Biometrika (2000), **87**, 4, *pp.* 789–804

Outliers in multivariate time series

By RUEY S. TSAY

Graduate School of Business, University of Chicago, Chicago, Illinois 60637, U.S.A.
ruey.tsay@gsb.uchicago.edu

DANIEL PEÑA

Department of Statistics and Econometrics, Universidad Carlos III, Cl. Madrid 126, 28903 Getafe, Spain
dpena@est-econ.uc3m.es

AND ALAN E. PANKRATZ

Department of Economics and Management, DePauw University, Greencastle, Indiana 46135, U.S.A.
graywolf@depauw.edu

Summary

This paper generalises four types of disturbance commonly used in univariate time series analysis to the multivariate case, highlights the differences between univariate and multivariate outliers, and investigates dynamic effects of a multivariate outlier on individual components. The effect of a multivariate outlier depends not only on its size and the underlying model, but also on the interaction between the size and the dynamic structure of the model. The latter factor does not appear in the univariate case. A multivariate outlier can introduce various types of outlier for the marginal component models. By comparing and contrasting results of univariate and multivariate outlier detections, one can gain insights into the characteristics of an outlier. We use real examples to demonstrate the proposed analysis.

Some key words: Additive outlier; Innovational outlier; Level shift; Temporary change.

1. Introduction

In the time series literature, outlier detection plays an important role in modelling, inference and even data processing because outliers can lead to model misspecification, biased parameter estimation and poor forecasts. As a specific example, outlier detection has become a key feature in recent advances in seasonal adjustment and in automatic, time series model identification; see the new adjustment procedure X-12 ARIMA of Findley et al. (1998), which is used by the U.S. Government, and the SEATS and TRAMO programs, which were recently adopted by the European Union and which are described in working papers of the European University Institute, Florence, by V. Gomez and A. Maravall. In an extension of the work of Fox (1972), four types of outlier have been proposed for univariate time series analysis. They are additive outliers, innovational outliers, level shifts and temporary changes. These four types of outlier affect an observed time series and its residual process differently; see Chang, Tiao & Chen (1988), Chen & Liu (1993), Tsay (1988) and the references therein. Several methods are available to detect outliers; see McCulloch & Tsay (1994) for a Bayesian approach and Chen & Liu (1993) and the references therein for non-Bayesian methods.

However, most outlier studies in time series analysis focus on a single series. A common practice for handling outliers in a multivariate process is to apply univariate techniques to the component series to remove outlier effects, then treat the adjusted series as outlier-free and model them jointly. This procedure encounters several difficulties. First, in a multivariate process, an outlier of a component may be caused by an outlier in the other components. Overlooking such a possibility can easily lead to overspecification of the number of outliers. Secondly, an outlier of moderate size affecting all the components may be unnoticed in the univariate analyses because univariate methods fail to combine information about the outlier among the component series. This outlier will be more easily detected in multivariate analysis. Thirdly, univariate detection procedures often use inferior estimation, because the joint dynamics of the series are not properly taken into account.

Pankratz (1993) considers additive and innovational outliers in a dynamic regression model with a single input and a single output. He classifies outliers in the input series as passed and non-passed outliers and uses a weighted average of least squares estimators to estimate non-passed outliers. The approach becomes complicated when there are multiple input or multiple output series.

In this paper we study outliers directly under a multivariate framework and analyse the effects of a multivariate outlier on the joint and marginal models. By comparing and contrasting results of univariate and multivariate detection methods, we can gain insight into the characteristics of an outlier. We shall demonstrate this later by a real example.

The paper is organised as follows. We generalise the four types of outlier to the multivariate case in § 2 and briefly discuss the effects of multivariate outliers on the joint and marginal models. The effects depend not only on the outlier size and the model, but also on the interaction between the two. In § 3 we present an iterative procedure for estimating multivariate outliers based on two test statistics. The first test statistic is a joint statistic that combines information across components, and the second test statistic is marginal and uses information contained in an individual component. In § 4 we use simulation to obtain finite-sample critical values and power of the test statistics. Section 5 contains two real examples.

2. Outliers in a vector time series

2·1. *Preliminaries*

Let $x_t = (x_{1t}, \ldots, x_{kt})'$ be a k-dimensional time series that follows a vector autoregressive integrated moving-average, ARIMA, model

$$\Phi(B)x_t = c + \Theta(B)\varepsilon_t, \qquad (1)$$

where

$$\Phi(B) = I - \Phi_1 B - \ldots - \Phi_p B^p, \quad \Theta(B) = I - \Theta_1 B - \ldots - \Theta_q B^q$$

are $k \times k$ matrix polynomials of finite degrees p and q, B is the backshift operator such that $Bx_t = x_{t-1}$, c is a k-dimensional constant vector, and $\{\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{kt})'\}$ is a sequence of independent and identically distributed Gaussian random vectors with zero mean and positive-definite covariance matrix Σ. We assume that $\Phi(B)$ and $\Theta(B)$ are left coprime and

matrix because x_t is a k-dimensional time series.

that all of the zeros of the determinants $|\Phi(B)|$ and $|\Theta(B)|$ are on or outside the unit circle. In addition, if $|\Phi(z)|=0$ for some $|z|=1$, we assume that the series x_t starts at a fixed time point t_0 with fixed initial values and initial innovations. The series x_t is asymptotically stationary if $|\Phi(z)| \neq 0$ for all $|z|=1$ and is unit-root nonstationary if $|\Phi(1)|=0$. Similarly, x_t is invertible if $|\Theta(z)| \neq 0$ for all $|z|=1$ and is non-invertible otherwise; see Li & Tsay (1998) for further information. In practice, the unit-root nonstationary case is of major interest. ?

CONSTRUCTION STEPS

For the vector ARIMA model in (1), define the autoregressive representation as

$$\Pi(B)x_t = c_0 + \varepsilon_t, \qquad (2)$$

→ AR REPRESENTATION

where $c_0 = \{\Theta(1)\}^{-1}c$ if x_t is invertible and is a function of t, c, Θ_i, initial values and initial innovations if x_t is non-invertible,

$$\Pi(B) = I - \sum_{i=1}^{\infty} \Pi_i B^i = \{\Theta(B)\}^{-1}\Phi(B),$$

and it is understood that $y_t = 0$ if $t < t_0$, where t_0 is the starting time point of the series. Also, define the moving-average representation as $\boxed{x_t = c_* + \Psi(B)\varepsilon_t}$, where

$$\Psi(B) = I + \sum_{i=1}^{\infty} \Psi_i B^i = \{\Phi(B)\}^{-1}\Theta(B),$$

→ MA REPRESENTATION

and $c_* = \{\Phi(1)\}^{-1}c = E(x_t)$ if x_t is stationary and is a function of t, c, Φ_i, initial values and initial innovations if x_t is unit-root nonstationary. Obviously, we have $\Pi(B)\Psi(B) = \Psi(B)\Pi(B) = I$, $\Pi(1)c_* = c_0$ and $\Psi(1)c_0 = c_*$.

Let $\xi_t^{(h)}$ be the indicator variable for time index h, that is $\xi_h^{(h)} = 1$ and $\xi_t^{(h)} = 0$ if $t \neq h$. Denote the observed time series by $y_t = (y_{1t}, \ldots, y_{kt})'$, and let $\omega = (\omega_1, \ldots, \omega_k)'$ be the size of the initial impact of an outlier on the series x_t. The four types of univariate outlier can be generalised to the multivariate case in a direct manner:

OBSERVED TIME SERIES (WITH OUTLIERS) REAL TIME SERIES (WITHOUT OUTLIERS DUE TO POOR MEASUREMENT)

$$y_t = x_t + \alpha(B)\omega\xi_t^{(h)}, \qquad (3)$$

where $\alpha(B) = \Psi(B)$ for a multivariate innovational outlier, $\alpha(B) = I$ for a multivariate additive outlier, $\alpha(B) = (1-B)^{-1}I$ for a multivariate level shift, and $\alpha(B) = \{D(\delta)\}^{-1}$ for a multivariate temporary change, where $D(\delta)$ is a $k \times k$ diagonal matrix with diagonal elements $\{(1-\delta_1 B), \ldots, (1-\delta_k B)\}$ and $0 < \delta_i < 1$. For simplicity, we shall assume that $\delta_1 = \ldots = \delta_k = \delta$.

2·2. *Effects of multivariate outliers* IMPORTANT ?

The effects of multivariate outliers in (3) are in general similar to those of the univariate case, but substantial differences exist in some cases. For illustration, consider the case of a multivariate innovational outlier. Suppose that the true model is $x_t = (I - \Theta_1 B)\varepsilon_t$ with $|\Theta_1| = 0$. This vector MA(1) model can occur in practice, especially when the dimension k is large and many of the elements in Θ_1 are zeros. If a multivariate innovational outlier occurs at time h with size ω belonging to the right null space of Θ_1, then, as $\Theta_1\omega = 0$, the outlier only affects a single observation at time h, and hence it is equivalent to a multivariate additive outlier. For higher-order models, the differences between multivariate and univariate cases can be more substantial.

Assume next that the model of x_t is known. Define a filtered series $\{a_t\}$ by

$$a_t = y_t - \sum_{i=1}^{p} \Phi_i y_{t-i} - c + \sum_{j=1}^{q} \Theta_j a_{t-j} \quad (t = t_0, t_0+1, \ldots),$$

where $y_t = x_t$ and $a_t = \varepsilon_t$ for $t < t_0$. By definition, if there exists no outlier, then $a_t = \varepsilon_t$. In the presence of outliers, $a_t \neq \varepsilon_t$ for some time points. Multiplying equation (3) on the left by $\Pi(B)$ and subtracting c_0 from both sides of the equation, we have

$$a_t = \varepsilon_t + \Pi(B)\alpha(B)\omega\xi_t^{(h)}. \tag{4}$$

In the univariate case the effect of an outlier on the filtered series depends only on $\Pi(B)\alpha(B)$. However, in the multivariate case the effect also depends on the interaction between $\Pi(B)\alpha(B)$ and ω. Again, we shall illustrate the situation by considering a simple, but important, model.

Suppose that x_t follows a unit-root nonstationary vector AR(p) model, which is commonly used in macro-economic applications, and that the outlier is a multivariate level shift at time $t = h$. In this case, $\Pi(B) = \Phi(B)$ in (1) and $\alpha(B) = (1 - B)^{-1}I$. Equation (4) then becomes $a_t = \varepsilon_t + \Pi^*(B)\omega\xi_t^{(h)}$, where the coefficient matrices Π_i^* of $\Pi^*(B)$ are $\Pi_i^* = \sum_{j=1}^{i}\Phi_j - I$ for $i = 1, \ldots, p$ and $\Pi_i^* = \Pi_p$ for $i > p$. In particular, we have $\Pi_p^* = -\Phi(1)$, which satisfies $|\Phi(1)| = 0$ under the unit-root assumption. If $\Phi(1) = 0$, then all component series of x_t are unit-root nonstationary, i.e. there is no cointegration, and the multivariate level shift only affects a_t for $t = h, \ldots, h + p - 1$. This is similar to the univariate case. If $\Phi(1) \neq 0$, then there is cointegration in x_t and we have $\Phi(1) = -\gamma\beta$, where γ and β are $k \times s$ and $s \times k$ matrices and s is the rank of $\Phi(1)$. Let $\beta_\perp$ be a $k \times (k - s)$ full-rank orthogonal matrix of β such that $\beta\beta_\perp = 0$. The effect of the multivariate level shift on a_t then depends on ω as follows.

(i) If ω is a linear combination of columns of $\beta_\perp$, then $\Pi_i^*\omega = -\Phi(1)\omega = 0$ for all $i \geqslant p$. The multivariate level shift only affects a_t for $t = h, \ldots, h + p - 1$. This is similar to the univariate case with a unit root.

(ii) If ω is not a linear combination of columns of $\beta_\perp$, then $\Phi(1)\omega \neq 0$ and hence $\Pi_i^*\omega \neq 0$ for $i \geqslant p$. The multivariate level shift then affects all a_t for $t \geqslant h$. Consequently, the unit roots do not affect the impact of the multivariate level shift on a_t, and the outlier effect is very different from that of the univariate case.

2·3. *Implications for marginal models*

From the definitions, a multivariate additive outlier or level shift or temporary change for x_t is also an additive outlier or level shift or temporary change for the component series x_{it} provided that $\omega_i \neq 0$. However, a multivariate innovational outlier can introduce different configurations for the marginal models of individual components. The implications depend on the vector model in (1), the outlier size ω and the interaction between ω and the model. We discuss two special cases.

For a vector moving-average model, a multivariate innovational outlier may introduce a univariate additive or innovational outlier or a patch of outliers for the marginal models. For illustration, consider a simple bivariate MA(1) model with a multivariate innovational outlier:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} 1 - \Theta_{11}B & -\Theta_{12}B \\ -\Theta_{21}B & 1 - \Theta_{22}B \end{bmatrix} \left(\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} + \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} \xi_t^{(h)} \right).$$

The marginal model for x_{1t} is a univariate MA(1) model, $x_{1t} = (1 - \theta B)e_{1t}$, say, where e_{1t} is a white noise sequence with mean zero and variance σ_e^2 and the parameters θ and σ_e^2 are determined by the relationships

$$(1 + \Theta_{11}^2)\sigma_{11} + \Theta_{12}^2\sigma_{22} + 2\Theta_{11}\Theta_{12}\sigma_{12} = (1 + \theta^2)\sigma_e^2, \quad -\Theta_{11}\sigma_{11} - \Theta_{12}\sigma_{12} = -\theta\sigma_e^2,$$

where σ_{ij} is the (i, j)th element of $\mathrm{cov}(\varepsilon_t)$. The marginal model for the observed series y_{1t} is then

$$y_{1t} = (1-\theta B)e_{1t} + \omega_1 \xi_t^{(h)} - (\Theta_{11}\omega_1 + \Theta_{12}\omega_2)\xi_{t-1}^{(h)}, \tag{5}$$

which may lead to several scenarios as follows.

(i) If $\Theta_{11}\omega_1 + \Theta_{12}\omega_2 = 0$, the model in (5) reduces to

$$y_{1t} = (1-\theta B)e_{1t} + \omega_1 \xi_t^{(h)} = x_{1t} + \omega_1 \xi_t^{(h)},$$

and the outlier becomes a univariate additive outlier at time h.

(ii) If $\Theta_{11}\omega_1 + \Theta_{12}\omega_2 = \theta\omega_1$, the univariate model for y_{1t} becomes

$$y_{1t} = (1-\theta B)(e_{1t} + \omega_1 \xi_t^{(h)}),$$

and hence the outlier is a univariate innovational outlier.

(iii) The parameter ω_2 in the last term of the right-hand side of equation (5) may assume values for which the outlier effect on y_{1t} cannot be written as $(1-\theta B)\omega_1\xi_t^{(h)}$. In this case, the multivariate innovational outlier introduces two consecutive outliers at time indices h and $h+1$ for the marginal model of y_{1t}.

Consider next a bivariate AR(1) model with a multivariate innovational outlier. The observed series becomes

$$\begin{bmatrix} 1-\Phi_{11}B & -\Phi_{12}B \\ -\Phi_{21}B & 1-\Phi_{22}B \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} \xi_t^{(h)} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix}.$$

Pre-multiplying the model by the adjoint matrix of the autoregressive matrix polynomial and after some algebra, one can easily obtain a univariate marginal model for each component. For instance, we have

$$y_{1t} = x_{1t} + \frac{\{\omega_1 + (\Phi_{12}\omega_2 - \Phi_{22}\omega_1)B\}\xi_t^{(h)}}{1 - g_1 B - g_2 B^2},$$

where $g_1 = \Phi_{11} + \Phi_{22}$, $g_2 = |\Phi_1|$, x_{1t} follows the ARMA(2, 1) model

$$(1 - g_1 B - g_2 B^2)x_{1t} = (1-\theta_1 B)e_{1t},$$

with θ_1 and $\mathrm{var}(e_{1t})$ determined by parameters of the vector AR(1) model, and $\{e_{1t}\}$ is a white noise series. The impact of the multivariate innovational outlier on x_{1t} can appear in several forms:

(a) if $\omega_1 + (\Phi_{12}\omega_2 - \Phi_{22}\omega_1)B = \omega_1(1-\theta_1 B)$, then x_{1t} has an innovational outlier at time index h;

(b) if $\{\omega_1 + (\Phi_{12}\omega_2 - \Phi_{22}\omega_1)B\}/(1 - g_1 B - g_2 B^2) = \omega_1$, then x_{1t} has an additive outlier at time index h;

(c) if $\{\omega_1 + (\Phi_{12}\omega_2 - \Phi_{22}\omega_1)B\}/(1 - g_1 B - g_2 B^2) = \omega_1/(1-\eta B)$, for some $0 < \eta < 1$, then x_{1t} has a temporary change at time index h.

In general, the disturbance becomes a special case of the intervention of Box & Tiao (1975).

In summary, a multivariate innovational outlier may introduce various outlier configurations for the component series. In some cases, it leads to a patch of outliers in the marginal component models with patch length determined by various factors. This result can help explain the empirical finding that univariate outlier detection often identifies consecutive outliers; see Example 1 of § 5.

The results of this subsection demonstrate that it is easier and more fruitful to study outliers in a multivariate framework. In addition, a vector series, when considered jointly, contains more information about an outlier than does a univariate series. Consider the case of a bivariate system (x_{1t}, x_{2t}) in which x_{1t} is the input and x_{2t} the output. Suppose an additive outlier is detected at time h in the univariate analysis of the input series. We cannot tell from the analysis whether the outlier was due to (I) a recording or measurement error or (II) an intervention that really changed the value of the series at time index h. However, if the multivariate analysis also shows significant outlier effects in the output series at the same time index, then it is more likely that the additive outlier is caused by an intervention that affects both series. On the other hand, if the multivariate analysis fails to show any significant outlier effect in the output series at time index h, then the chance of a recording error increases, because the analysis shows that the output series is consistent with the outlier-adjusted input series; see Example 2 of § 5 for an illustration.

3. A DETECTION PROCEDURE

In practice, the number, locations and types of outliers are unknown a priori, and we use an iterative procedure similar to that of the univariate case to detect multivariate outliers. Assuming no outlier at the very beginning, we build a multivariate ARIMA model for the series under study and let $\hat{a}_t$ be the estimated residuals and $\hat{\Pi}_i$ the estimated coefficients of the autoregressive representation. Then at each time point the effect of each type of outlier can be estimated as follows. For a multivariate innovational outlier at time index h, all information about the outlier is contained in $\hat{a}_h$, and we estimate the outlier size by using $\hat{\omega}_{I,h} = \hat{a}_h$, where the subscript I indicates innovational outlier. For the other types of outlier, the same estimation idea applies, and we shall give details for the case of a multivariate additive outlier only. In this case we have

STEP 1

STEP 2 $\hat{w}$

$$\hat{a}_t = \left(I - \sum_{i=1}^{\infty} \hat{\Pi}_i B^i\right) \xi_t^{(h)} \omega + \varepsilon_t = \left(\xi_t^{(h)} - \sum_{i=1}^{\infty} \hat{\Pi}_i \xi_{t-i}^{(h)}\right) \omega + \varepsilon_t,$$

→ estimation of a_t

for innovation outlier no need to use this formula

where $\varepsilon_t \sim N(0, \Sigma)$, and the estimator of ω is

$$\hat{\omega}_{A,h} = -\left(\sum_{i=0}^{n-h} \hat{\Pi}_i' \Sigma^{-1} \hat{\Pi}_i\right)^{-1} \sum_{i=0}^{n-h} \hat{\Pi}_i' \Sigma^{-1} \hat{a}_{h+i} \quad (\Pi_0 = -I),$$

ADDITIVE

→ estimator of w (effect of outlier) for ADDITIVE

which can readily be interpreted as a generalised least squares estimator, as in the context of multivariate seemingly unrelated regression model estimation. In addition, the covariance matrix of the estimator is $\Sigma_{A,h} = (\sum_{i=0}^{n-h} \hat{\Pi}_i' \Sigma^{-1} \hat{\Pi}_i)^{-1}$.

To test the significance of a multivariate outlier at time index h, we consider the null hypothesis $H_0: \omega = 0$ versus the alternative hypothesis $H_a: \omega \neq 0$. Two test statistics are used. The first is

STEP 3

$$J_{i,h} = \hat{\omega}_{i,h}' \Sigma_{i,h}^{-1} \hat{\omega}_{i,h},$$

→ FIRST TEST STATISTIC

where $i = I$, A, L or T, depending on the type of outlier: innovational, additive, level shift or temporary change. This statistic treats components of ω jointly. For a fixed h, and assuming that the model is known, $J_{i,h}$ is distributed as a chi-squared random variable with k degrees of freedom under the null hypothesis. The second test statistic is the maximum z-statistic, in absolute value, of the components of $\hat{\omega}_{i,h}$ when $\Sigma_{i,h}$ is known:

$$C_{i,h} = \max_{1 \leqslant j \leqslant k} |\hat{\omega}_{j,i,h}| / \sqrt{\sigma_{j,i,h}},$$

→ SECOND TEST STATISTIC

where $\hat{\omega}_{j,i,h}$ and $\sigma_{j,i,h}$ are the jth element of $\hat{\omega}_{i,h}$ and the (j, j)th element of $\Sigma_{i,h}$ respectively.

STEP 4

We define the overall test statistics as

$$J_{\max}(i, h_i) = \max_h J_{i,h}, \quad C_{\max}(i, h_i^*) = \max_h C_{i,h} \quad (i = I, A, L, T), \tag{6}$$

where h_i denotes the time index when the maximum of test statistic $J_{i,h}$ occurs and h_i^* denotes the time index when the maximum of $C_{i,h}$ occurs. Under the null hypothesis of no outlier in the sample and if we assume that the model of x_t is known, $J_{\max}(I, h_I)$ is the maximum of a random sample of size n from a chi-squared distribution with k degrees of freedom. Thus, the asymptotic distribution of $J_{\max}(I, h_I)$ can be obtained using the extreme value distribution. Each of the other three joint test statistics in (6) is the maximum of a dependent sample from a chi-squared distribution with k degrees of freedom. Their asymptotic distributions are therefore more complicated, depending on the serial dependence of $\{J_{i,h}\}$. From the estimation of the outlier parameter ω, it is seen that the serial correlations of $\{J_{L,h}\}_{h=1}^{n}$ are stronger than those of $\{J_{i,h}\}$ for $i = I$, A and T. This is because of the nondecaying weights induced by the operator $1/(1-B)$ so that $\hat{\omega}_{L,h}$ contains all of the filtered values $\hat{a}_t$ for $t \geqslant h$. Consequently, the asymptotic distribution of $J_{\max}(L, h_L)$ is more concentrated than those of the other three joint test statistics. The degree of concentration depends on the cumulative π-weights in equation (2). Therefore, the critical values of $J_{\max}(L, h_L)$ are in general smaller than those of the other joint test statistics.

For the component test statistics $C_{\max}(i, h_i^*)$, the critical values should be close to those commonly used in the univariate outlier detection provided that k is not too large, because these statistics are based on individual components. The only difference in the multivariate case is that the maximisation is evaluated across the k components as well as over the time indices. As with the joint test statistics, asymptotic distributions of $C_{\max}(i, h_i^*)$ also depend on the serial correlations of $\{C_{i,h}\}$. In practice, the true model is unknown, the above distributional properties are only approximations, and we use simulation to generate finite sample critical values of the two test statistics.

!! As in the univariate case, if a single joint statistic $J_{\max}(i, h_0)$ is significant at time index h_0, we identify a multivariate outlier of type i at h_0, where $i = I, A, L, T$. In the case of multiple significant joint test statistics, we identify the outlier type based on the test that has the smallest empirical p-value. For example, if $J_{\max}(A, h_A)$ has the smallest p-value at time index h_0 and the p-value is smaller than 0·05, then we identify an additive outlier at time index h_0 at the 5% significance level. When all of the four joint statistics are nonsignificant at a given level, we use the component statistics $C_{\max}(i, h_i^*)$ to check for additional outliers. This step ensures that no component outlier is overlooked. In some cases, the estimated outlier parameter $\hat{\omega}$ may also suggest that the identified outlier only affects some of the components. CRITICAL VALUES

Once an outlier is identified, its impact on the underlying time series is removed, using the results of § 2. The adjusted series is treated as a new dataset and the detecting procedure is iterated. We terminate the detection procedure when no significant outlier is detected. !! Then we recommend a joint estimation of the model parameters and detected outliers. If some outlier parameters are found to be nonsignificant in this joint estimation, they are deleted. The joint estimation is repeated until all the detected outliers are significant at the given level. ??

Some remarks on the proposed procedure are in order. First, as in the univariate case, a major contribution of the procedure is to identify data points that need further attention. If the percentage of outliers exceeds substantially the significance level used in the detection, then the entertained model may be inadequate and should be changed. For instance, a variance change or existence of conditional heteroscedasticity, which is not considered in this paper, may lead to a large number of outliers and one should consider such a model if necessary. Secondly, the proposed procedure detects a single outlier in an iteration to avoid overspecification of the number of outliers. Thirdly, when multiple outliers exist, the proposed procedure may encounter masking or smearing effects of the outliers. The final joint estimation of the proposed procedure is designed to reduce the chance of misidentification and the masking or smearing effects of multiple outliers. Fourthly, it could happen that an outlier affects different components differently in a vector time series, and this possibility is allowed in the proposed procedure. For instance, a strike can appear as an additive outlier on a production series but as a level shift on a sales series if it permanently affects the firm's market share. In this case at the time point of strike we may detect first a multivariate additive outlier that affects primarily the first component. We may also identify a multivariate level shift at the same time point in a subsequent iteration with a significant outlier parameter in the second component. Fifthly, some refinements of the proposed procedure are possible. For example, under the current procedure an identified outlier is assumed to have effects on all components of a time series and the estimated outlier effect $\hat{\omega}$ is used to remove outlier effects. It might be better to adjust only those components which have a significant t ratio in $\hat{\omega}_i$. For simplicity we do not adopt such a procedure.

Finally, instead of considering the four types of outlier separately, it is tempting to use a multivariate regression to detect jointly the existence of an outlier at a given time index h. The existence of an outlier at time index h can then be detected by testing jointly all $\omega_i = 0$. However, such an approach encounters serious multicollinearity, or identifiability, problems. First, as discussed in § 2, the marginal effect of a multivariate outlier may be equivalent to that of another type of outlier. In this case, the corresponding components in ω_i are not identifiable. Secondly, the interaction between outlier size and the model can also result in multicollinearity. Thirdly, there exists insufficient observations to estimate all ω_i when h is close to the end of the data span. Fourthly, the multivariate model structure can also lead to multicollinearity in outlier parameters. For instance, as discussed in Tiao & Tsay (1989), some linear combination of the component series may become white noise for which innovational outlier and additive outlier are equivalent. Such a linear combination leads to multicollinearity in the above multivariate regression.

4. Simulation study

In this section we investigate finite-sample critical values and power of the test statistics in (6) via simulation. We employ two vector AR(1) and AR(6) models to obtain empirical quantiles of the test statistics for $k = 2, 3$ and for sample sizes $n = 100$, 200 and 400. The two vector AR(1) models are in the form $x_t = \Phi x_{t-1} + \varepsilon_t$, with parameters given by

$$\Phi = \begin{bmatrix} 0{\cdot}2 & 0{\cdot}3 \\ -0{\cdot}6 & 1{\cdot}1 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1{\cdot}0 & 0{\cdot}2 \\ 0{\cdot}2 & 1{\cdot}0 \end{bmatrix}, \quad \Phi = \begin{bmatrix} 0{\cdot}2 & 0{\cdot}3 & 0{\cdot}0 \\ -0{\cdot}6 & 1{\cdot}1 & 0{\cdot}0 \\ 0{\cdot}2 & 0{\cdot}3 & 0{\cdot}6 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1{\cdot}0 & 0{\cdot}2 & 0{\cdot}2 \\ 0{\cdot}2 & 1{\cdot}0 & 0{\cdot}2 \\ 0{\cdot}2 & 0{\cdot}2 & 1{\cdot}0 \end{bmatrix}. \tag{7}$$

As a result of the normalisation by the matrix $\Sigma^{-\frac{1}{2}}$, the proposed detection statistics in (6) are scale-invariant. Therefore, the choice of covariance matrix is not critical in the simulation. The two coefficient matrices used in the simulation have eigenvalues (0·5, 0·8) and

(0·5, 0·6, 0·8) respectively. We chose these eigenvalues based on the results of Chang et al. (1988), who show that for univariate AR(1) models outlier effects on parameter estimation are larger when the coefficient is around 0·6.

The two vector AR(6) models are obtained as follows. For the bivariate case, we use the vector AR(6) model fitted to the gas-furnace series of Box, Jenkins & Reinsel (1994, p. 419); see also Example 1 of § 5. For the trivariate case, we add a third component that is independent of the bivariate model and follows the model

$$(1+0{\cdot}78B+0{\cdot}57B^2+0{\cdot}30B^3+0{\cdot}20B^4+0{\cdot}15B^5+0{\cdot}03B^6)x_{3t}=\varepsilon_{3t},$$

with $\sigma_{33}=0{\cdot}05$.

For a given model and sample size n, we generate 10 000 realisations. For each realisation, we estimate a vector autoregressive model with proper order by ordinary least squares, obtain the residuals and $\hat{\Sigma}$, and compute the test statistics in (6) using the estimated parameters. Tables 1 and 2 provide some empirical quantiles for the joint and component test statistics under the null hypothesis of no outlier in the data. From Tables 1 and 2, we make the following observations. First, as expected, empirical quantiles of $J_{\max}(L, h_L)$ are much smaller than those of the other three joint test statistics. This is

Table 1. *Simulation study. Empirical quantiles of the $J_{\max}(i, h_i)$ statistics in (6) based on 10 000 realisations. The models used are given in § 4*

Sample size	Test	Probability 50%	90%	95%	97·5%	99%	50%	90%	95%	97·5%	99%
		Vector AR(1)					Vector AR(6)				
		Bivariate case									
100	$J_{\max}(I, h_I)$	9·74	13·03	14·35	15·60	17·34	9·76	13·08	14·43	15·61	17·23
	$J_{\max}(A, h_A)$	9·70	13·07	14·32	15·57	16·96	9·74	13·63	15·08	16·63	18·59
	$J_{\max}(L, h_L)$	7·61	11·13	12·37	13·50	14·82	6·31	9·33	10·44	11·63	13·30
	$J_{\max}(T, h_T)$	9·58	12·95	14·27	15·43	17·05	9·33	12·92	14·34	15·65	17·58
200	$J_{\max}(I, h_I)$	11·20	14·66	16·01	17·47	19·06	10·80	14·24	15·54	16·82	18·41
	$J_{\max}(A, h_A)$	11·13	14·66	15·95	17·37	19·18	10·70	14·57	16·08	17·70	19·38
	$J_{\max}(L, h_L)$	8·37	12·18	13·49	14·81	16·40	6·20	9·15	10·36	11·53	13·17
	$J_{\max}(T, h_T)$	11·04	14·55	15·87	17·19	18·67	10·12	13·72	15·14	16·47	18·17
400	$J_{\max}(I, h_I)$	12·60	16·19	17·63	19·06	20·81	11·79	15·12	16·47	17·75	19·54
	$J_{\max}(A, h_A)$	12·56	16·21	17·64	18·86	20·81	11·47	15·26	16·71	18·25	20·16
	$J_{\max}(L, h_L)$	9·62	13·48	14·88	16·20	18·05	6·37	9·16	10·25	11·31	13·05
	$J_{\max}(T, h_T)$	12·57	16·13	17·53	18·96	20·83	10·58	14·18	15·53	17·02	18·80
		Trivariate case									
100	$J_{\max}(I, h_I)$	15·55	25·00	29·56	34·23	41·43	14·43	23·75	28·10	32·50	38·82
	$J_{\max}(A, h_A)$	15·50	25·08	20·49	34·00	41·81	14·66	25·10	30·14	34·98	42·23
	$J_{\max}(L, h_L)$	10·64	18·09	21·56	25·46	32·06	8·62	15·94	19·85	24·39	27·75
	$J_{\max}(T, h_T)$	15·48	25·14	30·05	34·81	42·18	14·03	24·14	29·24	36·61	41·12
200	$J_{\max}(I, h_I)$	19·20	28·45	32·10	36·72	42·10	16·44	26·94	31·85	36·77	43·16
	$J_{\max}(A, h_A)$	19·10	28·90	33·04	37·28	43·02	16·40	27·96	33·31	38·90	45·49
	$J_{\max}(L, h_L)$	12·12	19·99	23·24	26·43	31·51	8·71	16·22	20·28	24·13	29·81
	$J_{\max}(T, h_T)$	19·12	28·86	32·85	37·18	43·29	15·56	27·18	32·48	37·68	44·88
400	$J_{\max}(I, h_I)$	22·86	32·24	36·12	40·04	45·09	18·78	30·54	35·69	40·95	47·32
	$J_{\max}(A, h_A)$	22·96	32·66	36·49	40·32	45·77	18·11	30·70	36·11	41·61	48·89
	$J_{\max}(L, h_L)$	14·50	22·89	26·39	30·27	34·56	9·20	17·16	20·90	24·68	29·50
	$J_{\max}(T, h_T)$	23·10	32·79	36·67	40·67	45·47	17·11	30·01	35·08	40·66	48·01

particularly so for the vector AR(6) models, because the models have large cumulative π-weights. Secondly, quantiles of $J_{\max}(i, h_i)$ for $i = I, A, T$ are closer to each other and are less sensitive to the autoregressive order, implying that a common critical value can be used for these three test statistics. Thirdly, empirical quantiles of the component statistics $C_{\max}(i, h_i^*)$ are less affected by the autoregressive order, but seem more variable when the dimension increases. The quantiles of $C_{\max}(L, h_L^*)$ are also smaller than those of the other component statistics for $k = 2$, especially for the vector AR(6) models. Our simulation suggests that, except for level shift, 3·75 may serve as an approximate critical value at the 5% significance level for all sample sizes used in the bivariate study. This critical value is larger than the 3·0 or 3·5 used in univariate outlier detection; see Chen & Liu (1993) and the references therein. For the trivariate case, the variation in the empirical 95th percentiles of $C_{\max}(i, h_i^*)$ is relatively large, indicating that the critical values of $C_{\max}(i, h_i^*)$ depend on the dimension of x_t and should be adjusted accordingly in practice. In summary, the empirical critical values of the joint test statistics, especially $J_{\max}(L, h_L)$, depend on sample size, dimension and the model structure. Those of component test statistics $C_{\max}(i, h_i)$ are more stable. Theory and properties of these test statistics need further investigation. In practice, one can use simulation to obtain finite-sample critical values if necessary.

Table 2. *Simulation study. Empirical quantiles of the statistics $C_{\max}(i, h_i^*)$ in (6) based on 10 000 realisations. The models used are given in § 4*

Sample size	Test	50%	90%	95%	97·5%	99%	50%	90%	95%	97·5%	99%
		Vector AR(1)					Vector AR(6)				
		Bivariate case									
100	$C_{\max}(I, h_I^*)$	2·89	3·39	3·58	3·74	3·96	2·89	3·37	3·54	3·69	3·87
	$C_{\max}(A, h_A^*)$	2·89	3·39	3·57	3·73	3·94	2·99	3·58	3·78	3·98	4·22
	$C_{\max}(L, h_L^*)$	2·61	3·18	3·35	3·52	3·71	2·41	2·93	3·10	3·26	3·50
	$C_{\max}(T, h_T^*)$	2·87	3·37	3·55	3·74	3·95	2·76	3·29	3·47	3·64	3·86
200	$C_{\max}(I, h_I^*)$	3·11	3·60	3·78	3·95	4·15	3·05	3·52	3·69	3·84	4·03
	$C_{\max}(A, h_A^*)$	3·11	3·60	3·78	3·93	4·15	3·16	3·74	3·92	4·12	4·34
	$C_{\max}(L, h_L^*)$	2·74	3·33	3·50	3·68	3·88	2·38	2·89	3·06	3·22	3·46
	$C_{\max}(T, h_T^*)$	3·09	3·58	3·76	3·93	4·11	2·90	3·40	3·58	3·73	3·96
400	$C_{\max}(I, h_I^*)$	3·32	3·80	3·96	4·13	4·35	3·20	3·65	3·81	3·96	4·17
	$C_{\max}(A, h_A^*)$	3·31	3·80	3·97	4·12	4·32	3·30	3·83	4·01	4·21	4·43
	$C_{\max}(L, h_L^*)$	2·94	3·51	3·69	3·86	4·06	2·39	2·84	3·00	3·16	3·34
	$C_{\max}(T, h_T^*)$	3·31	3·78	3·95	4·12	4·34	3·00	3·48	3·66	3·84	4·05
		Trivariate case									
100	$C_{\max}(I, h_I^*)$	3·01	3·48	3·64	3·79	3·96	3·03	3·50	3·67	3·84	4·03
	$C_{\max}(A, h_A^*)$	3·24	3·93	4·18	4·44	4·74	3·35	4·28	4·66	5·04	5·52
	$C_{\max}(L, h_L^*)$	2·77	3·54	3·83	4·11	4·52	2·66	3·61	4·05	4·44	4·97
	$C_{\max}(T, h_T^*)$	3·23	3·94	4·18	4·45	4·78	3·18	4·15	4·54	4·91	5·41
200	$C_{\max}(I, h_I^*)$	3·22	3·69	3·85	4·02	4·20	3·19	3·66	3·82	3·98	4·15
	$C_{\max}(A, h_A^*)$	3·56	4·24	4·50	4·72	4·97	3·56	4·45	4·81	5·20	5·68
	$C_{\max}(L, h_L^*)$	2·90	3·67	3·93	4·18	4·49	2·64	3·63	4·05	4·40	4·87
	$C_{\max}(T, h_T^*)$	3·55	4·28	4·52	4·77	5·06	3·36	4·35	4·78	5·16	5·63
400	$C_{\max}(I, h_I^*)$	3·43	3·90	4·07	4·21	4·38	3·35	3·81	3·97	4·14	4·32
	$C_{\max}(A, h_A^*)$	3·86	4·56	4·80	5·03	5·34	3·76	4·72	5·07	5·40	5·81
	$C_{\max}(L, h_L^*)$	3·08	3·84	4·13	4·39	4·67	2·63	3·62	3·98	4·35	4·79
	$C_{\max}(T, h_T^*)$	3·87	4·64	4·89	5·12	5·41	3·57	4·60	4·97	5·32	5·75

Next, we use simulation to study the power of the proposed joint test statistics. The model used in the power study is a bivariate AR(1) model with Φ given in (7). However, the innovational covariance matrix is modified so that the variance of the individual innovation is unity and the correlation between innovations is $-0{\cdot}2$. The sample size used is 200. For each realisation, a single outlier is introduced at the time index $t = 100$ with outlier parameter $\omega = (3{\cdot}5, 3{\cdot}5)'$. For each type of outlier, we use the empirical 5% critical value from Table 1 and tabulate the number of realisations in which the corresponding test statistic exceeds the critical value. The powers based on 10 000 realisations are 89·1%, 96·9%, 100% and 92·1%, respectively, for multivariate innovational outlier, additive outlier, level shift and temporary change.

5. Applications

Example 1. The first example is the well-known gas furnace series of Box et al. (1994, p. 548). Denote the input gas rate in cubic feet per minute by X_t and the percentage of CO_2 in outlet gas by Y_t, both measured in 9-second time intervals. This series is commonly used in the literature as an example of transfer function models. There are 296 observations. For comparison, we also employ the univariate and transfer function models of Box et al. (1994, Ch. 11). Using the joint estimation and detection procedure of Chen & Liu (1993) and a critical value of 3·5, we obtain the models

$$(1 - 2{\cdot}273B + 1{\cdot}923B^2 - 0{\cdot}618B^3)X_t = -0{\cdot}002 + a_{1t}, \quad \hat{\sigma}_1 = 0{\cdot}129 \tag{8}$$

$$Y_t = 53{\cdot}08 + \frac{-0{\cdot}636B^3 - 0{\cdot}264B^4 - 0{\cdot}439B^5}{1 - 0{\cdot}570B} X_t + \frac{1}{1 - 1{\cdot}511B + 0{\cdot}579B^2} a_{2t}, \quad \hat{\sigma}_2 = 0{\cdot}195, \tag{9}$$

where $\hat{\sigma}_1$ and $\hat{\sigma}_2$ are the residual standard error of the input and output series respectively, after outlier adjustment. The detected outliers are given in Table 3. There are 7 and 6 outliers for models (8) and (9) respectively. If the critical value for outlier detection is set to 3·0, then there are 17 and 10 outliers, respectively, for the two models. A critical value of 3·5 corresponds approximately to an asymptotic 2·5% significance level. Note that the two temporary-change outliers at times 113 and 117 in the input series, which show opposite effects, may suggest a patch of outliers in the period 113–116. Similarly, there may be a patch of outliers from 265 to 269 in the output series.

Table 3. *Example* 1. *Outliers detected for the gas-furnace series using a univariate method with critical value* 3·5, *which is approximately at the* 2·5% *level*

Input series: gas rate				Transfer function for CO_2			
Time	Size	t ratio	Type	Time	Size	t ratio	Type
43	0·770	12·20	TC	199	0·915	6·08	LS
55	−0·718	−11·38	TC	236	−0·863	−4·42	IO
91	0·286	4·53	TC	265	1·481	7·59	IO
113	−0·479	−7·59	TC	266	0·729	3·74	IO
117	0·248	3·92	TC	267	0·454	4·23	AO
198	−0·534	−4·15	IO	269	−1·296	−6·33	IO
262	0·607	4·72	IO				

AO, additive outlier; IO, innovational outlier; LS, level shift; TC, temporary change.

We now turn to multivariate modelling. Using the chi-squared statistic of Tiao & Box (1981) and the Akaike information criterion, we adopt a bivariate AR(6) model for the series. The first component is the gas rate and the second component is the output CO_2 concentration. Applying the proposed detection procedure and, for comparison purposes, using 2·5% critical values obtained by interpolation from Tables 1 and 2, we summarise the detection results in Table 4. Twelve outliers are detected by the procedure. Once an outlier was detected, we removed its effects on the data and re-estimated the bivariate AR(6) model. The estimated outlier parameters $\hat{\omega} = (\hat{\omega}_1, \hat{\omega}_2)'$ of the 12 outliers are given in Table 5 along with t ratios of the estimates. Note that the detected multivariate temporary changes at $t = 43$ and 55 introduce large $J_{\max}(A, h_A)$ statistics at $t = 42$ and 54. This is understandable because, for a vector AR(6) model, the test statistic $J_{A,h}$ involves filtered values $\hat{a}_t$ for $t = h, h+1, \ldots, h+6$.

Table 4. *Example* 1. *Results of multivariate outlier detection for the gas-furnace series using a bivariate* AR(6) *model and* 2·5% *critical values. The number in parentheses for joint tests is the corresponding time index whereas those for component tests are time index and component index*

J

C

(*a*) *Joint test statistics*

J

Iterations	$J_{\max}(I, h_I)$	$J_{\max}(A, h_A)$	$J_{\max}(L, h_L)$	$J_{\max}(T, h_T)$	Outlier Time	Outlier Type
1	39·23 (265)	35·70 (42)	27·84 (199)	41·05 (43)	43	MTC
2	38·54 (265)	43·90 (54)	26·22 (199)	46·15 (55)	55	MTC
3	39·29 (265)	27·15 (264)	24·46 (199)	28·09 (264)	265	MIO
4	16·94 (199)	26·27 (113)	24·29 (199)	26·70 (113)	199	MLS
5	16·01 (269)	25·85 (113)	16·56 (113)	26·24 (113)	113	MTC
6	16·34 (262)	16·71 (235)	14·49 (288)	14·44 (261)	288	MLS
7	16·29 (262)	17·56 (235)	13·91 (287)	14·55 (91)	287	MLS
8	16·51 (236)	19·70 (235)	14·78 (236)	16·52 (235)	236	MLS
9	15·81 (262)	14·25 (197)	10·61 (82)	15·19 (91)	—	—
Crit.	17·29	17·98	11·42	16·73		

always compare the statistics with the critical values!

da dove li abbiam presi?

-> since the test now can't find any outlier, let's look at C

(*b*) *Component test statistics*

C

Iterations	$C_{\max}(I, h_I^*)$	$C_{\max}(A, h_A^*)$	$C_{\max}(L, h_L^*)$	$C_{\max}(T, h_T^*)$	Outlier Time	Outlier Type
9	3·95 (262, 1)	3·76 (197, 1)	−3·23 (82, 1)	3·85 (91, 1)	82	MLS
10	4·34 (262, 1)	4·09 (91, 1)	3·29 (262, 1)	4·24 (91, 1)	262	MIO
11	3·69 (198, 1)	3·96 (91, 1)	<3·19	4·10 (91, 1)	91	MTC
12	3·80 (198, 1)	4·07 (197, 1)	<3·19	4·11 (197, 1)	197	MTC
13	3·26 (266, 2)	−3·25 (116, 1)	<3·19	3·45 (117, 1)	—	—
Crit.	3·90	4·17	3·19	3·79		

stopping value!

MIO, multivariate innovational outlier; MLS, multivariate level shift; MTC, multivariate temporary change; Crit., critical value.

As shown by Tiao & Box (1981), the marginal models of the bivariate AR(6) model employed are close to those of equations (8) and (9). Therefore, the comparison of outlier detection between univariate and multivariate models can be made fairly, and we obtain some interesting results.

First, the multivariate method detects fewer outliers than the univariate methods. In addition, the outliers detected by the multivariate method are not a subset of those detected by univariate methods. For example, the level shifts at $t = 287$ and 288 are not detected by the univariate methods. This demonstrates that multivariate joint detection could be more powerful than univariate methods.

Secondly, the multivariate innovational outlier detected at $t = 265$ clearly highlights the discussion of § 2·3. Specifically, we observe the following: the t ratios of estimated outlier parameters in Table 5 show that the outlier occurred simultaneously to both components; as expected, this multivariate innovational outlier introduces a patch of outliers in the marginal model of the output series at time indices 266–269; the estimated outlier effect is negative in the input series and the transfer function model shows a negative relationship between the input and output series with a delay of 3 time periods, with the consequence that the outlier effects at time indices 266 and 267 of the output series are positive; as expected, the patch of outliers disappears under the multivariate framework.

Table 5. *Example* 1. *Estimates of outlier parameters for the gas-furnace series using a multivariate model*

Time	Type	$\hat{\omega}_1$	(t ratio)	$\hat{\omega}_2$	(t ratio)	Time	Type	$\hat{\omega}_1$	(t ratio)	$\hat{\omega}_2$	(t ratio)
43	MTC	0·683	(6·41)	−0·019	(−0·11)	55	MTC	−0·613	(−6·79)	0·049	(0·27)
265	MIO	−0·362	(−3·40)	1·396	(5·86)	199	MLS	−0·098	(−1·51)	0·866	(4·93)
113	MTC	−0·376	(−5·12)	−0·067	(−0·42)	288	MLS	0·154	(2·04)	0·587	(3·23)
287	MLS	0·130	(1·74)	0·578	(3·28)	236	MLS	0·069	(1·17)	−0·595	(−3·83)
82	MLS	−0·166	(−3·23)			262	MIO	0·565	(4·34)		
91	MTC	0·249	(4·10)			197	MTC	0·239	(4·11)		

MIO, multivariate innovational outlier: MLS, multivariate level shift; MTC, multivariate temporary change.

Thirdly, highly significant outliers detected by the univariate methods are also detected by the multivariate method; see the outliers at $t = 43$, 55, 113 and 265. Fourthly, the classification of outliers is rather consistent between univariate and multivariate methods. Fifthly, some minor time differences may occur between univariate and multivariate methods. For example, the univariate innovational outlier at $t = 198$ is shown as a temporary change at $t = 197$ in the multivariate case. Finally, t ratios in Table 5 show that no outlier occurred in the output series at $t = 43$ and 55. The significant temporary changes in the input series at these two time points suggest that the outlier effects are carried over from the input series to the output series. Consequently, the residuals of the output series do not contain additional information about these two outlying observations. These two multivariate temporary changes are referred to as passed outliers in Pankratz (1993).

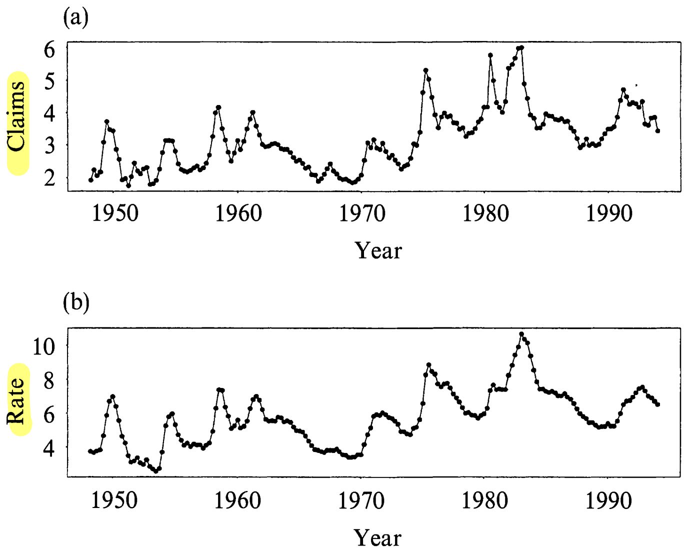
Young (1974) and Subba Rao & Tong (1974) analysed this dataset and suggested that the system is time-dependent. We did not consider time-dependent models in this paper.

Example 2. In this example, we consider the U.S. quarterly seasonally adjusted initial jobless claims and unemployment rate from 1948 to 1993. The initial jobless claims were divided by 100. Figure 1 shows the time plots of the data, which consist of 184 observations. Using the same models as in Montgomery et al. (1998) and the joint estimation-detection procedure of Chen & Liu (1993), we obtain the univariate models

$$(1 - 0{\cdot}30B)(1 - 0{\cdot}36B^4)(1 - B)y_{1t} = (1 - 0{\cdot}75B^4)a_{1t}, \quad \hat{\sigma}_1 = 0{\cdot}222,$$

$$(1 - 0{\cdot}66B)(1 - 0{\cdot}27B^4)(1 - B)y_{2t} = (1 - 0{\cdot}81B^4)a_{2t}, \quad \hat{\sigma}_2 = 0{\cdot}271,$$

where y_{1t} and y_{2t} are the initial claims and unemployment rate, respectively. The seasonal parameters in both models are highly significant, even though the data were seasonally adjusted. The detected outliers of the two models are given in Table 6. There are 4 and 2 outliers for y_{1t} and y_{2t}, respectively. → using univariate approach



→ again univariate time modelling.

Fig. 1. Example 2. Time plots of (a) U.S. quarterly initial jobless claims, divided by 100, and (b) unemployment rate for 1948–1993. The data were seasonally adjusted.

Table 6. *Example* 2. *Outliers detected for initial jobless claims and unemployment rates using univariate method with critical value* 3·5

Initial jobless claims				Unemployment rates			
Time	Size	t ratio	Type	Time	Size	t ratio	Type
108	1·137	5·13	IO	109	1·097	4·04	IO
130	1·493	9·08	TC	140	0·417	4·30	AO
136	0·999	5·43	LS				
141	−0·971	−5·25	LS				

AO, additive outlier; IO, innovational outlier; LS, level shift; TC, temporary change.

these 3 also identified by multi approach (see later)

→ probably induced by the IO in 108. Because no outlier in 109 using multi approach

Turning to multivariate detection, we employ a bivariate ARIMA model of the form

$$(I - \Phi_1 B - \Phi_2 B^2)(I - \Phi_4 B^4) y_t = c + (I - \Theta_4 B^4)\varepsilon_t. \quad (10)$$

The detection results are summarised in Table 7 based on the 5% empirical critical values from Table 1 for a vector AR(1) model and a sample size of 200. Only three outliers are detected. The estimated outlier parameters and their t ratios, in parentheses, are

130 108 136

[1·249 (6·16), 0·334 (1·84)], [1·080 (4·35), 0·563 (2·22)], [0·968 (4·14), 0·653 (2·62)]

for the outliers at $t = 130$, 108, 136, respectively. An examination of residual cross-correlation matrices indicates some minor significant correlations at lag 8, but these serial correlations disappear when the moving-average part is modified to $(I - \Theta_4 B^4 - \Theta_8 B^8)\varepsilon_t$. The parameter estimates of model (10) before and after outlier

→ · for 130 only the second component is significant
· for 108 & 136 both the components are significant: see later for further insights on this!

adjustment are shown in Table 8. The three detected outliers have marked effects on the seasonal parameters and the residual covariance matrix.

Table 7. *Example* 2. *Results of multivariate outlier detection for the initial jobless claim and unemployment using a bivariate seasonal* ARIMA *model and* 5% *critical values. The number in parentheses for joint tests is the corresponding time index whereas those for component tests are time index and component index*

(*a*) *Joint test statistics*

Iterations	$J_{\max}(I, h_I)$	$J_{\max}(A, h_A)$	$J_{\max}(L, h_L)$	$J_{\max}(T, h_T)$	Outlier Time	Outlier Type
1	31·66 (130)	36·35 (130)	<13·50	40·33 (130)	130	MTC
2	20·08 (108)	<13·50	<13·50	16·65 (141)	108	MIO
3	17·25 (136)	<13·50	<13·50	<13·50	136	MIO
4	15·53 (25)	<13·50	<13·50	14·05 (141)	—	—
Crit.	16·01	15·95	13·49	15·87		

(*b*) *Component test statistics*

Iterations	$C_{\max}(I, h_I^*)$	$C_{\max}(A, h_A^*)$	$C_{\max}(L, h_L^*)$	$C_{\max}(T, h_T^*)$	Outlier Time	Outlier Type
4	3·59 (25, 2)	<3·50	<3·50	3·61 (141, 1)	—	—
Crit.	3·78	3·78	3·50	3·76		

MIO, multivariate innovational outlier; MTC, multivariate temporary change; Crit., critical value.

Table 8. *Example* 2. *Parameter estimates of model* (10) *before and after multivariate outlier detection. The values in parentheses are standard errors*

c'	Φ_1		Φ_2		Φ_4		Θ_4		Σ	
				Before outlier adjustment						
0·139	1·31	−0·17	−0·25	0·11	0·13	−0·19	0·02	0·17	0·081	0·056
(0·089)	(0·10)	(0·09)	(0·11)	(0·07)	(0·18)	(0·08)	(0·22)	(0·15)		
0·054	0·59	1·16	−0·31	−0·36	0·06	−0·08	−0·07	−0·09	0·056	0·073
(0·086)	(0·09)	(0·09)	(0·11)	(0·09)	(0·13)	(0·21)	(0·17)	(0·23)		
				After outlier adjustment						
0·331	1·36	−0·22	−0·32	0·15	−0·21	−0·12	−0·55	0·18	0·049	0·036
(0·13)	(0·09)	(0·07)	(0·11)	(0·06)	(0·16)	(0·07)	(0·17)	(0·12)		
0·151	0·65	1·12	−0·34	−0·33	−0·06	−0·28	−0·22	−0·13	0·036	0·060
(0·13)	(0·10)	(0·08)	(0·12)	(0·08)	(0·15)	(0·17)	(0·18)	(0·20)		

Note that the multivariate model identifies only three outliers whereas the univariate models detect six outliers. In this example, the multivariate outliers form a subset of those identified by univariate methods. Secondly, the innovational outlier in the unemployment rate at time $t=109$ is caused by the innovational outlier in the initial jobless claims at $t=108$, because there exists no outlier at $t=109$ in the multivariate case. This example clearly demonstrates that an outlier in a component series may be induced by that of another component, and that detecting outliers separately for each individual component using a marginal model may result in overspecification of the number of outliers. Thirdly, the significance of the outlier parameters in both components at $t=108$ and 136 indicates that some external disturbances occurred in the U.S. economy at these two time points that affected both the initial jobless claims and unemployment rate. In other words, the impact of these disturbances on the unemployment rate cannot be fully accounted for by that on the initial claims. These two time points were the fourth quarters of 1974 and 1981, respectively, in which the U.S. economy was in recession as classified by the National Bureau of Economic Research. Thus, the significance of the positive estimates $\hat{\omega}$ at these two periods shows that the economic slow-downs in 1974 and 1981 caused both the initial jobless claims and unemployment rate to rise. In addition, the estimated effect on unemployment rate, $\hat{\omega}_2$, represents the additional effect of economic slow-down on unemployment rate beyond that induced by the impact on initial jobless claims. Such information is not evident if one only uses univariate outlier detection. Finally, the significance of both components of $\hat{\omega}$ at $t = 108$ and 136 also indicates a possible joint structural break of the series at these two points.

ACKNOWLEDGEMENT

The authors thank the editor and three referees for their careful and constructive comments that led to substantial improvements of the paper. The research of R. S. Tsay is supported in part by the National Science Foundation, the Chiang Ching-Kou Foundation and the Graduate School of Business, University of Chicago, and that of D. Peña by the Direccion General de Investigacion Cientifica y Tecnica of Spain.

REFERENCES

BOX, G. E. P., JENKINS, G. M. & REINSEL, G. (1994). *Time Series Analysis: Forecasting and Control*, 3rd ed. Englewood Cliffs, NJ: Prentice-Hall.

BOX, G. E. P. & TIAO, G. C. (1975). Intervention analysis with applications to environmental and economic problems. *J. Am. Statist. Assoc.* **70**, 70–9.

CHANG, I., TIAO, G. C. & CHEN, C. (1988). Estimation of time series parameters in the presence of outliers. *Technometrics* **3**, 193–204.

CHEN, C. & LIU, L. (1993). Joint estimation of model parameters and outlier effects in time series. *J. Am. Statist. Assoc.* **88**, 284–97.

FINDLEY, D. F., MONSELL, B. C., BELL, W. R., OTTO, M. C. & CHEN, B. C. (1998). New capabilities and methods of the X-12-ARIMA seasonal-adjustment program (with Discussion). *J. Bus. Econ. Statist.* **16**, 127–77.

FOX, A. J. (1972). Outliers in time series. *J. R. Statist. Soc.* B **34**, 350–63.

LI, H. & TSAY, R. S. (1998). A unified approach to identifying multivariate time series models. *J. Am. Statist. Assoc.* **93**, 770–82.

MCCULLOCH, R. E. & TSAY, R. S. (1994). Bayesian analysis of autoregressive time series via the Gibbs sampler. *J. Time Ser. Anal.* **15**, 235–50.

MONTGOMERY, A., ZARNOWITZ, V., TSAY, R. S. & TIAO, G. C. (1998). Forecasting the U.S. unemployment rate. *J. Am. Statist. Assoc.* **93**, 478–93.

PANKRATZ, A. (1993). Detecting and treating outliers in dynamic regression models. *Biometrika* **80**, 847–54.

SUBBA RAO, T. & TONG, H. (1974). Linear time-dependent systems. *IEEE Trans. Auto. Contr.* **19**, 735–7.

TIAO, G. C. & BOX, G. E. P. (1981). Modeling multiple time series with applications. *J. Am. Statist. Assoc.* **76**, 802–16.

TIAO, G. C. & TSAY, R. S. (1975). Model specification in multivariate time series (with Discussion). *J. R. Statist. Soc.* B **51**, 157–213.

TSAY, R. S. (1988). Outliers, level shifts, and variance change in time series. *J. Forecasting* **7**, 1–20.

YOUNG, P. C. (1974). Recursive approaches to time-series analysis. *Bull. Inst. Math. Appl.* **10**, 200–24.

[*Received November* 1998. *Revised February* 2000]