行政院國家科學委員會補助專題研究計畫 ■ **成果報告** □**期中進度報告**

偵測及認定追蹤資料中之介入模型分析-Spectral Whittle 估計法之應用

計畫類別：■ 個別型計畫　　□ 整合型計畫
計畫編號：NSC　97－2410－H－029－004－
執行期間：　97 年 08 月 01 日 至 98 年 07 月 31 日

計畫主持人：陳文典
共同主持人：
計畫參與人員：　蔣齡玲

成果報告類型(依經費核定清單規定繳交)：■精簡報告　□完整報告
執行單位：東海大學

中　華　民　國　98　年　07　月 31　日

偵測及認定追縱資料中之介入模型分析-Spectral Whittle 估計法之應用

中文摘要

由於一般追蹤資料模型皆使用大量的資料，非常容易有離群值的發生，這些離群值會影響到模型的估計及效率,因此在本研究中希望能建構出一個快速且有效率的方法，來檢測追蹤資料模型中的離群值。

本研究主要針對追蹤資料模型中離群值的偵測及認定，吾人使用了 spectral Whittle 估計法，根據 Chen(2006) 所發表的文章，將 Chang, Tiao, Chen(1998) 的介入模型分析方法應用在追蹤資料的分析上。在本研究中使用了固定效果模型，而且其誤差的型態可以延申到 ARFIMA(p,q)數列的型態。其中建立了概度比的檢定方法，可以快速的降低估計的時間及成本。

在本研究中，吾人將模型的應用推展到 AO (additive outlier),IO (Innovation outlier), 及 Level Shifts 的檢測,由 Monte Carlo 模擬的方式來驗証，當個別組數及時間長度增加時，估計式的檢定力。由結果分析可以發現模型具相當好的檢定能力。而在在實証的分析上，以台灣地區 12 家私人銀行為例，藉以了解模型的適用情形。

關鍵字: 介入模型、可加式離群值、創新性離群值、Level Shifts、Whittle 估計法、譜系分析、追蹤資料模型。

Detecting and identifying level shifts, additive outliers, and innovation outliers in a time series model

Abstract

This article provides a fast approximation to detect and identify the interventions in a time series model, in which level shifts and two fundamental types of outliers (additive (AO) and innovation (IO) outliers) are demonstrated. This research extends Chang, Tiao, and Chen (1988)'s procedure to a panel data model, and a modified inverse Fourier transform is used to construct the statistics, which is based on the spectral Whittle approach.

Through Monte Carlo experiments the consistency of the estimator is examined for the panel data model in which the dataset is contaminated with three types of interventions: AO, IO, and level shift. From the power tests we can see that the estimators are quite successful and powerful.

In the case study we focus on the performance of local private Taiwanese banks in the stock market from January 2000 to December 2006. By the panel data model we can easily estimate the impacts of the level shifts and outliers. In the result analysis Taitung Business Bank has the largest structure change with twice downward level shifts, i.e., it has the worse performance as time increases. The impacts are respectively -9.1882 and -10.0069, and the *t*-values are -4.655 and -5.2156. It makes sense that this debt-ridden bank was taken over by Central Deposit Insurance Corp. of Taiwan's government in December 2006.

Keyword: additive outlier, innovation outlier, level shift, long memory, panel data model.

1. **Introduction**

In a time series model the outliers can be regarded as being generated by dynamic intervention models. The structure of the underlying process plays an important role when detecting and identifying these outliers. A fundamental approach for the time series model was proposed by Fox (1972), and it demonstrate two typical types of outliers, additive and innovation outliers, with an AR(p) process. In economic system different types of interventions represent different types of dynamic influences of events. It is therefore important to detect and identify these interventions. In the estimation we remove these effects from the observations to better understand the underlying structure of the series. Therefore, many studies in the literature discuss the intervention detection and identification, such as Box and Tiao (1975), Chang, Tiao, and Chen (1988), Bianco, Ben, Martinez, and Yohai (2001), etc.

Because the interventions decrease the estimator's efficiency, many robust estimators have been developed, such as in Martin and Yohai (1986), Yohai (1987), and Pena and Yohai (1999). In practice, we may find a small fraction of atypical observations in the case when innovation or additive outliers may not significantly affect the robust estimates, but it is different in the case of level shifts. This causes a serious bias even after applying an appropriate robust estimator. Hence, Bianco, Ben, Martinez and Yohai (2001) proposed a diagnostic procedure for detecting level shifts as well as additive and innovation outliers in a regression model.

This article focuses on the level shifts and develops a fast approximation that can be used to detect and identify these interventions at unknown time points, in which we follow Chang, Tiao, and Chen (1988)'s useful iteration procedure and then extend to a panel data model. We compute the statistics herein based on the familiar spectral Whittle method. In conventional approaches we know that the spectral Whittle approach can save much time in recursive calculation and is easily operated in coordination with a periodogram, especially when the time length is long and the model is complicated (e.g., a mixed model with fractional AR and MA). In a time series model we have to test the outliers by the residuals that are demonstrated in the time domain. It does not seem easy to have a corresponding appropriate approach to being with a spectral model. This paper provides an approach to solve this problem.

2. Identification of level shift, additive outlier, and innovation outlier

This section develops statistics for detecting and identifying the outliers and level shifts, in which we extend Chang, Tiao, and Chen (1988)'s method to level shift detection and a quick approach in the spectral domain is developed. By a fast calculation we can apply the model on a large dataset.

Assume that an underlying process z_t follows an ARIMA(p,d,q) model, i.e.:

$$\phi(B)(1-B)^d z_t = \theta(B)a_t, \tag{1}$$

where $\phi(B) = 1-\phi_1 B - \cdots \phi_p B$, $\theta(B) = 1-\theta_1 B - \cdots - \theta_q B^q$, B is the backshift operator, $\{a_t\}$ is white noise, and $\mathrm{var}(a_t) = \sigma_a^2$. $\phi(B)$ and $\theta(B)$ represent the corresponding AR and MA polynomial operator functions with roots outside the closed unit circle, and a stationary process with $-0.5 \le d < 0.5$ wherein $d \in (0, 0.5)$ indicates a long memory process.

In the beginning let us regard an additive outlier. From Chang, Tiao, and Chen (1988) whereby a process contains an additive outlier at the sth observation, we can set $y_t = z_t + \alpha_{AO} P_t^{(s)}$ where $P_t^{(s)} = 1$ for $t = s$, and $P_t^{(s)} = 0$ otherwise. This can be expressed in a vector form:

$$P'_{AO,s} = \begin{bmatrix} 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \end{bmatrix}, \tag{2}$$

and the observations are:

$y = \alpha_{AO} P_{AO,S} + z$,

where $y' = [y_1 \quad y_2 \quad \cdots \quad y_T]$, $z' = [z_1 \quad z_2 \quad \cdots \quad z_T]$.

Following Chang, Tiao, and Chen (1988, eqn 2.2a) and (1), we can estimate the impact $\alpha_{AO,s}$ through the least square regression, i.e.:

$$\hat{\alpha}_{AO,s} = \sum_{j=0}^{T-s} \pi_j e_{j+s} \Big/ \sum_{j=0}^{T-s} \pi_j^2 , \qquad (3)$$

where $e_t = \pi(B) y_t$ and $\pi(B) = \phi(B)(1-B)^d \theta^{-1}(B) = 1 + \pi_1 B + \pi_2 B^2 - \cdots$. Here, $\pi(B)$ is a polynomial filter function that can transform $\{z_t\}$ to be a white noise process $\{a_t\}$. In finite sample observations we naturally apply the covariance matrix of z and $z' = [z_1 \quad z_2 \quad \cdots \quad z_T]$ to obtain the filter coefficients.

Let the covariance matrix of z be $\sigma_a^2 \Sigma(\Theta)$ and $\sigma_a^2 = \text{var}(a_t)$. The Cholesky decomposition of the inverse covariance matrix can be written as $\Sigma^{-1}(\Theta) = UU'$ where U is an upper triangular and it can be used to transform $\{z_t\}$ to be white noise, e.g., $U'z = a$. In the same way we can follow Chang, Tiao, and Chen (1988) to transform y_t to obtain the 'residual', i.e.:
$e = U'y$ where $e = [e_1 \quad e_2 \quad \cdots \quad e_T]$.

The corresponding filter coefficients can be obtained by $\tau' = P'_{AO,s} U$, which is approximately $[0 \quad \cdots \quad 0 \quad \pi_0 \quad \pi_1 \quad \cdots \quad \pi_{T-s}]$. Thus, relative to (3) the estimator of $\alpha_{AO,s}$ can be obtained by:

$$\hat{\alpha}_{AO,s} = \tau' e / \tau' \tau = \frac{P_{AO,s}' \Sigma^{-1}(\Theta) y}{P_{AO,s}' \Sigma^{-1}(\Theta) P_{AO,s}} . \qquad (4)$$

According to Beran (1994, page 116), we know:

$$\Sigma^{-1}(\Theta) = [\delta(m-l)]_{m,l=1,\ldots,T} \text{ and } \delta[m-l] \approx \frac{1}{T} \sum_{j \in J} \frac{1}{g(\lambda_j;\Theta)} e^{i(m-l)\lambda_j} , \qquad (5)$$

where $\Theta = (\phi_1, \ldots, \phi_p, d, \theta_1, \ldots, \theta_q)'$, $g(\lambda_j;\Theta) = \left| \Theta(e^{-i\lambda_j}) \Phi^{-1}(e^{-i\lambda_j})(1 - e^{-i\lambda_j})^{-d} \right|^2$,

$J = \{-M, \cdots, M\} - \{0\}$, and $M = [T/2]$, [.] is the floor integer. Here, we take away the zero frequency by mean correction to avoid the spectral density function exploding when z_t is a long memory process $d \in (0, 0.5)$.

Putting the equation (5) into (4) the estimator we obtain is:

$$\hat{\alpha}_{AO,s} \approx \left(2\sqrt{\frac{2\pi}{T}} \sum_{j=1}^{M} \mathrm{Re}\left[e^{is\lambda_j} \frac{w_y(\lambda_j;\Theta)}{g(\lambda_j;\Theta)} \right] \right) \Big/ \left(\frac{1}{M} \sum_{j=1}^{M} \frac{1}{g(\lambda_j;\Theta)} \right). \quad (6)$$

According to Chang, Tiao, and Chen (1988) the variance of $\tilde{\alpha}_{AO,s}$ is:

$$\mathrm{var}\left(\tilde{\alpha}_{AO}\right) = \sigma_a^2 \Big/ \sum_{j=0}^{T-s} \pi_j^2 .$$

This can be approximately by:

$$\mathrm{var}\left(\tilde{\alpha}_{AO}\right) \approx \sigma_a^2 \left(\frac{2}{T} \sum_{j=1}^{M} \frac{1}{g(\lambda_j;\Theta)} \right)^{-1} .$$

Thus, the statistics can be established by:

H_0 vs. H_1:

$$\tilde{\eta}_{AO,s} = \frac{\tau\tilde{\alpha}_{AO,s}}{\tilde{\sigma}_a} = \left(2\sqrt{\frac{2\pi}{T}} \sum_{j=1}^{M} \mathrm{Re}\left[e^{is\lambda_j} \frac{w_y(\lambda_j)}{g(\lambda_j;\Theta)} \right] \right) \Big/ \left(\tilde{\sigma}_a \sqrt{\frac{2}{T} \sum_{j=1}^{M} \frac{1}{g(\lambda_j;\Theta)}} \right). \quad (7)$$

From (6) and (7) we see the coefficients and statistics can be calculated by the modified inverse Fourier transform, which can save much time especially when applied to a huge dataset.

Aside from the additive outlier, another important type of outlier has been popularly discussed in the literature. This is the innovation outlier (IO), in which the impact not only affects the particular observation, but also the subsequent observations, which can be seen in Fox (1972) and Chang, Tiao, and Chen (1988). In the following we focus on IO and develop its corresponding estimator and statistics. If the impact affects the *s*th observation and the subsequent observations, then we rewrite (2) in the following vector form:

$$P'_{IO,s} = [0 \quad \cdots \quad 0 \quad \tau_0 \quad \tau_1 \quad \cdots \quad \tau_{T-s}],$$

where τ_0 start at the sth element. If it is an IO, then the coefficients have the following relationship:

$$\tau(B) = \tau_0 + \tau_1 B + \tau_2 B^2 + \cdots = \theta(B)\phi^{-1}(B)(1-B)^{-d}.$$

According to (4), the impact can be written as:

$$\tilde{\alpha}_{IO,s} = \frac{P_{IO,s}{}'\Sigma^{-1}(\Theta)y}{P_{IO,s}{}'\Sigma^{-1}(\Theta)P_{IO,s}}.$$

Putting (5) into the above equation we obtain an estimator:

$$\tilde{\alpha}_{IO,s} = 2\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\mathrm{Re}\left[\frac{w_y(\lambda_j)}{g(\lambda_j;\Theta)}\frac{\theta(e^{i\lambda_j})}{\phi(e^{i\lambda_j})(1-e^{i\lambda_j})^d}e^{is\lambda_j}\right].$$

Following Chang, Tiao, and Chen (1988) we obtain the corresponding variance:

$$\mathrm{var}\left(\tilde{\alpha}_{s,IO}\right) = \sigma_a^2.$$

Therefore, we establish the statistics:

$$H_0 \text{ vs. } H_2\text{:} \quad \tilde{\eta}_{IO,s} = \frac{2}{\tilde{\sigma}_a}\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\mathrm{Re}\left[\frac{w_y(\lambda_j)}{g(\lambda_j;\Theta)}\frac{\theta(e^{i\lambda_j})}{\phi(e^{i\lambda_j})(1-e^{i\lambda_j})^d}e^{is\lambda_j}\right]. \quad (8)$$

From the above estimator and statistics we see the modified inverse Fourier transform still can be used to obtain the coefficients quickly which is slightly different from (6) and (7).

Aside from AO and IO, the level shift is also a typical event that one often encounters in time series models. Relative to (2), we set the impact vector as a step function as the following:

$$P_{L,s} = [0 \quad \cdots \quad 0 \quad 1 \quad 1 \quad \cdots \quad 1],$$

where the unit elements begin at the sth element. Thus, following (4) we obtain the estimator as:

$$\hat{\alpha}_{L,s} = \frac{P_{L,s}{}'\Sigma^{-1}(\Theta)y}{P_{L,s}{}'\Sigma^{-1}(\Theta)P_{L,s}}.$$

In the same way, we put (5) into the equation and obtain the estimator as:

$$\tilde{\alpha}_{L,s} = 2\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\mathrm{Re}\left[\frac{w_y(\lambda_j)}{g(\lambda_j;\Theta)}\left(\frac{1-e^{i(T-s+1)\lambda_j}}{1-e^{i\lambda_j}}\right)e^{is\lambda_j}\right]\left(\frac{2}{T}\sum_{j=1}^{M}\frac{1}{g(\lambda_j;\Theta)}\left(\frac{\sin(T-s+1)\lambda_j/2}{\sin\lambda_j/2}\right)^2\right)^{-1}.$$

This can be represented by an alternative form:

$$\tilde{\alpha}_{L,s} = 2\sqrt{\frac{2\pi}{T}}\left(\sum_{j=1}^{M}\mathrm{Re}\left[\frac{w_y(\lambda_j)}{g(\lambda_j;\Theta)}\frac{e^{is\lambda_j}}{1-e^{i\lambda_j}}\right]-\sum_{j=1}^{M}\frac{w_y(\lambda_j)}{g(\lambda_j;\Theta)}\frac{e^{i\lambda_j}}{1-e^{i\lambda_j}}\right)\times$$
$$\left(\frac{2}{T}\sum_{j=1}^{M}\frac{1}{g(\lambda_j;\Theta)}\left(\frac{\sin(T-s+1)\lambda_j/2}{\sin\lambda_j/2}\right)^2\right)^{-1}.$$

From the above equation we see the first term can be obtained by a modified inverse Fourier transform. In practice, it saves much time in calculation. The relative variance of the estimator can also be obtained by a least square method, that is:

$$\mathrm{var}\left(\tilde{\alpha}_{L,s}\right) = \left(P'_{L,s}\Sigma^{-1}(\Theta)P_{L,s}\right)^{-1}\sigma_a^2.$$

This is approximately by using (5) as:

$$\mathrm{var}\left(\tilde{\alpha}_{L,s}\right) = \sigma_a^2\left(\frac{2}{T}\sum_{j=1}^{M}\frac{1}{g(\lambda_j;\Theta)}\left(\frac{\sin(T-s+1)\lambda_j/2}{\sin\lambda_j/2}\right)^2\right)^{-1}.$$

The corresponding statistics are established by:

H_0 vs. H_3:

$$\tilde{\eta}_{L,s} = \frac{\tilde{\alpha}_{L,s}}{\sqrt{\mathrm{var}\,\tilde{\alpha}_{L,s}}}$$
$$\approx \frac{\left(2\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\mathrm{Re}\left[\frac{w_y(\lambda_j)}{g(\lambda_j;\Theta)}\frac{e^{is\lambda_j}}{1-e^{i\lambda_j}}\right]-2\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\frac{w_y(\lambda_j)}{g(\lambda_j;\Theta)}\frac{e^{i\lambda_j}}{1-e^{i\lambda_j}}\right)}{\tilde{\sigma}_a\sqrt{\left(\frac{2}{T}\sum_{j=1}^{M}\frac{1}{g(\lambda_j;\Theta)}\left(\frac{\sin(T-s+1)\lambda_j/2}{\sin\lambda_j/2}\right)^2\right)}}.$$

From the above equation we see that the first term in the numerator part is obtained by the modified inverse Fourier transform, and the second term is a particular case when $s = 1$. Thus, we calculate the statistics $\tilde{\eta}_{L,s}$ at each time $s = 1,\ldots,T$ very fast.

3. **Iterative procedure for outliers' detection and identify**

We now follow Chang, Tiao, and Chen (1988)'s method to implement the iterative procedure for the outliers' detection and identification. At the first stage of this procedure, the fractional ARIMA model is estimated by the observed time series y_t in the Whittle approach and assuming that the series contains no outliers.

Relative to the previous section we calculate the statistics $\tilde{\eta}_{AO,t}$, $\tilde{\eta}_{IO,t}$, and $\tilde{\eta}_{L,t}$ which can be calculated for each time t = 1, 2,…, T. Following Chang, Tiao, and Chen (1988) we calculate the largest statistics by $\tilde{\eta}_s = \max_t[\max(|\tilde{\eta}_{IO,t}|, |\tilde{\eta}_{AO,t}|, |\tilde{\eta}_{L,t}|)]$. If $\tilde{\eta}_s > c$ where c is a pre-specified critical value, then we remove the intervention's effect. Typically, the value c could be 3.0, 3.5, or 4.0. If $\tilde{\eta}_s = |\tilde{\eta}_{AO,s}| > c$ which indicates an additive outlier, then we remove the effect from the original data, i.e., $y_s^* = y_s - \tilde{\alpha}_{AO,s}$; else if $\tilde{\eta}_s = |\tilde{\eta}_{IO,s}| > c$, then $y_{s+r}^* = y_{s+r} - \tilde{\alpha}_{IO,s}\tau_r$ for r = 0, …, T-s. Otherwise, if $\tilde{\eta}_s = |\tilde{\eta}_{L,s}| > c$, then $y_{s+r}^* = y_{s+r} - \tilde{\alpha}_{L,s}$ for r = 0, …, T-s. The preceding steps are then repeated until all outliers are identified.

After removing these outliers, we calculate the model with a traditional Whittle approach:

$$Q_{(0)} = \frac{4\pi}{T} \sum_{j=1}^{[T/2]} \frac{I_{y^*}^{(T)}(\lambda_j)}{g(\lambda_j;\Theta)}, \text{ where } -0.5 < d < 0.5,$$

where $y^* = \begin{bmatrix} y_1^* & y_2^* & \cdots & y_T^* \end{bmatrix}'$ and $Q_{(0)} \to \sigma_a^2$ if $T \to \infty$. Here, σ_a^2 is the variance of the noise a_t. Through the first-order condition, $\partial Q_{(0)} / \partial\Theta \big|_{\Theta=\tilde{\Theta}_{(0)}} = 0$, a Taylor expansion can be obtained by:

$$\left.\frac{\partial Q_{(0)}}{\partial \Theta}\right|_{\Theta=\tilde{\Theta}_{(0)}} = \left.\frac{\partial Q_{(0)}}{\partial \Theta}\right|_{\Theta=\Theta_0} + \left.\frac{\partial^2 Q_{(0)}}{\partial \Theta \partial \Theta'}\right|_{\Theta_1} \left(\tilde{\Theta}_{(0)} - \Theta_0\right),$$

where $\left|\Theta_1 - \Theta_0\right| \le \left|\tilde{\Theta}_{(0)} - \Theta_0\right|$ and the covariance matrix can be obtained:

$$COV\left(\tilde{\Theta}_{(0)}\right) = \left(\left.\frac{\partial^2 Q_{(0)}}{\partial \Theta \partial \Theta'}\right|_{\Theta=\Theta_1}\right)^{-1} COV\left(\left.\frac{\partial Q_{(0)}}{\partial \Theta}\right|_{\Theta=\Theta_0}\right)\left(\left.\frac{\partial^2 Q_{(0)}}{\partial \Theta \partial \Theta'}\right|_{\Theta=\Theta_1}\right)^{-1}.$$

As Θ_1 locates between $\tilde{\Theta}_{(0)}$ and Θ_0, and $\tilde{\Theta}_{(0)} \to \Theta_0$ if $T \to \infty$, then we can use $\tilde{\Theta}_{(0)}$ to instead. Under the stationary assumption we express the covariance as (see the proof in Appendix A):

$$COV\left(\tilde{\Theta}_{(0)}\right) \approx \left[\left(\sum_{j=1}^{M} \eta_j \eta_j'\right)_{\Theta=\tilde{\Theta}_{(0)}}\right]^{-1},$$

where $\eta_j = \partial \ln g(\lambda_j;\Theta) / \partial \Theta$.

4. Extension to a panel data model

This section extends the approach to a panel data model where Chen (2006)'s spectral Whittle method is used, in which the remainder disturbance is a mixed fractional ARIMA(p,d,q) model. First, let us consider a function with a one-way panel data model:

$$y_{kt} = x_{kt}'\beta + \mu_k + v_{kt} \quad \text{for } k = 1,\dots,N, \text{and } t = 1,\dots,T, \tag{9}$$

where k denotes individuals and t is time. The k subscript represents the cross-section dimension, t denotes the time-series dimension, β shows the regression coefficients if there are r exogenous variables (thus, β is $r\times1$), x_{kt} is the kt-th observation explanatory variable that is $r\times1$, and $x_{kt}' = [x_{kt,1},\dots,x_{kt,r}]$, μ_k is the individual effect, and v_{kt} is the remainder disturbance. Here, $v_k \sim N(0,\sigma_a^2\Sigma(\Theta))$ and v_k indicates the vector of the

kth individual's remainder disturbance and it has T elements. If we represent the panel data model in a vector form, then:

$y = x\beta + z_\mu \mu + v$.

In this equation y is an $NT\times 1$ vector, x is an $NT\times r$ matrix such that r is the number of explanatory variables, v is an $NT\times 1$ vector, $z_\mu = I_N \otimes l_T$, and $l_T = [1,\ldots,1]'$, $\mu = [\mu_1, \ldots, \mu_N]'$. According to Chen (2006, (3) and (7)), if the model is a fixed effects panel data model, then μ_k is a constant parameter. The quadratic form for the spectral Whittle objective function can then be expressed as:

$$Q = (y - x\beta - z_\mu \mu)'(I_N \otimes \Sigma^{-1}(\Theta))(y - x\beta - z_\mu \mu)$$
$$= \frac{4\pi}{NT} \sum_{k=1}^{N} \sum_{j=1}^{M} \frac{\left| w_{y_k}(\lambda_j) - w'_{x_k}(\lambda_j)\beta \right|^2}{g(\lambda_j;\Theta)}, \qquad (10)$$

where $w_{y_k}(\lambda_j)$ is the Fourier transform of the kth individual's dependent variable and $w_{x_k}(\lambda_j)$ is a $r\times 1$ vector that is the Fourier transform of the kth individual's explanatory variables.

If the remainder disturbance includes an additive outlier at the kth individual, then according (9) we can express the remainder disturbance as:

$u_k = \alpha_{AO} P_{AO,s} + v_k$,

where $y_k = x_k \beta + \mu_k l_T + u_k$ for $k = 1,\ldots,N$. In the same way, if the remainder disturbance includes an IO or level shift, then it can be represented respectively as:

$u_k = \alpha_{IO} P_{IO,s} + v_k$ and $u_k = \alpha_L P_{L,s} + v_k$.

Following (3), (4), and (5) the outlier estimators are obtained respectively as:

$$\tilde{\alpha}_{AO,ks} \approx \left(2\sqrt{\frac{2\pi}{T}} \sum_{j=1}^{M} \mathrm{Re}\left[e^{is\lambda_j} \frac{\left(w_{y_k}(\lambda_j) - w'_{x_k}(\lambda_j)\tilde{\beta} \right)}{g(\lambda_j;\Theta)} \right] \right) \Bigg/ \left(\frac{1}{M} \sum_{j=1}^{M} \frac{1}{g(\lambda_j;\Theta)} \right),$$

$$\tilde{\alpha}_{IO,ks} \approx 2\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\mathrm{Re}\left[e^{is\lambda_j}\frac{\left(w_{y_k}(\lambda_j)-w'_{x_k}(\lambda_j)\tilde{\beta}\right)}{g(\lambda_j;\Theta)}\frac{\theta(e^{i\lambda_j})}{\phi(e^{i\lambda_j})(1-e^{i\lambda_j})^d}\right],$$

and

$$\tilde{\alpha}_{L,ks} = 2\sqrt{\frac{2\pi}{T}}\left(\sum_{j=1}^{M}\mathrm{Re}\left[\frac{\left(w_{y_k}(\lambda_j)-w'_{x_k}(\lambda_j)\beta\right)}{g(\lambda_j;\Theta)}\frac{e^{is\lambda_j}}{1-e^{i\lambda_j}}\right]-\sum_{j=1}^{M}\frac{\left(w_{y_k}(\lambda_j)-w'_{x_k}(\lambda_j)\beta\right)}{g(\lambda_j;\Theta)}\frac{e^{i\lambda_j}}{1-e^{i\lambda_j}}\right)\times$$
$$\left(\frac{2}{T}\sum_{j=1}^{M}\frac{1}{g(\lambda_j;\Theta)}\left(\frac{\sin(T-s+1)\lambda_j/2}{\sin\lambda_j/2}\right)^2\right)^{-1},$$

where $\tilde{\alpha}_{AO,ks}$, $\tilde{\alpha}_{IO,ks}$, and $\tilde{\alpha}_{L,ks}$ represent the AO, IO, and level shift impacts at the sth observation of the kth individual, respectively. Among them we estimate β and Θ through (10). The corresponding statistics are thus represented as:

$$\tilde{\eta}_{AO,ks} = \left(2\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\mathrm{Re}\left[e^{is\lambda_j}\frac{\left(w_{y_k}(\lambda_j)-w'_{x_k}(\lambda_j)\tilde{\beta}\right)}{g(\lambda_j;\Theta)}\right]\right)\Bigg/\left(\tilde{\sigma}_{(0)}\sqrt{\frac{2}{T}\sum_{j=1}^{M}\frac{1}{g(\lambda_j;\Theta)}}\right),$$

$$\tilde{\eta}_{IO,ks} = \frac{2}{\tilde{\sigma}_a^2}\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\mathrm{Re}\left[e^{is\lambda_j}\frac{\left(w_{y_k}(\lambda_j)-w'_{x_k}(\lambda_j)\tilde{\beta}\right)}{g(\lambda_j;\Theta)}\frac{\theta(e^{i\lambda_j})}{\phi(e^{i\lambda_j})(1-e^{i\lambda_j})^d}\right],$$

and

$$\tilde{\eta}_{L,ks} = \frac{\tilde{\alpha}_{L,ks}}{\sqrt{\mathrm{var}\,\tilde{\alpha}_{L,ks}}}$$
$$\approx \frac{\left(2\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\mathrm{Re}\left[\frac{\left(w_{y_k}(\lambda_j)-w'_{x_k}(\lambda_j)\tilde{\beta}\right)}{g(\lambda_j;\Theta)}\frac{e^{is\lambda_j}}{1-e^{i\lambda_j}}\right]-2\sqrt{\frac{2\pi}{T}}\sum_{j=1}^{M}\frac{\left(w_{y_k}(\lambda_j)-w'_{x_k}(\lambda_j)\tilde{\beta}\right)}{g(\lambda_j;\Theta)}\frac{e^{i\lambda_j}}{1-e^{i\lambda_j}}\right)}{\tilde{\sigma}_a\sqrt{\left(\frac{2}{T}\sum_{j=1}^{M}\frac{1}{g(\lambda_j;\Theta)}\left(\frac{\sin(T-s+1)\lambda_j/2}{\sin\lambda_j/2}\right)^2\right)}}.$$

Following section 3 we estimate the model by an iterative procedure, in which the outliers are removed from y_{kt} for k=1,…, N and t = 1,..., T and the objective function is obtained as:

$$Q_{(0)} = \frac{4\pi}{NT}\sum_{k=1}^{N}\sum_{j=1}^{M}\frac{\left|w_{y_k^*}(\lambda_j)-w'_{x_k}(\lambda_j)\tilde{\beta}\right|^2}{g(\lambda_j;\Theta)},$$

where $Q_{(0)} \to \sigma_a^2$. Here, we shall note that if we minimize $Q_{(0)}$ to obtain $\tilde{\beta}_{(0)}$ and $\tilde{\Theta}_{(0)}$, due to $\tilde{\beta}_{(0)}$ being asymptotically orthogonal to $\tilde{\Theta}_{(0)}$, i.e, $E\left[\partial^2 Q_{(0)} / \partial\beta_{(0)} \partial\Theta'_{(0)}\right] = E\left[\left(\partial Q_{(0)} / \partial\beta_{(0)}\right)\left(\partial Q_{(0)} / \partial\Theta'_{(0)}\right)\right] = 0$, then we calculate $\tilde{\beta}_{(0)}$ and $\tilde{\Theta}_{(0)}$ separately. The properties of $\tilde{\Theta}_{(0)}$ have been discussed in the previous section and the covariance matrix is equivalent to:

$$COV\left[\tilde{\Theta}_{(0)}\right] \approx \frac{1}{N}\left[\sum_{j=1}^{M} \eta_j \eta'_j\right]^{-1}_{\Theta=\tilde{\Theta}_{(0)}}.$$

The corresponding covariance matrix of $\tilde{\beta}_{(0)}$ is obtained as:

$$\begin{aligned}
\operatorname{var}\left(\tilde{\beta}_{(0)}\right) &= \left[E\left(\frac{\partial^2 Q_{(0)}}{\partial\beta\partial\beta'}\right)\right]^{-1} E\left(\frac{\partial Q_{(0)}}{\partial\beta}\frac{\partial Q_{(0)}}{\partial\beta'}\right)\left[E\left(\frac{\partial^2 Q_{(0)}}{\partial\beta\partial\beta'}\right)\right]^{-1} \\
&\approx \sigma_a^2\left(2\pi\sum_{k=1}^{N}\left(\sum_{j=1}^{M}\frac{2\operatorname{Re}\left(I_{x_k}(\lambda_j)\right)}{g(\lambda_j;\Theta_{(0)})}\right)\right)^{-1},
\end{aligned}$$

which is devised from:

$$E\left[\frac{\partial Q_{(0)}}{\partial\beta_{(0)}}\frac{\partial Q_{(0)}}{\partial\beta'_{(0)}}\right] = 4\sigma_a^2 x'(I_N \otimes \Sigma^{-1}(\Theta_{(0)}))x \quad \text{and} \quad \frac{\partial^2 Q_{(0)}}{\partial\beta_{(0)}\partial\beta'_{(0)}} = 2x'(I_N \otimes \Sigma^{-1}(\Theta_{(0)}))x.$$

5. Simulations and power test

In this section we use Monte Carlo experiments to examine the power of the iterative outliers' detection procedure. A panel data model with a mixed remainder disturbance model is applied for the evaluations, in which two types of outliers, AO and IO, and level shifts are scattered in the dataset. As it is a time consuming task we only choose a simple model and we believe that the estimator has the same power when the parameters values are changed.

A panel data model is constructed by $y_{kt} = x_{kt}\beta + \mu_k + v_{kt}$ and the remainder disturbance is used with $(1-\phi B)(1-B)^d v_{kt} = (1-\theta)\varepsilon_{kt}$, for k = 1, …, N, t =1,…, T, where μ_k ~ NID(0,1) and ε_{kt} ~ NID(0,1). In this simulation, without loss of generality, we set the parameters with ϕ = 0.6, θ = 0.3, and d = 0.1, and the cross-section dimensions (group numbers or individual numbers) are increased from 15, to 20, and then to 25. At the same time, the time-series dimensions T are increased from 280, to 300, and then to 320.

The impacts of the outliers are implemented by $\alpha_L = \alpha_{IO} = \alpha_{AO} = 10$, and 10 events are encountered in the dataset, which includes four additive outliers, four innovation outliers, and two level shifts. The additive outliers take place at the 50th and 100th observations of the 5th and 15th individuals, and the innovation outliers are encountered at the 120th and 180th observations of the 8th and 12th individuals, and eventually the level shifts set at the 130th observations of the 10th and 13th individuals.

Table 1 exhibits the power test. From Table 1 we see that though the model estimated by the raw data with these extreme outliers ($\alpha_{IO} = \alpha_{AO} = \alpha_L = 10$) could be invalid, after the proposed iterative procedure the estimator employs a consistent property. We see that when (N, T) increases from (15, 280), to (20, 350), and then to (25, 320), the rejected probabilities for the null hypothesis ϕ = 0.6 are around 0.05. By contrast, it is obviously different for the fault null hypothesis ϕ = 0.5 that the rejected probability increases from 0.367, to 0.472, and then to 0.558. The likely properties are also found in other null hypotheses. Overall, we see that the proposed estimators are not affected by these extreme outliers, and the larger the sample size is, the more powerful the estimator will be.

Table 2 shows the rate of the correct identification. From the table we see that the true IO is easier to be misidentified when compared to AO. For instance, we see that the percentage of correct identification as an AO at (k, t) = (5, 50) of (N, T) = (15, 280) is 97.9% and the rest is misidentified as IO. In contrast, the true IO at (k, t) = (8, 120) of (N, T) = (15,280) is 90.8%, the misidentification rate is 9.2%, and these are all identified as AO, which is greater than 2.1%. In the other estimators, they have the same result. The reason could be that the estimations of parameters in each iteration have to assume that the model is free from outliers. In fact, they are not really free from the outliers. Therefore, when we apply these estimators to obtain these statistics, it will reduce the precision of outliers' detection especially for the estimation in IO. However, in the detections of level shifts we see that they are totally correct. The reason could be that the level shifts are quite different from the other types of outliers and are easy to be distinguished.

6. Case study

In this section we investigate Taiwanese local banks. Different from international business banks, the local private banks are basically conservative and most of them are family businesses. Here, we investigate 12 local banks and they are respectively Chang Hwa Bank, Hsinchu International Bank, Taitung Business Bank, Taichung Bank, The Chinese Bank, Taiwan Business Bank, Kaohsiung Bank, Cosmos Bank, Union Bank of Taiwan, Far Eastern International Bank, Tachong Bank, and Entie Bank. Weekly data between January 2000 and December 2006 from the Taiwan Stock Exchange are available for this research.

In the beginning a traditional market model for the i-th stock can be represented as:

$R_{it} = \beta_i R_{mt} + \mu_i + u_{it}$ for $i=1,\ldots, N$,

where μ_i shows the individual effects that can be regarded as an intercept or a constant term. Term β_i is the beta coefficient, a measure of systematic risk. Term R_{it} is the rate of return on stock i, and R_{mt} is the rate of return on the market portfolio, represented here by TAIWX (Taiwan Weighted Index).

Following the proposed iteration estimation we remove the outliers and level shifts. AIC (Akaike information criterion) and SBC (Schwartz Bayesian criterion) are used to choose the best model. Table 3 demonstrates the results. From the table we choose the best model $p = 2$, $q = 1$. In the following the proposed model and the traditional model are estimated for comparison:

The raw model:

$$R_{kt} = \underset{(0.0225)}{0.8327} R_{mt} + \mu_k + u_{kt}, \text{ and } (1-B)^{\overset{0.2235}{(0.0885)}}(1 - \underset{(0.0870)}{0.7804} B^1 - \underset{(0.0387)}{0.0408} B^2) u_{kt} = (1 - \underset{(0.0165)}{0.9257} B)\varepsilon_{kt},$$

where $N = 12$, $T = 359$, AIC = 3.3937, SBC = 3.4011, MSE = 29.7079.

The proposed model:

$$R_{kt} = \underset{(0.0186)}{0.7901} R_{mt} + \mu_k + u_{kt}, \text{ and } (1-B)^{\overset{0.2899}{(0.1071)}}(1 - \underset{(0.0922)}{0.6286} B - \underset{(0.0308)}{0.0775} B^2) v_{kt} = (1 - \underset{(0.0229)}{0.8945} B)\, \varepsilon_{kt},$$

where $N = 12$, $T = 359$, AIC = 3.0027, SBC = 3.0101, and MSE = 20.094.

The MSE of the proposed model is 20.094, which is obviously smaller than the raw model 29.7079. From the beta coefficient, we see that both beta coefficients are smaller than unit (0.8327 and 0.7901), which indicates these securities are defensive. Table 4 shows the outliers and level shifts locations. In this model three times of level shifts took place. Two of the level shifts occurred at February 2000 and October 2006 by Taitung Business Bank, and the other one came from Taichung Bank in January 2000. Their corresponding impacts are respectively -9.1882, -10.0069, and -10.7557 and the t-values are -4.655, -5.2156, and -4.463. This indicates these two banks have the largest structure changes and worse performances as time increases. Furthermore, Table 4 shows that there are 35 and 27 times of IO and AO scattered over these private local banks, and Taitung Business Bank occupies the largest amount of IO and AO at respectively 14 and 16. This shows it encountered radical changes during the study period. From a general viewpoint, Taitung Business Bank's reputation is questionable and it has been taken over by Central Deposit Insurance Corp. of Taiwan's government in December 2006 due to being debt-ridden.

Concluding remark

This article provides a useful approach to detect and identify the interventions in a panel data model, in which the remainder disturbance with a fractional ARMA model is implemented. In the simulation we apply a dataset encountering level shifts, additive outliers, and innovation outliers, and from the result analysis we see it is quite a successful approach when the data contain different types of interventions. Furthermore, through the case study we see it is successfully applied on local banks. From the result analysis we can easily evaluate each bank's dynamic performance in the stock market when compared to the others.

References:

1. Beran, J. (1994), Statistics for long-memory process, (New York: Chapman & Hall).
2. Bianco, A.M., Ben, M.G., E.J. Martinez, and V.J. Yohai (2001), Outlier detection in regression models with ARIMA errors using robust estimates, Journal of Forecasting, Vol. 20, pages 565-579.
3. Box, G.E.P. and Tiao, G.C. (1975), Intervention analysis with applications to economic and environmental problems, Journal of the American Statistical Association, Vol. 70, Number 349, pages 70-79.
4. Chen, W.D. (2006), Testing for Spurious Regression in a Panel Data Model with the Individual Number and Time Length Growing, Journal of Applied Statistics, Vol. 33, No. 88, pages 759-772.
5. Fox, A.J. (1972), Outliers in Time Series, Journal of the Royal Statistical Society, Series B, 43, 350-363.
6. Chang, I., Tiao, G.C., and Chen, C. (1988), Estimation of Time Series Parameters in the Presence of Outliers, Technometrics, Vol. 30, No. 2, pages 193-204.
7. Martin, R.D., and Yohai, V. (1986), Influence functions for time series, The Annals of Statistics, Vol. 14, No. 3, 781-818.
8. Pena, D. and Yohai, V. (1999), A fast procedure for outlier diagnostics in large regression problems, Journal of the American Statistical Association, Vol. 94, No. 446, Theory and Methods.
9. Rosenblatt, M. (2000), Gaussian and non-gaussian linear time series and random fields, (New York: Springer-Verlag).
10. Yohai, V. (1987), High breakdown-point and high efficiency robust estimates for regression, The Annals of Statistics, Vol. 15, No. 20, 642-656.

Appendix A

As y_t^* indicates that y_t has been removed the outliers, let $\pi(B)y_t^* = a_t^*$, and then according to (5) we obtain the derivative as:

$$\frac{\partial Q_{(0)}}{\partial \Theta} = -\frac{4\pi}{T}\sum_{j=1}^{M}\frac{I_{y^*}(\lambda_j)}{g^2(\lambda_j;\Theta)}\frac{\partial g(\lambda_j;\Theta)}{\partial \Theta} = -\frac{4\pi}{T}\sum_{j=1}^{M}\frac{I_{y^*}(\lambda_j)}{g(\lambda_j;\Theta)}\eta_j = -\frac{4\pi}{T}\sum_{j=1}^{M} I_{a^*}(\lambda_j)\eta_j ,$$

$$\frac{\partial^2 Q_{(0)}}{\partial \Theta \partial \Theta'} = \frac{8\pi}{T}\sum_{j=1}^{M}\frac{I_{y^*}(\lambda_j)}{g^3(\lambda_j;\Theta)}\frac{\partial g(\lambda_j;\Theta)}{\partial \Theta}\frac{\partial g(\lambda_j;\Theta)}{\partial \Theta'} - \frac{4\pi}{T}\sum_{j=1}^{M}\frac{I_{y^*}(\lambda_j)}{g^2(\lambda_j;\Theta)}\frac{\partial^2 g(\lambda_j;\Theta)}{\partial \Theta \partial \Theta'} .$$

As $\sum_{j=1}^{M}\ln g(\lambda_j;\Theta) = 0$, we obtain:

$$\sum_{j=1}^{M}\frac{1}{g(\lambda_j;\Theta)}\frac{\partial^2 g(\lambda_j;\Theta)}{\partial \Theta \partial \Theta'} = \sum_{j=1}^{M}\frac{\partial \ln g(\lambda_j;\Theta)}{\partial \Theta}\frac{\partial \ln g(\lambda_j;\Theta)}{\partial \Theta'} .$$

From Robsenblatt (2000) we know that $E\left[I_{y^*}(\lambda_j)\right] \to \frac{\sigma_a^2}{2\pi} g(\lambda_j;\Theta)$ for $T \to \infty$, and thus we obtain:

$$E\left(\frac{\partial^2 Q_{(0)}}{\partial \Theta \partial \Theta'}\right) \to \frac{2\sigma_a^2}{T}\sum_{j=1}^{M}\eta_j \eta_j' \quad \text{for } T \to \infty,$$

where $\eta_j = \partial \ln g(\lambda_j;\Theta) / \partial \Theta$. As a_t^* can be regarded as a white noise, therefore we obtain:

$$I_{a^*}(\lambda_j) \sim \frac{\sigma_a^2}{2\pi}\frac{1}{2}\chi_2^2 , \text{ and } \operatorname{cov}\left(I_{a^*}(\lambda_j), I_{a^*}(\lambda_k)\right) = 0 \quad \text{for } j \neq k.$$

From the first condition we know that $E\left[\frac{\partial Q_{(0)}}{\partial \Theta}\right] = -E\left[\frac{4\pi}{T}\sum_{j=1}^{M} I_{a^*}(\lambda_j)\eta_j\right] = 0$, and therefore we have the following relationship:

$$
\begin{aligned}
\operatorname{cov}\left[\frac{\partial Q_{(0)}}{\partial \Theta}\right] &= E\left[\frac{\partial Q_{(0)}}{\partial \Theta}\frac{\partial Q_{(0)}}{\partial \Theta'}\right] \\
&= E\left[\left(-\frac{4\pi}{T}\sum_{j=1}^{M}\frac{I_{y^*}(\lambda_j)}{g(\lambda_j;\Theta)}\eta_j\right)\left(-\frac{4\pi}{T}\sum_{j=1}^{M}\frac{I_{y^*}(\lambda_j)}{g(\lambda_j;\Theta)}\eta_j'\right)\right] \\
&= E\left[\left(-\frac{4\pi}{T}\sum_{j=1}^{M}I_{a^*}(\lambda_j)\eta_j\right)\left(-\frac{4\pi}{T}\sum_{j=1}^{M}I_{a^*}(\lambda_j)\eta_j'\right)\right] \\
&= \frac{4\sigma_a^4}{T^2}\sum_{j=1}^{M}\eta_j\eta_j'.
\end{aligned}
$$

That is:

$$
\begin{aligned}
COV\left(\tilde{\Theta}_{(0)}\right) &= \left(\left.\frac{\partial^2 Q_{(0)}}{\partial \Theta \partial \Theta'}\right|_{\Theta=\Theta_1}\right)^{-1} COV\left(\left.\frac{\partial Q_{(0)}}{\partial \Theta}\right|_{\Theta=\Theta_0}\right)\left(\left.\frac{\partial^2 Q_{(0)}}{\partial \Theta \partial \Theta'}\right|_{\Theta=\Theta_1}\right)^{-1} \\
&\approx \left[\left(\sum_{j=1}^{M}\eta_j\eta_j'\right)_{\Theta=\tilde{\Theta}_{(0)}}\right]^{-1}.
\end{aligned}
$$

Table 1. Empirical size and power test based on 1000 replications for the model $y_{kt} = x_{kt}\beta + \mu_k + v_{kt}$, where $(1-\phi B)v_{kt}=(1-B)^{-d}(1-\theta B)\varepsilon_{kt}$, $\mu_k \sim NID(0,1)$, $\varepsilon_{kt} \sim NID(0, 1)$. The sizes are respectively (N, T) = (15, 280), (20, 300), and (25, 320), and the parameters are respectively $\phi = 0.6$, $\theta = 0.3$, and $d = 0.1$. The outlier parameters are $\alpha_{IO} = \alpha_{AO} = \alpha_{LS} = 10$ and the additive outliers are scattered at the 50th and 100th observations for the 5th, and 15th individuals, and simultaneously the innovation outliers are set at the 120th and 180th observations for the 8th and 12th individuals, and the level shifts are set at 130 observation of the 10th and 13th individuals. The critical value c = 4.0. The power test is based on 1000 replications, the significance level is at 5%, and the two-tailed standard normal distribution test is used.

H_0: $\phi =$	*0.40*	*0.45*	*0.50*	*0.55*	*0.60*	*0.65*	*0.70*	*0.75*	*0.80*	*0.85*	*0.90*
(15, 280)	0.763	0.602	0.367	0.157	0.051	0.091	0.335	0.774	0.969	0.994	0.999
(20, 300)	0.874	0.723	0.472	0.196	0.058	0.153	0.514	0.909	0.994	1.000	0.999
(25,320)	0.954	0.831	0.558	0.224	0.055	0.179	0.692	0.976	0.999	1.000	1.000
H_0: $\theta =$	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45	0.50	0.55	1.30
(15, 280)	0.946	0.846	0.545	0.153	0.064	0.292	0.694	0.931	0.991	0.999	1.000
(20, 300)	0.989	0.942	0.703	0.245	0.060	0.395	0.838	0.984	1.000	1.000	1.000
(25,320)	0.999	0.985	0.853	0.308	0.045	0.451	0.938	0.998	1.000	1.000	1.000
H_0: $d =$	-0.10	-0.05	0.00	0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40
(15, 280)	0.883	0.730	0.449	0.202	0.096	0.123	0.379	0.762	0.943	0.999	1.000
(20, 300)	0.953	0.825	0.583	0.242	0.076	0.183	0.546	0.910	0.999	1.000	1.000
(25,320)	0.991	0.925	0.664	0.281	0.073	0.217	0.713	0.980	0.999	1.000	1.000
H_0: $\beta =$	0.80	0.85	0.90	0.95	1.00	1.05	1.10	1.15	1.20	1.25	1.30
(15, 280)	1.000	1.000	1.000	0.921	0.041	0.921	1.000	1.000	1.000	1.000	1.000
(20, 300)	1.000	1.000	1.000	0.982	0.048	0.982	1.000	1.000	1.000	1.000	1.000
(25,320)	1.000	1.000	1.000	0.998	0.051	0.997	1.000	1.000	1.000	1.000	1.000

The means of these estimates of ϕ for different sizes (15, 280), (20, 300), and (25, 320) are respectively 0.5874, 0.5915, and 0.5934, and their standard deviations are respectively 0.0732, 0.0587, and 0.0475. For θ, the means are respectively 0.2881, 0.2907, and 0.2932, and standard deviations are respectively 0.0477, 0.0398, and 0.0305. For d, the means are respectively 0.1007, 0.0988, and 0.0993, and standard deviations are respectively 0.0674, 0.0555, and 0.0433. For β, the means are respectively 0.9997, 0.9999, and 0.9997, and standard deviations are respectively 0.0148, 0.0124, and 0.0104. This indicates that even the data contain the outliers, and the estimator by using the iterative procedure to remove these outliers still employs consistent properties.

Table 2. The identifications for the outliers' simulations, which are based on 1000 replications for the model $y_{kt} = x_{kt}\beta + \mu_k + v_{kt}$, where $(1-\phi B)v_{kt} = (1-B)^{-d}(1-\theta B)\varepsilon_{kt}$, $\mu_k \sim NID(0,1)$, $\varepsilon_{kt} \sim NID(0, 1)$. The sizes are respectively (N, T) = (15, 280), (20, 300), and (25, 320), and the parameters are respectively $\phi = 0.6$, $\theta = 0.3$, and $d = 0.1$. The outlier parameters are $\alpha_{IO} = \alpha_{AO} = \alpha_{LS} = 10$ and the additive outliers are scattered at the 50th and 100th observations for the 5th, and 15th individuals, and the innovation outliers are set at the 120th and 180th observations for the 8th and 12th individuals, and the level shift are set at 130 at the 10th and 13th individuals. The critical value $c = 4.0$.

(*N, Time*)	True outlier	Percentage for correct and wrong identifications				
(15, 280)	AO	Location (*k*, *t*)	(5,50)	(5,100)	(15,50)	(15,100)
		Correct (identified as AO)	97.9%	98.2%	97.3%	98.4%
		Wrong (identified as IO)	2.1%	1.8%	2.7%	1.6%
	IO	Location (*k*, *t*)	(8,120)	(8,180)	(12,120)	(12,180)
		Correct (identified as IO)	90.8%	91.8%	91.6%	90.6%
		Wrong (identified as AO)	9.2%	8.2%	8.4%	9.4%
	level shift	Location (*k*,*t*)	(10,130)		(13,130)	
		Correct (identified as LS)	100.0%		100.0%	
(20, 300)	AO	Location (*k*, *t*)	(5,50)	(5,100)	(15,50)	(15,100)
		Correct (identified as AO)	97.5%	98.1%	97.8%	97.4%
		Wrong (identified as IO)	2.5%	1.9%	2.2%	2.6%
	IO	Location (*k*, *t*)	(8,120)	(8,180)	(12,120)	(12,180)
		Correct (identified as IO)	91.6%	91.8%	91.3%	91.6%
		Wrong (identified as AO)	8.4%	8.2%	8.7%	8.4%
	level shift	Location (*k*,*t*)	(10,130)		(13,130)	
		Correct (identified as LS)	100.0%		100.0%	
(25, 320)	AO	Location (*k*, *t*)	(5,50)	(5,100)	(15,50)	(15,100)
		Correct (identified as AO)	98.5%	98.2%	97.7%	98.0%
		Wrong (identified as IO)	1.5%	1.8%	2.3%	2.0%
	IO	Location (*k*, *t*)	(8,120)	(8,180)	(12,120)	(12,180)
		Correct (identified as IO)	91.2%	91.1%	92.4%	92.9%
		Wrong (identified as AO)	8.8%	8.9%	7.6%	7.1%
	level shift	Location (*k*,*t*)	(10,130)		(13,130)	
		Correct (identified as LS)	100.0%		100.0%	

In the table we can see the identifications of AO rarely fail. For instance, when (N,Time) = (15, 280) the rate of correctly identification at the location $k = 5$, $t = 50$ is 97.9%, and similar results are seen in the other locations. In contrast, IO identifications are easy to be misidentified, e.g., the correct identification rate at $k = 8$, $t = 120$ of (N, Time) = (15, 280) is 90.8%. At the same time, we also see that the identifications of level shifts are 100% correct.

Table 3. The summary of AICs (or SBCs) for different remainder disturbance ARIMA(p,d,q) models.

Model: $R_{kt} = \beta R_{mt} + \mu_k + v_{kt}$,

$$(1-\phi_1 B - ... - \phi_p B^p)(1-B)^d v_{kt} = (1-\theta_1 B - ... - \theta_q B^q)\varepsilon_{kt} ,$$

where $\varepsilon_{kt} \sim NID(0, \sigma_\varepsilon^2)$.

	$q = 0$	$q = 1$	$q = 2$
$p = 0$	3.0098 (3.0128)	3.0047 (3.0091)	3.0058 (3.0117)
$p = 1$	3.0042 (3.0087)	3.0074 (3.0133)	3.0063 (3.0137)
$p = 2$	3.0064 (3.0123)	<u>3.0027</u> (<u>3.0101</u>)	3.0031 (3.0120)
$p = 3$	3.0409 (3.0483)	3.0285 (3.0374)	3.0086 (3.0189)

Figures in the parentheses are SBC, and p and q indicate respectively the autoregressive and moving average orders, whereby d could be a real number and the interventions have been removed.

Table 4. The AO and IO are estimated for the bank industry, in which we implement a dataset with 12 banks. Weekly data from January 2000 to December 2006 are used, in which 359 time periods and 12 banks' data are obtained from the Taiwan Stock Exchange.

Individuals	Type	Time points
Chang Hwa Bank	IO	55
Hsinchu International Bank	AO	119, 160, 352
Taitung Business Bank	AO	48, 55, 101, 109, 117, 141, 149, 163, 201, 208, 245, 254, 339, 340, 352, 357
	IO	4, 84, 116, 119, 140, 143, 150, 159, 160, 166, 244, 330, 331, 334
	Level shift	7, 353
Taichung Bank	AO	48, 113, 160
	IO	119
	Level shift	5
The Chinese Bank	AO	357
Taiwan Business Bank	AO	117
	IO	119, 160, 215, 297
Kaohsiung Bank	IO	351
Cosmos Bank	AO	119
	IO	4,50,116,227
Union Bank of Taiwan	IO	4,26,160
Far Eastern International Bank	IO	4,90
Tachong Bank	AO	163
	IO	4,120,160
Entie Bank	AO	27
	IO	4,163

There are three times of level shifts, respectively at the 7^{th} and 353^{th} observations of Taitung Business Bank and the 5^{th} observation of Taichung Bank. The critical value is 4.0.