi'an REINV and dynamic factors

$$
\begin{aligned}
&LNY_t = -5.90 - 0.05LNX2_{t-1} + 0.67LNX4_t + 0.87LNX7_t \\
&+[MA(2) = 0.9806, BACKCAST = 1989] \qquad (8) \\
&T = (-13.14418, 2.494131, 3.622721, 9.063560, 56.81093,) \\
&R^2 = 0.9972, Loglikelihood = 19.8595, D.W. = 0.9986, AIC = -1.0614, SC = -0.8235
\end{aligned}
$$

4.7 Error correction model of Xi'an REINV and Dynamic Factors

$$
\begin{aligned}
&\Delta LNY_t = -0.07(LNY_{t-1} + 0.11LNX4_{t-1} - 0.47LNX7_{t-1} - 5.09) \\
&-0.097(LNX2_{t-1} + 6.45LNX4_{t-1} + 6.697LNX7_{t-1} - 2.16 * @TREND - 133.397) \\
&-0.22\Delta LNY_{t-1} - 0.0005\Delta LNY_{t-2} + 0.0003\Delta LNX2_{t-1} + 0.09\Delta LNX2_{t-2} + 0.81\Delta LNX4_{t-1} + 0.67\Delta LNX4_{t-2} \\
&+0.56\Delta LNX7_{t-1} + 0.29\Delta LNX7_{t-2} - 0.04 \qquad (9) \\
&R^2 = 0.8070, F = 6.2729, AIC = 0.9682, SC = -1.3303, Loglikelihood = 119.7407
\end{aligned}
$$

The actual values, fitted values, and residual sequences in model 7, 8, and 9 refer to figure 1.



ARDL Model Long-term Equilibrium Model Error Correction Model

Figure 1

Notes: The data used above (excluding generated data) is the standardized data of original.

Source: National Bureau of Statistics, China Statistical Yearbook 1988-2016. China Population (and Employment) Statistical Yearbook 1988-2016.

5. Conclusion

Through the analysis above, the long-term development and short-term fluctuations of real estate investment (REINV) conform to the following rules: firstly, continue their own developmental inertias; secondly, the impact on Xi’an REINV mainly comes from the comprehensive effect of social fixed investment, real estate stock and per capita consumption expenditure. In the long-term equilibrium model, real estate stock has the biggest influence on Xi’an REINV, and is negatively correlated. In the short-term fluctuation model, the fluctuations of social fixed assets investment and per capita consumption expenditure in the early stage will certainly bring about a large-scale fluctuations in Xi’an current REINV, and at the same time, social fixed assets investment in Xi’an REINV presents a strong dependence with per capita consumption expenditure. It also shows that when the REINV in Xi’an is higher than the equilibrium value in the previous period, the REINV will be lower than the equilibrium value in the current period. Conversely, when REINV in the previous period is lower than the equilibrium value, the REINV will be higher than equilibrium value in the current period. This kind of balance will control Xi'an REINV in the long-term equilibrium model based on the relation of REINV with social fixed assets investment, real estate stock, and per capita consumption expenditure, as the proportion of 0.29, 0.099, 0.67 in the previous two years, and 0.56, 0.0003, 0.081 in the previous year, and finally gets balance within two years.

Acknowledgement

Project: Yueqian Science and Technology Industry- University Cooperative Education Project of Ministry of Education in 2017(Project No. 201702071032)

References

[1] Han Jian. The impact of real estate industry on the national economic development[J]. Journal of Hohai University (Philosophy and Social Science Edition).2004,((l):24-27.

[2] Liu Hong.Dynamic Economic Effects of Urban Real Estate Investment in China[J].Economics and Management Research.2006,(3):49-53

[3] M.Ball,T.Morrison.Housing Investment Fluctuation:An International comparison[J].Housing, Theory and Society,2000,17(1):3-13.

[4] Wang Songtao, Liu Hongyu. Theoretical research and empirical analysis of real estate development investment level [J]. Construction Economy.2006,(6):60-63.

[5] Yu Mingxuan. Analysis of Real Estate Investment [M]. Beijing: Capital University of

Economics and Business Press.2004.

[6] Li Zinai, Pan Wenqing. Econometrics [M]. Beijing: Higher Education Press.2005.

[7] He Jingtong, Li Feng. Factors Affecting Independent Innovation——Analysis of Chinese Inter-provincial Data Based on BACE Method[J].Nankai Economic Research.2007,(3):68-79.

[8] Wang Songtao, Chen Wei, Chen Yi, Fang Ming.Study on the Influencing Factors of Real Estate Development Investment Level——Comparative Study on Real Estate Investment and Development in Beijing, Shanghai, Tianjin and Shenzhen[J]. Construction Economy.2007, (8): 17-21.

[9] Xiang Weimin, Li Jiao. Research on the relationship between urbanization level and real estate investment growth rate [J]. Journal of Chongqing Jianzhu University.2007,(29)106-109.

[10] He Zongln. Maekawa Koichi. On spurious Grange Causality[J]. Economic Letter, 2001, 73: 307-313.