Study on the Spillover Effect of Housing Price in Yangtze River Delta
Based on VAR Model

Meng-meng BAI*
Shanghai University, Baoshan district, Shanghai, China
*Corresponding author

Keywords: Housing Price, Spillover Effect, VAR Model, Co-integration Method, Cumulative Effect.

Abstract. This paper analyzed housing prices’ spillover effect of four cities. Through theoretical model and empirical analysis, we can conclude that there is a certain correlation between the housing prices, and Shanghai’s housing price fluctuation has a greater impact on other cities. Nanjing and Hefei have a strong correlation and spillover effect with each other but the spillover effect on Hangzhou is not obvious. Hangzhou can produce larger and more durable influence on its own. This paper therefore is of great practical significance in formulating real estate policies.

Introduction

China’s property price has been rising at a modest rate of 3.5% since the reform in 1998. But during 2003 to 2008, it maintained a high growth rate of nearly 10.5%, and in 2016, the overall increase in house prices was 13 times. Although property price has been regulated, prices still maintain strong growth momentum and show regional clustering. The eastern coastal region such as the Yangtze river delta has become the most obvious regional effects of the nationwide housing market.

The Yangtze river delta is one of China’s important economic advancement area, which in total includes 24 cities. The house pricing fluctuations has consistency in one region, but the relevance among this fluctuation trend is not clear, so in this paper we will study the cumulative effect and the spillover effect among four main cities of Yangtze river delta.

Literature Review

Through the calculation of housing prices in the UK in 1960-1990, Holmans AE [1] found that the first change in house prices happened in London-based southern region and thus affecting the entire British. Van Steenkiste I, P Hiebert [2] proposed that when the regional housing prices show a strong volatility, the linkage then show the spillover effect. Wu Weiwei [3] defines the spillover effect as "the conduction and the degree of interaction of the price fluctuation between two regions". In the empirical study on the housing price linkage and spillover effect, British scholars used the grander causality test and co-integration method and found that house prices have significant correlation. After 2005, most scholars used VAR, GVAR model and CCE estimation method to test the spillover effect in the US. Costello [4] used the eight VAR models to estimate the equilibrium value of house prices in Australia, and then used the residual as a measure of house price deviation. Yu Yihua [5] used GVAR model on the study of spillover effect and the impact of housing prices on the cross-regional impact of inflation.

Theoretical Model

The VAR model was introduced by Christopher Sims in 1980. It uses current variables to make regression analysis and to estimate the dynamic relationship of the joint endogenous variables without any prior constraint. The mathematical model of unconstrained VAR model is:

$$y_t = A_1y_{t-1} + A_2y_{t-2} + \cdots + A_py_{t-p} + bx_t + \epsilon_t$$ (1)
$y_t$ is a $K$-dimensional endogenous variable, $x_t$ is the $D$-dimensional exogenous variable, and $P$ is the lag order. $\varepsilon_t$ is a $K$-dimensional perturbation vector, which can be auto-correlated. The shock variable $\varepsilon_t$ is a white noise variable, cause $\varepsilon_t$ has no structural meaning.

**Empirical Analysis**

The data is chosen from Huatai database, and the samples were monthly real estate prices of Shanghai, Hangzhou, Nanjing and Hefei. Fixed price index is set to eliminate inflation influence.

As is shown in Figure 1, the housing prices trend in three cities except Hangzhou has a strong dynamic consistency. This verifies spillover effect to a certain extent.

![Figure 1. Price trend in 2001-2014.](image)

VAR model is based on smooth time series data in order to avoid Pseudo-regression phenomenon. Logarithm does not change the nature of the original data, and it can eliminate the heteroscedasticity. The results of ADF unit root method for the stability are shown in Table1.

<table>
<thead>
<tr>
<th>variable</th>
<th>ADF Statistics</th>
<th>1% Critical value</th>
<th>5% Critical value</th>
<th>10% Critical value</th>
<th>conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>ERPsh</td>
<td>-0.227630</td>
<td>-4.057910</td>
<td>-3.119910</td>
<td>-2.701103</td>
<td>Unstable</td>
</tr>
<tr>
<td>ΔERPsh</td>
<td>-3.190667</td>
<td>-4.057910</td>
<td>-3.119910</td>
<td>-2.701103</td>
<td>Stable</td>
</tr>
<tr>
<td>ERPhz</td>
<td>-0.630625</td>
<td>-4.057910</td>
<td>-3.119910</td>
<td>-2.701103</td>
<td>Unstable</td>
</tr>
<tr>
<td>ΔERhz</td>
<td>-3.014933</td>
<td>-4.057910</td>
<td>-3.119910</td>
<td>-2.701103</td>
<td>Stable</td>
</tr>
<tr>
<td>ERPhf</td>
<td>1.973561</td>
<td>-3.646342</td>
<td>-2.954021</td>
<td>-2.615817</td>
<td>Unstable</td>
</tr>
<tr>
<td>ΔERPhf</td>
<td>-3.364363</td>
<td>-3.646342</td>
<td>-2.954021</td>
<td>-2.615817</td>
<td>Stable</td>
</tr>
<tr>
<td>ERPnj</td>
<td>0.095926</td>
<td>-4.057910</td>
<td>-3.119910</td>
<td>-2.701103</td>
<td>Stable</td>
</tr>
<tr>
<td>ΔERPnj</td>
<td>-2.861509</td>
<td>-4.057910</td>
<td>-3.119910</td>
<td>-2.701103</td>
<td>Stable</td>
</tr>
</tbody>
</table>

The sequence of ERPsh, ERPhz, ERPhf and ERPnj is non stationary because the ADF value is greater than 1%, 5% and 10%. Therefore the original hypothesis cannot be rejected. The above variables respectively are taken the first difference, in which the ADF value of ΔERPsh is less than 1%, so it is stable. The ADF value of ΔERhz, ΔERPnj is less than 10%, ΔERPhf is less than 5%. The first order differential sequence is a positive time series, therefore, these four variables satisfy the conditions for constructing the co-integration equation. First, the VAR equation was established, and the second order lag was selected according to the SIC criterion.
\[
\begin{pmatrix}
\text{ERP}_{sh} \\
\text{ERP}_{hz} \\
\text{ERP}_{hf} \\
\text{ERP}_{nj}
\end{pmatrix}
= 
\begin{pmatrix}
0.037 \\
-0.27 \\
0.08 \\
0.016
\end{pmatrix}
+ 
\begin{pmatrix}
0.37 & -0.27 & 0.08 & 0.016 \\
-0.08 & 0.64 & 0.54 & -0.29 \\
-0.15 & 0.006 & 0.2 & 0.61 \\
0.25 & 0.012 & 0.59 & 0.81
\end{pmatrix}
\begin{pmatrix}
\text{ERP}_{sh_{t-1}} \\
\text{ERP}_{hz_{t-1}} \\
\text{ERP}_{hf_{t-1}} \\
\text{ERP}_{nj_{t-1}}
\end{pmatrix}
+ 
\begin{pmatrix}
0.037 \\
-0.27 \\
0.08 \\
0.016
\end{pmatrix}
+ 
\begin{pmatrix}
0.37 & -0.27 & 0.08 & 0.016 \\
-0.08 & 0.64 & 0.54 & -0.29 \\
-0.15 & 0.006 & 0.2 & 0.61 \\
0.25 & 0.012 & 0.59 & 0.81
\end{pmatrix}
\begin{pmatrix}
\text{ERP}_{sh_{t-2}} \\
\text{ERP}_{hz_{t-2}} \\
\text{ERP}_{hf_{t-2}} \\
\text{ERP}_{nj_{t-2}}
\end{pmatrix}
+ 
\begin{pmatrix}
\varepsilon_{1t} \\
\varepsilon_{2t} \\
\varepsilon_{3t} \\
-\varepsilon_{4t}
\end{pmatrix}
\text{(2)}
\]

The AIC value is 4.2 and the SC value is 5.8, and the goodness of the five equations is better which respectively are 0.92, 0.573, 0.776, 0.91. Then, we carry out Johansen co-integration test.

Table 2. Johansen co-integration test results.

<table>
<thead>
<tr>
<th>Null hypothesis</th>
<th>Trace statistics</th>
<th>Trace statistic threshold</th>
<th>Maximum eigenvalue</th>
<th>Trace statistic threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Co-variance equation</td>
<td>5% P value</td>
<td>10% P value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>63.55020</td>
<td>47.85613</td>
<td>0.0009</td>
<td>32.95681</td>
</tr>
<tr>
<td>&lt;=1</td>
<td>30.59340</td>
<td>29.79707</td>
<td>0.0404*</td>
<td>16.34482</td>
</tr>
<tr>
<td>&lt;=2</td>
<td>14.24857</td>
<td>15.49471</td>
<td>0.0764</td>
<td>13.13356</td>
</tr>
<tr>
<td>&lt;=3</td>
<td>1.115016</td>
<td>3.841466</td>
<td>0.2910</td>
<td>1.115016</td>
</tr>
</tbody>
</table>

The trace statistics and the maximum characteristic quantity value of the Johansen co-integration test show that the original hypothesis is rejected at 5% and that 1 or 2 associations exist. Therefore, there is a long-term co-integration relation between \( \Delta \text{ERP}_{sh}, \Delta \text{ERP}_{hz}, \Delta \text{ERP}_{hf}, \Delta \text{ERP}_{nj} \).

\[
\Delta \text{ERP}_{sh} = 0.253 \Delta \text{ERP}_{hz} + 0.280559 \Delta \text{ERP}_{nj} + 0.432474 \Delta \text{ERP}_{hf} - 14.47
\text{(3)}
\]

It can be deduced that Shanghai’s housing price in the long term has a greater negative impact on Hefei and Nanjing, while housing price of Nanjing have a positive impact on Hangzhou.

The co-integration relationship can be established after the error correction model, with the hysteresis of the residual term ECM as the error correction term, the regression equation can be set as shown in the Table 3. The error correction model can be used to correct the short-term deviation.

Table 3. Error correction results.

<table>
<thead>
<tr>
<th>Error Correction:</th>
<th>D ( \Delta \text{ERP}_{sh} )</th>
<th>D ( \Delta \text{ERP}_{hz} )</th>
<th>D ( \Delta \text{ERP}_{hf} )</th>
<th>D ( \Delta \text{ERP}_{nj} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.09932</td>
<td>0.018376</td>
<td>0.041202</td>
<td>0.024668</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{sh} ))</td>
<td>-0.137718</td>
<td>-0.596567</td>
<td>0.057415</td>
<td>-0.193749</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{hz} ))</td>
<td>-0.079243</td>
<td>0.657400</td>
<td>-0.171039</td>
<td>0.298936</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{hf} ))</td>
<td>-0.237630</td>
<td>0.437468</td>
<td>-0.087643</td>
<td>-0.088465</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{nj} ))</td>
<td>0.040436</td>
<td>0.236064</td>
<td>-0.036515</td>
<td>0.096854</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{nj} ))</td>
<td>0.044613</td>
<td>-0.161200</td>
<td>-0.004335</td>
<td>0.093286</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{nj} ))</td>
<td>0.477888</td>
<td>-2.944619</td>
<td>-0.440687</td>
<td>-0.748878</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{nj} ))</td>
<td>0.301979</td>
<td>-1.604848</td>
<td>-0.193634</td>
<td>-0.414741</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{nj} ))</td>
<td>0.640530</td>
<td>0.268374</td>
<td>0.532371</td>
<td>0.325377</td>
</tr>
<tr>
<td>D(( \Delta \text{ERP}_{nj} ))</td>
<td>-0.125104</td>
<td>0.414499</td>
<td>0.107745</td>
<td>0.218115</td>
</tr>
</tbody>
</table>

The error correction of each equation is negative, in line with the reverse correction mechanism. And \( \Delta \text{ERP}_{sh}, \Delta \text{ERP}_{hz}, \Delta \text{ERP}_{hf}, \Delta \text{ERP}_{nj} \) will adjust the deviation at the ratio of \(-0.137718, -0.596567, 0.057415, -0.193749 \) when each variable deviates from the long-term equilibrium.

We already know there is a long-term stable co-integration between the first-order differences variables. Therefore we use Granger causality to study the causal relationship.
This test was based at significance level of 10%, therefore we found that Shanghai’s housing prices change is the Granger reasons for Hefei and Nanjing, but not for Hangzhou.

Likewise, Hefei is the Granger reasons for Nanjing, Nanjing is the cause of Shanghai.

The impulse response function describes the impact of an endogenous variable has on the other endogenous variables, showing the reactions after shock. Figure2 is the corresponding impulse response curve, the horizontal axis represents the number of traces, and the vertical axis represents the dependent variables’ response degree to explanatory variables. The tracking period is set as 25.

It is shown in Fig.2 that Hangzhou, Hefei and Nanjing have positive response, but from 4th period, Shanghai’s housing price reaction gradually weaken, and at 20th period it tends to stabilize. This indicating that the housing prices in Shanghai is both short-term and long-term factor for other cities. Hangzhou’s housing prices produced a positive impact on Shanghai, Hefei, Nanjing, and the duration is long. Hefei’s housing price shock has a short-term strong impact on other cities. Likewise Nanjing’s house price has a positive impact.

Then we decompose the variances. The results are shown in Fig 3. Shanghai’s house price fluctuations have an impact on its future changes to which degree it reached 100%, but later it maintains a stable level of 20%. Hefei’s housing price declines fast, which remained basically stable after 3th period. 90% of Hangzhou’s early rising house prices results from itself, but after the 8th period it maintains at 50%.

For Shanghai, its impact on Hefei and Nanjing’s price is significant while has no contribution to Hangzhou. Hefei’s impact on Hangzhou is also very low. Shanghai’s housing prices in the 2nd
period reached the maximum, but after the decline in the 5th period maintains at 20%. The contribution of Nanjing to Hefei’s housing prices in the 5th period maintains a high level of 45%.

Conclusion and Suggestion
Shanghai’s housing price has a clear self-accumulation effect and spillover effect at the same time, thus driving the price rising in Nanjing and Hefei, so we should pay more attention to housing prices trend in Shanghai. Hangzhou’s housing price is largely affected by itself. Meanwhile, Nanjing and Hefei have a distinct spillover effect on each other, if the government wants to regulate one of real estate markets, analyzing the other one’s price in advance is necessary, thus making appropriate adjustments.

References