An Empirical Study of Cointegration with Structural Breaks of Shanghai and Shenzhen Stock Markets

YANG Jiping  SUN Lisha

(School of Economics and Management, Beihang University, Beijing 100083, China)

Abstract: Using cointegration techniques allowing for structural breaks, the paper tests the long-run equilibrium relationships between Shanghai and Shenzhen stock markets. The conclusion has been reached that the cointegration relationship between them does have a structural break during the sample period. Through establishing error correction model and forecasting the future Shanghai synthesis index, it has been demonstrated that cointegration model with regime shifts can well characterize the equilibrium relationships between the two stock markets when there exists a structural break.

Keywords: cointegration, structural break, regime shift, Shanghai and Shenzhen stock markets

1. Introduction

Many studies have shown that there is a cointegration relationship between Shanghai and Shenzhen stock markets. For example, Shi Daimin (2002) explored the cointegration relationships among Shanghai synthesis index, Shanghai 30 Index, Shenzhen composition index and Shenzhen synthesis index and found that there was a cointegration relation between Shanghai synthesis index and Shenzhen synthesis index during the sample period from January 1, 1993 to July 31, 2001[1]. Chen Shoudong, Han Guangzhe and Jing Wei (2003) also proved that Shanghai synthesis index and Shenzhen synthesis index were cointegrated during the sample period from November 20, 1992 to July 12, 2002[2]. Yin Ling (2005) explored the Shanghai index, Shenzhen composition index and other different indices and draw a conclusion that there was a cointegration relationship between Shanghai index and Shenzhen composition index[3].

However most of the empirical studies fail to notice that the cointegration relationship may have a structural break during the sample period. The structural breaks of cointegration relationship mean the significant change of the cointegration parameters or the change of the existence of cointegration relationships. Since the development of stock markets can be affected by policies, system, economies and other factors of a country, the cointegration relationships may have a structural break at some special time when a large fluctuation appears.

Joen and Von Furstenberg (1990) investigated daily stock price indices in Tokyo, Frankfurt, London and New York for the period January 1986 through November 1988 and found a significant structural change with regard to the correlation structure and leadership since the stock market crash of October 1987[4].

Gregory, Nason, and Watt (1996) gave a proof that the power of the conventional ADF test with no allowance for regime shifts falls sharply[5], which means the conventional ADF test tends
to accept the null hypothesis of no cointegration when there was a structural break.

Perron (1989), Banerjee, Lumsdaine, and Stick (1992), Perron and Vogelsang (1992), Zivot and Andrews (1992) tested the null of a unit root in univariate time series against the alternative of stationarity, while allowing for a structural break in the deterministic component of the series [6-9]. Based on this, Gregory and Hansen (1996) developed a residual-based cointegration approach that allows for regime shifts [10].

There are some researches in China about cointegration with structural breaks. Yang Baochen and Zhang Shiying (2002) defined cointegration according to three different types. Through analyzing the relationship between the national income and the total consumption of the society in China they found that the distribution patterns and proportion of the national income changed largely and this appeared as the level shifts in structure after China’s reform and opening [11, 12]. Sun Qinghua and Zhang Shiying (2003), Zhang Shiyung (2004) proposed model-free technology to test the cointegration relationships with structural breaks. They used the algorithm of neural networks and hierarchical genetic programming to analyze and detect cointegration with structural breaks in multivariate nonlinear time series. They also used this method to test different indices of Shanghai stock market and found that the relationships among five different indices changed largely for several times in 1998 [13,14].

It is essential to test the cointegration relationships with structural breaks. Firstly, we can know more about the relationships among variables. Secondly, we can establish and estimate the cointegration equation and error correction model more exactly.

2. Basic theory of cointegration model with regime shifts

In the conventional cointegration test, the cointegration model is: \( Y_t = a + bX_t + \epsilon_t \), in which \( X_t, Y_t \) are integration time series with order of \( d \) and \( \{ \epsilon_t \} \) is residual series, and the test is the residual-based one in which the null hypothesis is no cointegration against the alternative that the relation is cointegrated. With this method we can deduce that there is no cointegration between variables if the test fails to reject the null hypothesis for a sample period. In fact, this may be falsely concluded because of the existence of structural breaks.

Yang Baochen and Zhang Shiyung (2002) defined three types of cointegration with structural breaks. They are cointegration with parameter changes, partly cointegration and cointegration with mechanism changes. Simply speaking, cointegration with parameter changes means the parameters of the cointegration equation happen to change at some time, but the cointegration relationship still exists. Partly cointegration means the cointegration relationship exists before or after some time but disappears in other periods. Cointegration with mechanism changes means the former cointegration relationship is destroyed because new variables enter the system and they form a new type of cointegration relationship [11].

Define \( y_{1t} \) and \( y_{2t} \) as \( m \)-vectors, for the cointegration with parameter changes, Gregory and Hansen (1996) developed three models as follows [10]:

\[
\begin{align*}
  y_{1t} &= \mu_1 + \mu_2 \varphi_{1t} + \alpha^T y_{2t} + \epsilon_t, t = 1, \cdots, n \\
  y_{1t} &= \mu_1 + \mu_2 \varphi_{1t} + \beta t + \alpha^T y_{2t} + \epsilon_t, t = 1, \cdots, n \\
  y_{1t} &= \mu_1 + \mu_2 \varphi_{1t} + \beta t + \alpha_1^T y_{2t} + \alpha_2^T y_{2t} \varphi_{1t} + \epsilon_t, t = 1, \cdots, n
\end{align*}
\]
where $\alpha$ is a cointegration vector, $\{e_t\}$ is residual series, $n$ is the sample size, and $\varphi_{\tau t}$ is the dummy variable:

$$
\varphi_{\tau t} = \begin{cases} 
0 & \text{if } t \leq [n \tau] \\
1 & \text{if } t > [n \tau] 
\end{cases}, \quad \tau \in (0,1)
$$

The symbol $[*]$ denotes integer part. $T = [nt]$ is the breakpoint when cointegration may happen to change, and cointegration relationship is one type before time $T$ and another after. It can be realized by giving different values to the dummy variable.

Model (1) represents that there is a level shift in the cointegration relationship, i.e. there is a change in the intercept $\mu$, while $\mu_1$ represents the intercept before the shift, and $\mu_2$ represents the change in the intercept at the time of the shift. Model (2) is added a time trend $\beta t$ on the basis of Model (1). Model (3) allows the slope vector to shift as well in which $\alpha_1$ denotes the cointegration slope coefficients before the regime shift, and $\alpha_2$ denotes the change in the slope coefficients. Model (3) includes all of the three changes in intercept, time trends and slope coefficients.

We can divide the method of testing cointegration with structural breaks into two steps: first, for each possible breakpoint $T = [nt]$, estimate the models (1)-(3) by OLS, yielding the residual series $\{e_t\}$ from which we can get the values of ADF test statistic. The statistic of the cointegration test with allowance of regime shifts is the smallest value of the conventional ADF test statistic across all values of every possible breakpoint. The new test statistic we use is $ADF^*$, $ADF^* = \inf ADF(\tau)$; Second, compare the value of $ADF^*$ test statistic and the critical value given by Gregory and Hansen (1996) using Monte Carlo simulation method. If the value of $ADF^*$ test statistic is smaller than the critical value, we can conclude that there exists a cointegration relationship. But we should notice that in the general cointegration test proposed by Gregory and Hansen (1996) the null hypothesis (no cointegration) is the same with the conventional cointegration method, while the alternative is different, it contains the Engle-Granger model as a special subcase. So if the test rejects the null hypothesis under the new method we can conclude that there is cointegration but we don’t know whether it is cointegration with structural breaks or not. To detect the breakpoints we need do a further test. If we reject the null hypothesis under the new method and fail to reject the null hypothesis under the conventional method, we can say that the cointegration relationship with structural breaks exists.

3. Empirical study on Shanghai and Shenzhen stock markets

(1) Data selecting

Sample data used in this study comprise daily stock prices of Shanghai synthesis index and Shenzhen composition index at market close (sample period June 1, 1995 through June 28, 1996) and we obtained the data on the sohu website (http://business.sohu.com). Take every day as a possible breakpoint, so there are 260 possible breakpoints in all (excluding the first two and the last points).

The sample period we select is shorter and we don’t use the latest data for two main reasons. Firstly, the cointegration test with allowance of structural breaks is applicable for existing only one breakpoint, long sample period and more than one breakpoint will affect the test; Secondly,
according to the division of the process of stock market’s development by Ma Xiangqian and Ren Ruoen (2002) we select a period with higher possibility of having breakpoints to test whether the cointegration relationship changed with the development of stock markets.

(2) Conventional cointegration test for the whole sample period

First of all, we make curves of the closing price of Shanghai synthesis index and Shenzhen composition index using Evies3.1. \((p_{shang} \text{ and } p_{shen})\) denote daily closing price of Shanghai synthesis index and Shenzhen composition index, respectively.)

Fig.1 shows that Shanghai synthesis index and Shenzhen composition index have similar trends. Then we do the unit root test using Eviews3.1 and selecting the test equation with intercept and time trends and the lagged difference equal to 1. The result is shown in Table 1:

<table>
<thead>
<tr>
<th></th>
<th>ADF test statistic</th>
<th>Critical value</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1%</td>
<td>5%</td>
</tr>
<tr>
<td>Shanghai synthesis index</td>
<td>Level</td>
<td>-0.3323</td>
<td>-3.9965</td>
</tr>
<tr>
<td></td>
<td>1st difference</td>
<td>-12.6681</td>
<td>-3.9966</td>
</tr>
<tr>
<td>Shenzhen composition index</td>
<td>Level</td>
<td>0.5429</td>
<td>-3.9965</td>
</tr>
<tr>
<td></td>
<td>1st difference</td>
<td>-10.6935</td>
<td>-3.9966</td>
</tr>
</tbody>
</table>

From Table 1, we can conclude that the time series of Shanghai synthesis index and Shenzhen composition index are both I(1).

Then we use conventional method, Engle-Granger two-stage estimating method, to test cointegration between Shanghai synthesis index and Shenzhen composition index.

Step 1: Using OLS to estimate cointegration equation: using Evies3.1 (the software we used
below is the same)

\[ pshang = 467.8 + 0.157 \ pshen + e_t \]  \hspace{1cm} (4)

where the values in the brace are the p-values of the t-test statistic.

Step 2: Testing the unit root of residual series and the result is shown in Table 2:

<table>
<thead>
<tr>
<th></th>
<th>ADF test Statistic</th>
<th>Critical value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>Residual series</td>
<td>-1.6201</td>
<td>-2.5735</td>
</tr>
</tbody>
</table>

Table 2 shows that, under the significant level of 10%, there is cointegration between Shanghai and Shenzhen stock markets; while under the significant level of 1%, we fail to arrive at the same conclusion.

(3) Cointegration test with allowance of structural breaks

The test equations are models (1)-(3) which have been mentioned above.

For every possible breakpoint, calculate the values of ADF test statistic of each model. According to the properties of residual series we choose no time trends and intercept for the test equation, and according to the AIC and BIC principle, we choose the lagged difference equal to 1. The values of ADF test statistic of each model are expressed in Fig.2: (Fig.2 and Fig.3 below are drawn using Excel)

![Fig.2 Values of ADF test statistic of each model](image)

Through calculation the values of ADF test statistic of each model, we can get the value of ADF* test statistic, \( ADF^* = \inf ADF(r) \). For model (1) and (2) the time point according to ADF* test statistics is at December 7, 1995 and July 20, 1995, and for model (3) the time point is at April 26, 1996. They are compared in Table 3:
Table 3 Results of the test for cointegration with allowance of structural breaks

<table>
<thead>
<tr>
<th>Cointegration models with regime shifts</th>
<th>ADF* test statistic</th>
<th>Time point</th>
<th>Critical value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>1%</td>
</tr>
<tr>
<td>(1)</td>
<td>-3.95831</td>
<td>———</td>
<td>-5.13</td>
</tr>
<tr>
<td>(2)</td>
<td>-4.03023</td>
<td>———</td>
<td>-5.45</td>
</tr>
<tr>
<td>(3)</td>
<td>-5.9549</td>
<td>96.4.26</td>
<td>-5.47</td>
</tr>
</tbody>
</table>

Note: the critical values are from reference [10]

From Table 3 we can see that under the significant level of 1% model (3) rejects the null hypothesis. Having concluded that there is no cointegration under the significant level of 1% using the conventional test, we can conclude that there is cointegration with a structural break between Shanghai and Shenzhen stock markets during the sample period June 1, 1995 to June 28, 1996. The cointegration relationship happened to change at the end of April 1996 and it changed in intercept, time trends and coefficient.

4. Comparing the forecasting results of two cointegration models

In this part, we use two cointegration equations to establish error correction models. Through forecasting and comparing the precision we can find which model can characterize of the data better.

We use the closing price of Shanghai synthesis index and Shenzhen composition index over the sample period June 1, 1995 to June 28, 1996 to forecast the next seven trading days’ closing price of Shanghai synthesis index.

First, establishing error correction model:

We can get the residual series \{\varepsilon_t\} from the cointegration equation (4) for the whole sample period, using OLS to estimate the parameters of the error correction model below, we can get:

\[
\Delta p_{\text{shang}} = -0.828 + 0.328 \Delta p_{\text{shen}} - 0.008 ECM_{t=1} + \varepsilon_t \tag{5}
\]

In equation (5), \(\Delta p_{\text{shang}}\) and \(\Delta p_{\text{shen}}\) denote the change of closing price at time \(t\), \(ECM_{t=1}\) is the error correction term at time \(t-1\), \(ECM_{t=1} = p_{\text{shang}_{t=1}} - 0.157p_{\text{shen}_{t=1}} - 467.8\). The values in brace are \(p\)-values of \(t\)-test statistic of the parameters. \(R^2=0.4510\), the \(p\)-value of the \(F\)-test statistic is 0.0000, DW=2.0443, so the goodness of fit for the equation is satisfying.

We use the same method to establish the error correction model with structural breaks. The value of the dummy variable \(\varphi_t\) is definite because the position of the breakpoint is known.

First we can estimate the cointegration equation:

\[
p_{\text{shang}} = 11.9 + 0.6356 p_{\text{shen}} + 378.0085 \varphi - 0.5173 t - 0.3850 \varphi, shen + \varepsilon_t \tag{6}
\]

Then we can get the residual series to estimate the error correction model:

\[
\Delta p_{\text{shang}} = -0.8053 + 0.3158 \Delta p_{\text{shen}} - 0.0991 ECM_{t=1} + \varepsilon_t \tag{7}
\]

The \(R^2\) of equation (7) is 0.4704, the \(p\)-value of the \(F\)-test statistic is 0.0000, and DW=2.1103.

Now we use two error correction models: equation (5) and (7) to forecast the future closing price of Shanghai synthesis index. We repeat this in the same way. First, use the sample period of
June 1, 1995 through June 28, 1996 to estimate the error correction model, and then forecast the price of July 1, 1996. Then put the price forecasted to the sample period and repeat the steps above to forecast the price of July 2, 1996, … , repeat this process for seven times, so we can get the forecasted price of future seven trading days. The results of the two models are shown in Table 4:

Table 4 Comparisons of the forecasting results using two models

<table>
<thead>
<tr>
<th>Date</th>
<th>The forecasted price</th>
<th>The real price</th>
<th>Forecasting precision</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Conventional model</td>
<td>Model with regime shifts</td>
<td></td>
</tr>
<tr>
<td>1996.07.01</td>
<td>765.16</td>
<td>763.31</td>
<td>761.12</td>
</tr>
<tr>
<td>1996.07.02</td>
<td>774.15</td>
<td>770.40</td>
<td>759.71</td>
</tr>
<tr>
<td>1996.07.03</td>
<td>796.41</td>
<td>791.24</td>
<td>786.13</td>
</tr>
<tr>
<td>1996.07.04</td>
<td>806.97</td>
<td>800.03</td>
<td>785.43</td>
</tr>
<tr>
<td>1996.07.05</td>
<td>825.64</td>
<td>817.29</td>
<td>785.38</td>
</tr>
<tr>
<td>1996.07.08</td>
<td>841.37</td>
<td>831.60</td>
<td>807.34</td>
</tr>
<tr>
<td>1996.07.09</td>
<td>853.02</td>
<td>841.51</td>
<td>817.61</td>
</tr>
</tbody>
</table>

To see their relationships more clearly we put the forecasted prices and the real price in Fig. 3.

Fig. 3 Curves of forecasted price and real price
5. Conclusion

Through establishing error correction models using two cointegration models and forecasting the price of Shanghai synthesis index, we can clearly see that the error correction model established by cointegration model with regime shifts has a higher precision compared with the one established by conventional cointegration model. It provides a validation of the conclusion we get in the third part that the cointegration relationship between Shanghai and Shenzhen stock markets happened to change during the sample period, and the cointegration model with regime shifts can characterize the equilibrium relationship more exactly when there is a structural break.

We can also argue that the cointegration relationship with structural breaks can describe the relationship between the two stock markets better from other two aspects:

1. From the significance level of cointegration tests, the cointegration exists under the level of 10% by using the conventional testing method, while 1% by using the model with regime shifts.

2. From the $p$-values of the $t$-test statistic of the error correction term, it is 0.4624 in equation (5); and it is 0.0017 in equation (7), which reflects the error correction term is significant. In other words, in the error correction model established using cointegration model with regime shifts, the long-run factor has significant explanatory ability for the change of dependent variable.

In this paper, we studied the sample period of June 1995 to June 1996, and the breakpoint is at the end the April 1996. We can find some evidence to explain the cause of the breakpoint. Ma Xiangqian and Ren Ruoen (2002) divided the developing progress of China’s stock markets into three periods: 1992-1993, the initial stage; 1993-1995, the interim; 1996-2000, developing stage [15]. It means that in 1996 the stock markets of China changed from one stage to another. So the equilibrium relationship between them is likely to change at the same time. The breakpoint we find is just in April 1996, and it may be caused by the change of the development style.

References

8. Perron, P. and Vogelsang, Timothy J.. Nonstationarity and level shifts with an application to


