MARCEL DEKKER, INC. • 270 MADISON AVENUE • NEW YORK, NY 10016

ECONOMETRIC REVIEWS
Vol. 21, No. 3, pp. 353–381, 2002

ANALYSIS OF VECTOR AUTOREGRESSIONS IN THE PRESENCE OF SHIFTS IN MEAN

Serena Ng[1] and Timothy J. Vogelsang[2]

[1]Department of Economics, Johns Hopkins University,
Baltimore, Maryland
E-mail: serena.ng@jhu.edu
[2]Department of Economics, Cornell University, Ithaca, New York
E-mail: tjv2@cornell.edu

ABSTRACT

This paper considers the implications of mean shifts in a multivariate setting. It is shown that under the additive outlier type mean shift specification, the intercept in each equation of the vector autoregression (VAR) will be subject to multiple shifts when the break dates of the mean shifts to the univariate series do not coincide. Conversely, under the innovative outlier type mean shift specification, both the univariate and the multivariate time series are subject to multiple shifts when mean shifts to the innovation processes occur at different dates. We consider two procedures, the first removes the shifts series by series before forming the VAR, and the second removes intercept shifts in the VAR directly. The pros and cons of both methods are discussed.

Key Words: Trend break; Structural change; Causality tests; Forecasting

1 INTRODUCTION

In recent years, many univariate statistics have been developed to test for the presence of structural breaks in stationary and nonstationary time series. When applied to macroeconomic data, the evidence suggests that breaks in the form of a shift in mean and/or the trend function have occurred in many series. For example,

DOI: 10.1081/ETC-120015788 0747-4938 (Print); 1532-4168 (Online)

www.dekker.com

Perron[1] analyzed the Nelson-Plosser data set and found that many series are stationary around segmented means and/or trends. Structural change in many of these series was confirmed by Vogelsang[2] and Chu and White[3] using direct tests for shifts in trend. Vogelsang[4] found evidence of a mean shift in the unemployment rate. Series for international output were also found to have segmented trends by Banerjee et al.,[5] Ben-David and Papell[6] and Perron.[7] In spite of these findings, many vector autoregressions (VARs) continue to be estimated as though there were no breaks in the series. If there are breaks in the univariate series, it seems natural that the breaks should also appear in a multivariate system. While it is fairly well known how an unstable mean affects univariate analysis, much less is known about how unstable means affect multivariate analysis.

This paper considers additive outlier (AO) and innovational outlier (IO) mean shifts in a vector time series. This terminology is borrowed from the outlier and intervention analysis of Box and Tiao.[8] We explore the ramifications of unstable means for estimation, inference and forecasting using VARs. We show that some parameter estimates of the VAR can remain consistent even when there are omitted mean shifts, but only under very restrictive assumptions about the causal structure of the vector time series. In general, inference based on the estimates of the VAR is invalid when mean shifts are omitted. Our analysis thus provides a theoretical explanation for the simulations reported in Lutkepohl[9] and Bianchi[10] which show that Granger causality (GC) tests over-reject in finite samples when mean shifts are ignored. We show that the extent of over-rejection depends on whether the mean shift is of an AO or an IO type.

We also consider the case when there is one mean shift in each series but they occur at different break dates. We show that AO type mean shifts will induce multiple intercept shifts in each equation of the VAR but only a single shift in the univariate representation of the series. In contrast, IO type mean shifts will induce multiple mean shifts in both the univariate representation of the series and the reduced form VAR. We then consider strategies for removing the breaks when it is not known *a priori* whether the mean shifts are AO or IO. The two choices are (i) remove the breaks before estimating the dynamic parameters, and (ii) estimate the mean shifts simultaneously with the autoregressive parameters. We show that if the break dates are the same, strategy (i) is inefficient when the data are of the IO type, while strategy (ii) is appropriate but inefficient for AO data. When the break dates differ, some demeaning procedures are invalid. We discuss the relative merits of those that are robust even when the break dates are unknown.

2 THE MODEL

There are two common approaches to modeling mean shifts in a univariate time series which are based on models with outliers.[11,12] The first is the

AO approach[1] which specifies a series, y_t, as the sum of a deterministic component $\mu + \delta DU_t$ and a stochastic component, $z_t = az_{t-1} + e_t$, where $DU_t = 1(t > T_B)$ and $1(\cdot)$ is the indicator function, T_B is the break date and e_t is a white noise process. The second is the IO approach which models the break as occurring to the mean of the innovation series. That is, $y_t = \mu + v_t$, $v_t = av_{t-1} + e_t + \delta DU_t$. Thus, y_t can equivalently be represented as $y_t = \mu(1-a) + \delta DU_t + ay_{t-1} + e_t$. The IO model has the advantage of permitting mean shifts to occur gradually over time, but has the disadvantage that the unconditional mean of the series depends on the dynamics of the noise component.

The AO and IO approaches to modeling mean shifts in univariate processes provide a natural starting point to modeling mean shifts in multivariate time series. Let $y_t = (y_{1t}, \ldots, y_{nt})'$ be a $n \times 1$ vector for $t = 1, \ldots, T$. The jth time series is denoted $Y_j = (y_{j1}, \ldots, y_{jT})'$. An AO mean shift model can be specified as follow:

$$y_t = \mu + \delta DU_t + z_t \tag{1}$$

$$\Gamma z_t = Az_{t-1} + e_t \tag{2}$$

where y_0 is non-random, e_t is an $n \times 1$ vector of white noise sequences with finite fourth moments and $E(e_t e_t')$ a diagonal matrix, μ is a vector of constants, $DU_t = (dU_{1t}, \ldots, dU_{nt})'$ is an $n \times 1$ vector of break dummies with $dU_{it} = 1(t > T_{B_i})$ where T_{B_i} is the break date for the series i, and δ is a $n \times n$ diagonal matrix with $\text{diag}(\delta) = (\delta_1, \ldots, \delta_n)'$. In the special case when $T_{B_i} = T_B \; \forall i$, then $DU_t = \imath dU_t$, where $\imath$ is the unit vector. The contemporaneous correlation in z_t is given by the $n \times n$ matrix Γ. The matrix A is $n \times n$ with elements a_{ij}. We focus on a lag one or first order VAR to minimize complicated notation. The generalization to higher order lags is straight forward. The roots of $|I_n - BL|$ are assumed to lie outside the unit circle where $B = \Gamma^{-1}A$. We denote the (i, j)th element of B by b_{ij}, and the ith row of B by B_i'. For future reference, we also define D_t from $DU_t = DU_{t-1} + D_t$. Thus, the ith element of D_t is defined as $d_{it} = 1 \; (t = T_{B_i} + 1)$ for $i = 1, \ldots, n$.

By transforming the above model in standard fashion, the VAR representation of (1) and (2) is

$$y_t = \mu^* + \delta^A DU_t + \gamma^A d_t + By_{t-1} + u_t \tag{3}$$

where $\mu^* = I_n - B$, $\delta^A = (I_n - B)\delta$, $\gamma^A = B\delta$, $u_t = \Gamma^{-1}e_t$. The corresponding vector moving-average (VMA) representation is

$$y_t = \mu + \delta DU_t + \sum_{i=0}^{\infty} B^i u_{t-i}$$

An IO mean shift model can be specified as:

$$y_t = \mu + v_t \tag{4}$$

$$\Gamma v_t = Av_{t-1} + e_t + \delta DU_t \tag{5}$$

[1]In the outlier literature what we are labelling an AO mean shift would be called a level shift.

To contrast the difference with the AO model, it is useful to rewrite the IO model as:

$$y_t = \mu + (\Gamma - AL)^{-1}\delta DU_t + z_t$$
$$z_t = \Gamma z_{t-1} + u_t$$

The reduced form VAR corresponding to the IO model is:

$$y_t = \mu^* + \delta^I DU_t + By_{t-1} + u_t \tag{6}$$

where $\mu^* = I_n - B$, and $\delta^I = \Gamma^{-1}\delta$. The VMA representation of the IO model is

$$y_t = \mu + \sum_{i=0}^{\infty} B^i(\delta^I DU_{t-i} + u_{t-i})$$

The VMA representations reveal three important differences between the AO and the IO models. First, the unconditional means for Y_i in the AO case are invariant to mean shifts in Y_j for $i \neq j$. This is not the case with IO data. Second, the impact of the IO mean shifts changes over time. In contrast, the short and long run impact of AO mean shifts are the same. Third, the effect of AO mean shifts falls relative to the unconditional variance of y_t as B moves closer to the unit circle. In the IO case, the effect of mean shifts and the unconditional variance of the series both increase with B. Thus the size of the AO mean shifts is relatively less sensitive to B.

2.1 Omitted Mean Shifts

To understand the effects of omitted mean shifts for estimation of the VAR, consider without loss of generality estimation of the ith equation:

$$y_{it} = \hat{\mu}_i^* + \hat{B}_i' y_{t-1} + \hat{u}_{it}^* \tag{7}$$

when the data are generated by:

$$y_{it} = \mu_i^* + B_i' y_{t-1} + u_{it}^*$$

In the AO case we have $u_{it}^* = u_{it} + \delta_i^{A\prime} DU_t + \gamma_i^{A\prime} d_t$. In the IO case we have $u_{it}^* = u_{it} + \delta_i^{I\prime} DU_t$. Therefore, unless δ_i^A and δ_i^I are zero vectors, regressions based upon (7) that do not take into account mean shifts in the data are misspecified. The implications for the least squares estimates are summarized in the following two theorems.

Theorem 1
Let $y_t = (y_{1t}, \ldots, y_{nt})'$, $t = 1, \ldots, T$ be generated by the AO model (1) and (2). Let the parameters $B_i = (b_{i1}, \ldots, b_{in})'$ be estimated from (7) by OLS to yield $\hat{B}_i = (\hat{b}_{i1}, \ldots, \hat{b}_{in})'$. Let $\lambda_i = T_{B_i}/T$, and for each i, λ_i is constant as T increases. Let Λ and $\ominus$ be $n \times n$ matrices with $\Lambda_{ij} = 1 - \min(\lambda_i, \lambda_j) - (1 - \lambda_i)(1 - \lambda_j)$, $\ominus = E(z_t z_t')$, and define $C = \delta\Lambda\delta' + \ominus$. Then as $T \to \infty$, $p\lim(\hat{B}_i - B_i) = C^{-1}\delta\Lambda\delta_i^A$.

Theorem 1 shows that the least squares estimates of B_i are inconsistent when mean shifts are omitted except when δ_i^A is a zero vector. In general, a mean shift in any one series will induce bias in all the estimates of the VAR. The exceptional case is the following:

Corollary 2

Under the assumptions of Theorem 1, $p\lim \hat{B}_i = B_i$ only if the following two conditions hold: (a) either $\delta_i = 0$ or $b_{ii} = 1$, and (b) either $\delta_j = 0$ or $b_{ij} = 0$ for $j \neq i$. If, in addition, $DU_{it} = DU_t\ \forall i$, then $p\lim \hat{B}_i = B_i$ if $\iota'\delta_i{}^A = 0$ or $C^{-1}\delta\iota = 0_{n\times 1}$.

One possibility for $\delta_i{}^A$ to be zero is $\delta_i = 0$ and $b_{ij} = 0\ \forall j \neq i$. That is, Y_i is not Granger caused by any other series in the system and Y_i does not experience a mean shift. If the break dates coincide, $\Lambda = \lambda(1-\lambda)\iota\iota'$, $\lambda = T_B/T$, T_B being the common break date. Then $p\lim \hat{B}_i = B_i$ if $\iota'\delta_i{}^A = 0$. This occurs when the mean shifts and dynamics are such that the intercept to the Y_i equation of the VAR is not affected by mean shifts.[2] These, however, are very special cases.

The results for the IO model are given in the next Theorem.

Theorem 3

Let $y_t = (y_{1t}, \ldots, y_{nt})'$, $t = 1, \ldots, T$ be generated by the IO model (4) and (5). Let the parameters $B_i = (b_{i1}, \ldots, b_{in})'$ be estimated from (7) by OLS to yield $\hat{B}_i = (\hat{b}_{i1}, \ldots, \hat{b}_{in})'$. Let $\lambda_i = T_{B_i}/T$ remain constant as T increases. Let $g(i)$ be the $n \times n$ matrix associated with the ith lag of $g(L) = (I - BL)^{-1}$. Define

$$Q = \sum_{i=0}^{\infty} g(i)\delta\Lambda\delta' g(i)' + \ominus$$

where Λ and $\ominus$ are defined as in Theorem 1. Then as $T \to \infty$, $p\lim(\hat{B}_i - B_i) = Q^{-1}g(1)\delta\Lambda\delta_i{}^I$.

Clearly, the parameters estimated from the ith equation will generally not be consistent except when $\delta_i{}^I$ is a zero vector. The exceptional case is given by:

Corollary 4

Under the assumptions of Theorem 2, $p\lim \hat{B}_i = B_i$ if for every $j = 1, \ldots, n$, either $\delta_j = 0$ or $[\Gamma^{-1}]_{ij} = 0$. If, in addition, $DU_{it} = DU_t\ \forall i$, then $p\lim \hat{B}_i = B_i$ if $\iota'\delta^I = 0$ or $g(1)\delta\iota = 0_{n\times 1}$.

In the IO model, regression (7) is always misspecified as long as $\delta_i \neq 0$. However, if $g(1)\delta\iota$ is a zero vector, the unconditional means of the series are eventually unaffected by the IO mean shifts. The transitory nature of the mean shifts allows the OLS estimates to remain consistent.

An important use of bivariate autoregressions is testing for Granger Causality. The theoretical results indicate the circumstances in which omitted mean shifts will adversely affect GC tests. Consider estimation of a first order VAR:

$$y_{2t} = \hat{\mu}_2^* + \hat{b}_{21}y_{1t-1} + \hat{b}_{22}y_{2t-1} + \hat{u}_{2t}^*$$

[2] The mean shifts have the property of a co-feature in the sense of Engle and Kozicki.[12]

The GC test is the square of the Wald test for the null of $b_{21} = 0$.

Theorem 5

Suppose the data are generated under the AO model (1) and (2). Then under the null hypothesis of no GC ($b_{21} = 0$) the following holds as $T \to \infty$.

1. If $\delta_2 = 0$ then GC $\Rightarrow \chi_1{}^2$;
2. If $\delta_2 \neq 0$ then $T^{-1}\mathrm{GC} = O_p(1) > 0$, unless $\delta_1 = 0$, $b_{12} = 0$ and $\theta_{12} = 0$ in which case GC $= O_p(1)$.

From Corollary 2, a sufficient condition for consistency of $\hat{b}_{21}$ is $\delta_2 = 0$. Consistent estimates of b_{21} (but not b_{22}) can be obtained even if $\delta_2 \neq 0$, but this would require Y_1 and Y_2 to be asymptotically orthogonal. Even in that case, size distortions could still arise when the variance of $\hat{b}_{21}$ is biased. In general, we should expect the GC statistic to reject non-causality even when Y_1 does not Granger cause Y_2 when mean shifts are omitted. Similar results hold for the IO model.

2.2 Simulations for Bivariate Vector Autoregression

In this subsection we illustrate the practical implications of omitted means shifts for GC tests in a bivariate VAR. Without loss of generality, we focus on the null hypothesis that Y_1 does not Granger cause Y_2, i.e., the maintain null hypothesis is $b_{21} = 0$. We generated series according to (1) and (2) using $T = 200$ and $T_B = 100$. Four possible combinations of (δ_1, δ_2) were considered: (0,0), (1,0), (0,1), (1,1). This allows for the possibility that a break occurs in none, one, or both series. We assumed throughout that $\Gamma = I$ so that $B = A$. The errors, e_t, are i.i.d draws from a standard bivariate normal distribution using the rndn( ) function in Gauss with seed = 999. Since the variances of the e_{it}s are unity, the magnitude of the mean shift is measured in terms of the standard deviations of the e_{it}s. In all cases, 2500 replications were used.

Figures 1 (AO) and 2 (IO) are typical of the simulations results. The top panels of both figures show that when $\delta' = (0, 0)$ and $(1, 0)$, the GC test has an exact size close to the nominal size of 5%.[3] The lower panels of Figs. 1 and 2 show that size distortions are larger when both series are subject to mean shifts rather than to Y_2 alone. Size distortions in AO models (Fig. 1) are smaller the closer is b_{22} to the unit circle. This is because when b_{22} is unity and $b_{21} = 0$, Y_2 is a random walk, and a mean shift only induces a one time outlier in the first differences of the data. On the other hand, negative serial correlation in y_{2t} reduces the unconditional variance of the series and effectively increases the relative magnitude of the break. In such cases, a 100% rejection rate is possible even if there is no causal relationship in the data.

[3]We considered three values of $b_{11}(-0.5, 0.0, 0.5)$. Values for b_{12} and b_{22} are taken from the parameter set $-0.6, -0.4, -0.2, 0.0, 0.2, 0.4, 0.6$ for each value of b_{11}. All parameterizations have eigenvalues of A that lie inside the unit circle with unequal roots.

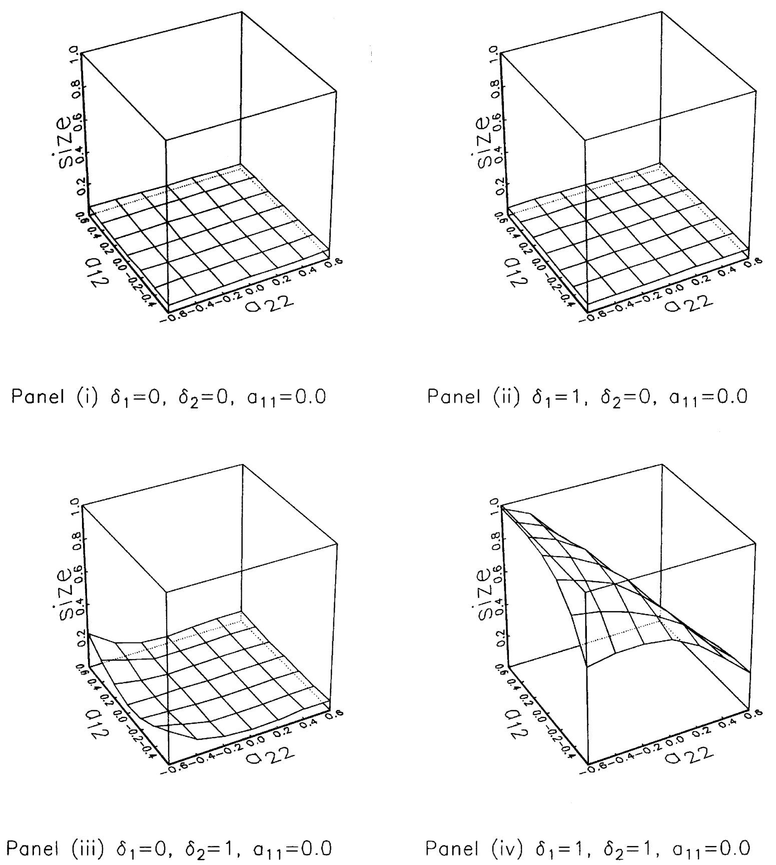

Figure 1. AO: Size of Granger casuality test, $T = 200$.

Now consider the IO model in Fig. 2. While it is still true that the unconditional variance of Y_2 increases as b_{22} approaches one, the magnitude of the mean shift also increases. When $b_{22} = 1$ the mean shift becomes a trend shift and y_{2t} becomes a unit root series with a slope shift. Thus, the magnitude of size distortions also reflects this discontinuity in the functional form of the deterministic components of the series.

Figures 1 and 2 also illustrate how given a set of dynamic parameters, omitting AO and IO type mean shifts can have rather different quantitative implications, even in the simple bivariate case with common break dates.

Because omitted mean shifts lead to inconsistent estimates of the VAR, forecasts and impulse response functions based upon the OLS estimates will also be inconsistent. The obvious solution to removing the least squares bias is to account for the mean shifts when estimating the model. In the next section, we will show that the effectiveness of detecting and removing mean shifts also depends on the model type on hand.

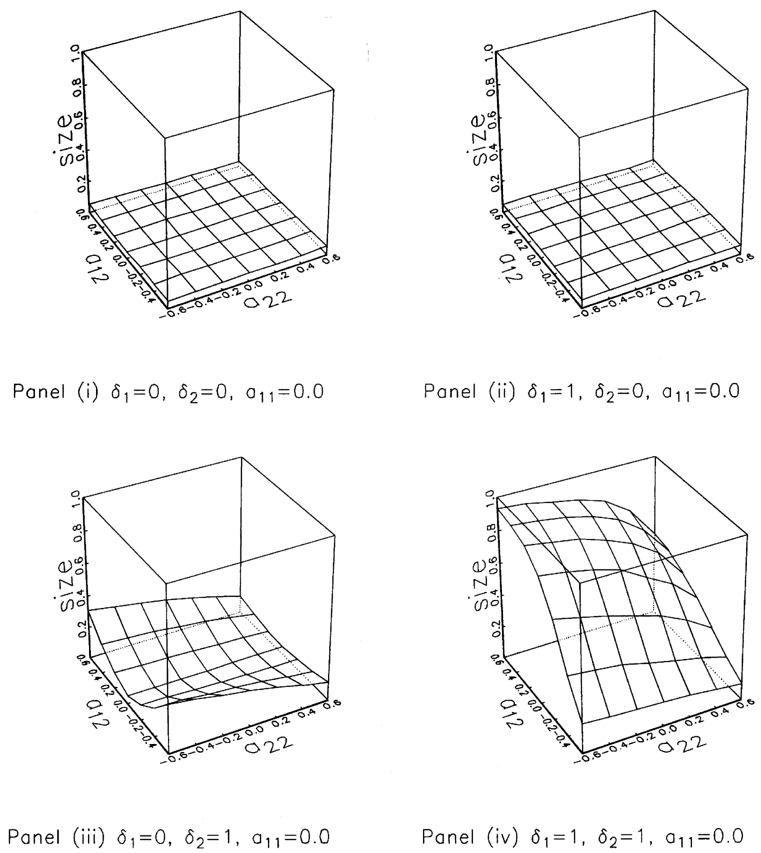

Figure 2. IO: Size of Granger casuality test, $T = 200$.

3 HOW TO REMOVE THE BREAKS?

Suppose a practitioner has an idea of the number of breaks and either knows the breaks dates or has estimated them. Then, there are two strategies for removing the mean shifts. The first is estimate the mean shifts, remove them from the data, and then form a VAR for the demeaned data. Examples include Blanchard and Quah[13] and Gambe and Joutz[14] where a mean shift in output growth in 1973 was suspected, and the break was removed from the series prior to estimating the VAR. We refer to this as the two-step method. The second approach is to work directly with the VAR by adding intercept shifts to the equations of the VAR. We refer to this as the one-step approach. From a practitioner's point of view, is one method preferred over the other?

Consider first the two-step method under which the VAR is fitted to the data with one mean shift removed from each series:

$$\hat{y}_{it} = \hat{B}_i' \hat{y}_{t-1} + \hat{u}_{it} \tag{8}$$

where $\hat{y}_{it} = y_{it} - \hat{\mu}_i - \hat{\delta}_i DU_{it}$, $\hat{\mu}_i$ and $\hat{\delta}_i$ are obtained from OLS estimation of (1). From (3), the AO model written in terms of data with no mean shifts is:

$$\breve{y}_{it} = B_i' \breve{y}_{t-1} + u_{it}$$

where $\breve{y}_{it} = y_t - \mu_i - \delta_i DU_{it}$. Clearly, this differs from (8) only to the extent that $\hat{y}_{it}$ is based on $\hat{\mu}_i$ and $\hat{\delta}_i$ rather than μ_i and δ_i. The two-step procedure is thus correct for the AO model, whether or not the break dates coincide.

A typical equation of the IO model in terms of demeaned data $\breve{y}_{it} = y_{it} - \mu_i - (\delta_i{}^I)' DU_t$ is:

$$\breve{y}_{it} = B_i' \breve{y}_{t-1} + B_i'(\delta_i{}^I)' DU_{t-1} + u_{it}$$

Since $\breve{y}_{it-1}$ is uncorrelated with DU_{t-1}, omitting DU_{t-1} from a VAR in $\breve{y}_{it}$ will only have efficiency implications. But except when the break dates coincide, $\breve{y}_{it}$ is the result of removing mean shifts occurring to the entire data vector from y_{it}. Thus, when the data is IO, using $\hat{y}_{it}$ which only removes one mean shift to form the VAR will generally be inappropriate.

Now consider the one-step procedure which estimates a VAR using the raw data but allows for one intercept shift per equation:

$$y_{it} = \hat{\mu}_i^* + \hat{\delta}_i^{*\prime} dU_{it} + \hat{B}_i' y_{t-1} + \hat{u}_{it} \tag{9}$$

A typical IO equation is

$$y_{it} = \mu_i^* + (\delta_i{}^I)' DU_t + B_i' y_{t-1} + u_{it}$$

When the break dates coincide, (9) is clearly the correct specification for the IO model. However, when the break dates are different, $(\delta_i{}^I)' DU_t$ will generally exhibit multiple shifts. One would need to account for multiple intercept shifts to every equation in the VAR.

Now, suppose the data is generated under the AO model. A typical equation is

$$y_{it} = \mu^* + (\delta_i{}^A)' DU_t + (\gamma_i{}^A)' d_t + B_i' y_{t-1} + u_{it}$$

If the break dates coincide, the equation differs from (9) only in that d_t is omitted. But this dummy for a one time outlier has asymptotically negligible effects on the remaining estimates, so that while the one-step procedure is inefficient if d_t is omitted, it nevertheless will deliver consistent estimates for the slope and intercept parameters when the DGP is an AO model. Now if the break dates do not coincide, $(\delta_i{}^A)' DU_t$ will exhibit multiple shifts. Then following arguments analogous to the IO case, the one-step procedure would need to allow for multiple intercept shifts.

The main observations can be summarized as follows. When the break dates are the same, both the one and two step procedures are appropriate. However, the two step procedure is more efficient for AO data while the one step procedure is more efficient for IO data. When the break dates differ, the two step procedure that removes one mean shift per series remains correct for AO data but is incorrect for IO data. The one step procedure with one intercept shift is incorrect for either data type. Both strategies can be made robust to mean shifts at different dates by controlling for multiple shifts. However, the two step procedure with multiple shifts is inefficient for AO data with single shifts.

This discussion clearly illustrates that distinct known break dates or unknown break dates seriously complicates estimation strategies. One could always take the robust route and include n breaks per series or equation. But, this approach becomes computationally burdensome as n increases and could be inefficient if break dates coincide. To help provide some practical advice, we conducted extensive finite sample simulations for a bivariate VAR. The goal here is to compare the one and two-step procedures in the simplest possible environment. We hope this analysis stimulates further simulation studies of higher order VARs.

For a bivariate VAR, we consider six procedures of removing mean shifts from the data. As a matter of notation, the procedures are labeled $xSyzB$, where x is 1 for one-step procedures and 2 for two-step procedures. If the break dates are estimated from the VAR, $y = V$; if the break dates are estimated based on (1) (univariate method), $y = U$. If no breaks are included, $z = 0$; if one break is included per equation or series, $z = 1$; if two breaks are included per equation or series, $z = 2$. Recall that $DU_t = (dU_{1t}, dU_{2t})'$. Details of break date estimation are given in Section 3.2.

The following are the one-step procedures.[4]

1. 1SU2B: Given the break dates, the VAR analysis is based on the regressions

$$y_{it} = \mu_i + \delta_i dU_{1t} + \pi_i dU_{2t} + b_{i1} y_{1t-1} + b_{i2} y_{2t-1} + u_{it}, \quad i = 1, 2 \tag{10}$$

If the break dates are unknown, they are estimated using the regressions

$$y_{it} = \mu_i + \delta_i dU_{1t} + \pi_i dU_{2t} + z_{it}, \quad i = 1, 2 \tag{11}$$

2. 1SU1B: Given the break dates, the VAR analysis is based on the regressions

$$y_{it} = \mu_i + \delta_i dU_{1t} + b_{i1} y_{1t-1} + b_{i2} y_{2t-1} + u_{it}, \quad i = 1, 2 \tag{12}$$

If the break dates are unknown, they are estimated using the regressions

$$y_{it} = \mu_i + \delta_i dU_{it} + z_{it}, \quad i = 1, 2 \tag{13}$$

3. 1SV1B: Given the break dates, the VAR analysis is based on regression (12). If the break dates are unknown, they are estimated using regression (12).
4. 1SV2B: Given the break dates, the VAR analysis is based on regression (10). If the break dates are unknown, they are estimated using regression (11).

[4]The one step procedure can also be augmented to include lags of dU_t (or d_t). These are irrelevant lags in an IO setting but could improve the efficiency of the estimates in the AO case.

5. 1SV0B: No breaks are included in the model and the VAR analysis is based on the regressions

$$y_{it} = \mu_i + b_{i1}y_{1t-1} + b_{i2}y_{2t-1} + u_{it}, \quad i = 1, 2$$

The following are the two-step procedures. Let $\hat{y}_{it}$ generically denote OLS residuals from the regression of y_{it} on deterministic regressors. For all two-step procedures, the VAR analysis is based on the regression:

$$\hat{y}_{it} = b_{i1}\hat{y}_{1t-1} + b_{i2}\hat{y}_{2t-1} + u_{it}, \quad i = 1, 2$$

1. 2SU1B: Given the break dates, $\hat{y}_{it}$ is obtained using regression (13). If the break dates are unknown, they are estimated using regression (13).
2. 2SU2B: Given the break dates, the $\hat{y}_{it}$ is obtained using regression (11). If the break dates are unknown, they are estimated using regression (11).
3. 2SU0B: No breaks are included in the estimation and the $\hat{y}_{it} = y_{it} - T^{-1}\sum_{t=1}^{T} y_{it}$.

To assess the merits of the various procedures, we use the trace of the mean-squared error (MSE) of the h step ahead forecasts as the metric for comparison. This amounts to summing the h step ahead squared forecasts errors of y_{1t} and y_{2t}. We report results for $h = 5$. With smaller h, it is difficult to see substantial differences in the procedures. Results with $h = 1$, 2, 3, 4 are available upon request. Our experiments are based on 2500 replications using $N(0, 1)$ errors in Gauss with seed = 999. Without loss of generality, Γ_{11} and Γ_{22} are normalized to 1. We also set $a_{21} = 0.3$ and $a_{22} = 0.6$. By varying Γ_{12}, Γ_{22}, a_{11}, and a_{12}, we obtain 16 parameterizations for given δ_1, δ_2, T_{B_1} and T_{B_2}. We then consider four combinations of δ_1 and δ_2, and two combinations of the break dates. In all cases, we generate 205 observations. Estimations are based on the first $T = 200$ observations. We then evaluate the out-of-sample forecasts for the next five periods. We organize the simulations results according to whether the break dates are treated as known or unknown.

3.1 Known Break Dates

When the break dates are known, several of the procedures become equivalent. If the break dates are the same for both series, then, with the exception of procedure 1SV0B, the one-step procedures are equivalent (two breaks cannot be included in a regression if the break dates are the same). Similarly, procedures 2SU1B and 2SU2B are equivalent when the break dates are the same. If the break dates are not the same but known, then procedures 1SU1B and 1SV1B are equivalent, as are procedures 1SU2B and 1SV2B.

The results are reported in Table 1 (same break dates) and Table 2 (different break dates). To make comparisons across the procedures easier, we report MSE relative to the MSE for procedure 2SU2B. In other words, for a given

Table 1. MSE Relative to 2SU2B; Known Break Dates: $TB_1 = 100$, $TB_2 = 100$; $a_{21} = 0.3$, $a_{22} = 0.6$

				One Step					Two Steps		
a_{11}	a_{12}	Γ_{12}	Γ_{21}	U2B	U1B	V1B	V2B	0B	U1B	U2B	0B
No Break											
0.60	0.30	0.00	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.60	0.00	0.00	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.30	0.30	0.00	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.30	0.00	0.00	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.60	0.30	0.00	0.30	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.60	0.00	0.00	0.30	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.30	0.30	0.00	0.30	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.30	0.00	0.00	0.30	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.60	0.30	0.30	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.60	0.00	0.30	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.30	0.30	0.30	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.30	0.00	0.30	0.00	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.60	0.30	0.30	0.30	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.60	0.00	0.30	0.30	1.00	1.00	1.00	1.00	0.97	1.00	1.00	0.97
0.30	0.30	0.30	0.30	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
0.30	0.00	0.30	0.30	1.00	1.00	1.00	1.00	0.98	1.00	1.00	0.98
AO DGP							$\delta_1 = 1$, $\delta_2 = 1$				
0.60	0.30	0.00	0.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
0.60	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.07	1.00	1.00	1.08
0.30	0.30	0.00	0.00	1.00	1.00	1.00	1.00	1.05	1.00	1.00	1.05
0.30	0.00	0.00	0.00	1.00	1.00	1.00	1.00	1.11	1.00	1.00	1.11
0.60	0.30	0.00	0.30	1.00	1.00	1.00	1.00	1.06	1.00	1.00	1.07
0.60	0.00	0.00	0.30	1.00	1.00	1.00	1.00	1.08	1.00	1.00	1.09
0.30	0.30	0.00	0.30	1.00	1.00	1.00	1.00	1.09	1.00	1.00	1.10

0.30	0.00	0.00	0.30	1.00	1.00	1.00	1.00	1.11	1.00	1.00	1.11
0.60	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.06	1.00	1.00	1.06
0.60	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.12	1.00	1.00	1.12
0.30	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.09	1.00	1.00	1.10
0.30	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.13	1.00	1.00	1.14
0.60	0.30	0.30	0.30	1.00	1.00	1.00	1.00	1.06	1.00	1.00	1.06
0.60	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.06	1.00	1.00	1.07
0.30	0.30	0.30	0.30	1.00	1.00	1.00	1.00	1.08	1.00	1.00	1.09
0.30	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.08	1.00	1.00	1.09
IO DGP							$\delta_1 = 1,\ \delta_2 = 1$				
0.60	0.30	0.00	0.00	0.94	0.94	0.94	0.94	1.09	1.00	1.00	1.10
0.60	0.00	0.00	0.00	0.99	0.99	0.99	0.99	1.36	1.00	1.00	1.40
0.30	0.30	0.00	0.00	0.99	0.99	0.99	0.99	1.28	1.00	1.00	1.32
0.30	0.00	0.00	0.00	0.99	0.99	0.99	0.99	1.38	1.00	1.00	1.42
0.60	0.30	0.00	0.30	0.99	0.99	0.99	0.99	1.26	1.00	1.00	1.29
0.60	0.00	0.00	0.30	1.00	1.00	1.00	1.00	1.25	1.00	1.00	1.27
0.30	0.30	0.00	0.30	1.00	1.00	1.00	1.00	1.29	1.00	1.00	1.31
0.30	0.00	0.00	0.30	1.00	1.00	1.00	1.00	1.27	1.00	1.00	1.30
0.60	0.30	0.30	0.00	0.99	0.99	0.99	0.99	1.25	1.00	1.00	1.28
0.60	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.37	1.00	1.00	1.39
0.30	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.29	1.00	1.00	1.32
0.30	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.35	1.00	1.00	1.37
0.60	0.30	0.30	0.30	1.00	1.00	1.00	1.00	1.17	1.00	1.00	1.19
0.60	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.17	1.00	1.00	1.18
0.30	0.30	0.30	0.30	1.00	1.00	1.00	1.00	1.20	1.00	1.00	1.22
0.30	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.21	1.00	1.00	1.22

Table 2. MSE Relative to 2SU2B; Known Break Dates: $TB_1 = 75$, $TB_2 = 150$; $a_{21} = 0.3$, $a_{22} = 0.6$

				One Step					Two Steps		
a_{11}	a_{12}	Γ_{12}	Γ_{21}	U2B	U1B	V1B	V2B	0B	U1B	U2B	0B
AO DGP							$\delta_1 = 0,\ \delta_2 = 1$				
0.60	0.30	0.00	0.00	1.00	1.02	1.02	1.00	1.01	1.01	1.00	1.01
0.60	0.00	0.00	0.00	1.00	0.98	0.98	1.00	1.08	0.98	1.00	1.09
0.30	0.30	0.00	0.00	1.01	1.02	1.02	1.01	1.05	0.99	1.00	1.05
0.30	0.00	0.00	0.00	1.00	0.99	0.99	1.00	1.12	0.99	1.00	1.12
0.60	0.30	0.00	0.30	1.00	1.03	1.03	1.00	1.10	0.98	1.00	1.10
0.60	0.00	0.00	0.30	1.00	0.98	0.98	1.00	1.10	0.98	1.00	1.10
0.30	0.30	0.00	0.30	1.00	1.02	1.02	1.00	1.13	0.99	1.00	1.13
0.30	0.00	0.00	0.30	1.00	0.99	0.99	1.00	1.13	0.99	1.00	1.13
0.60	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.08	0.98	1.00	1.09
0.60	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.14	0.98	1.00	1.14
0.30	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.12	0.99	1.00	1.12
0.30	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.16	0.99	1.00	1.17
0.60	0.30	0.30	0.30	1.00	1.00	1.00	1.00	1.09	0.98	1.00	1.09
0.60	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.05	0.97	1.00	1.06
0.30	0.30	0.30	0.30	1.00	1.00	1.00	1.00	1.12	0.99	1.00	1.12
0.30	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.10	0.98	1.00	1.10
							$\delta_1 = 1,\ \delta_2 = 1$				
0.60	0.30	0.00	0.00	1.00	1.02	1.02	1.00	0.98	1.01	1.00	0.98
0.60	0.00	0.00	0.00	1.00	0.98	0.98	1.00	1.08	0.98	1.00	1.09
0.30	0.30	0.00	0.00	1.01	1.02	1.02	1.01	1.04	0.99	1.00	1.05
0.30	0.00	0.00	0.00	1.00	0.99	0.99	1.00	1.14	0.99	1.00	1.14
0.60	0.30	0.00	0.30	1.00	1.04	1.04	1.00	1.09	0.98	1.00	1.09
0.60	0.00	0.00	0.30	1.00	0.98	0.98	1.00	1.11	0.98	1.00	1.11
0.30	0.30	0.00	0.30	1.00	1.02	1.02	1.00	1.13	0.99	1.00	1.13
0.30	0.00	0.00	0.30	1.00	0.99	0.99	1.00	1.14	0.99	1.00	1.15
0.60	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.07	0.98	1.00	1.08

0.60	0.00	0.30	0.00	1.00	0.99	0.99	1.00	1.16	0.98	1.00	1.16
0.30	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.12	0.99	1.00	1.12
0.30	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.18	0.99	1.00	1.19
0.60	0.30	0.30	0.30	1.00	1.00	1.00	1.00	1.08	0.98	1.00	1.09
0.60	0.00	0.30	0.30	1.00	0.99	0.99	1.00	1.07	0.97	1.00	1.08
0.30	0.30	0.30	0.30	1.00	1.00	1.00	1.00	1.12	0.99	1.00	1.12
0.30	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.11	0.98	1.00	1.12
IO DGP							$\delta_1=0,\ \delta_2=1$				
0.60	0.30	0.00	0.00	0.93	0.92	0.92	0.93	1.16	1.29	1.00	1.20
0.60	0.00	0.00	0.00	1.00	0.98	0.98	1.00	1.55	0.98	1.00	1.57
0.30	0.30	0.00	0.00	0.99	0.98	0.98	0.99	1.48	1.17	1.00	1.52
0.30	0.00	0.00	0.00	1.00	0.99	0.99	1.00	1.66	0.99	1.00	1.69
0.60	0.30	0.00	0.30	1.00	0.99	0.99	1.00	1.44	1.22	1.00	1.47
0.60	0.00	0.00	0.30	1.00	0.99	0.99	1.00	1.57	0.98	1.00	1.60
0.30	0.30	0.00	0.30	1.00	0.99	0.99	1.00	1.57	1.11	1.00	1.60
0.30	0.00	0.00	0.30	1.00	0.99	0.99	1.00	1.67	0.99	1.00	1.70
0.60	0.30	0.30	0.00	1.00	1.01	1.01	1.00	1.51	0.98	1.00	1.53
0.60	0.00	0.30	0.00	1.00	1.01	1.01	1.00	1.66	1.16	1.00	1.67
0.30	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.60	0.99	1.00	1.63
0.30	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.71	1.09	1.00	1.73
0.60	0.30	0.30	0.30	1.00	1.01	1.01	1.00	1.51	0.98	1.00	1.53
0.60	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.75	1.23	1.00	1.78
0.30	0.30	0.30	0.30	1.00	1.01	1.01	1.00	1.58	0.99	1.00	1.61
0.30	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.77	1.12	1.00	1.80
							$\delta_1=1,\ \delta_2=1$				
0.60	0.30	0.00	0.00	0.93	0.92	0.92	0.93	1.11	1.36	1.00	1.14
0.60	0.00	0.00	0.00	1.00	0.98	0.98	1.00	1.61	0.98	1.00	1.67
0.30	0.30	0.00	0.00	0.99	0.98	0.98	0.99	1.43	1.19	1.00	1.49
0.30	0.00	0.00	0.00	1.00	0.99	0.99	1.00	1.71	0.99	1.00	1.77
0.60	0.30	0.00	0.30	1.00	0.99	0.99	1.00	1.41	1.24	1.00	1.45
0.60	0.00	0.00	0.30	1.00	0.99	0.99	1.00	1.51	0.98	1.00	1.55
0.30	0.30	0.00	0.30	1.00	0.99	0.99	1.00	1.56	1.11	1.00	1.60

(*continued*)

Table 2. Continued

				One Step					Two Steps		
a_{11}	a_{12}	Γ_{12}	Γ_{21}	U2B	U1B	V1B	V2B	0B	U1B	U2B	0B
0.30	0.00	0.00	0.30	1.00	0.99	0.99	1.00	1.61	0.99	1.00	1.64
0.60	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.46	0.98	1.00	1.51
0.60	0.00	0.30	0.00	1.00	1.01	1.01	1.00	1.68	1.16	1.00	1.71
0.30	0.30	0.30	0.00	1.00	1.00	1.00	1.00	1.56	0.99	1.00	1.61
0.30	0.00	0.30	0.00	1.00	1.00	1.00	1.00	1.72	1.09	1.00	1.75
0.60	0.30	0.30	0.30	1.00	1.01	1.01	1.00	1.40	0.98	1.00	1.42
0.60	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.48	1.22	1.00	1.50
0.30	0.30	0.30	0.30	1.00	1.01	1.01	1.00	1.48	0.99	1.00	1.51
0.30	0.00	0.30	0.30	1.00	1.00	1.00	1.00	1.58	1.12	1.00	1.60

parameterization, we divide all the MSE by the MSE for 2SU2B. Thus, the entries for 2SU2B will always be 1.0. Entries less than 1.0 indicate a procedure that gives more accurate forecasts than 2SU2B whereas entries greater than 1.0 indicate a procedure that gives less accurate forecasts than 2SU2B. When there are no mean shifts, data generated by the AO and the IO models are observationally equivalent and we only report one set of results.

The first panel of Table 1 gives results when there are no breaks in the data. The procedures that omit the breaks have the lowest MSE as expected. However, including breaks does not significantly increase forecast error MSE even when two breaks are included in each series/equation. The second and third panels give results for breaks in both series that occur at the same dates, $T_B = 100$ (halfway through the sample) for the AO and IO DGPs respectively. Omitting the breaks leads to much higher MSE. Even including two breaks when only one break per series/equation is required, does not significantly increase MSE. These results hold whether the data is AO is IO. Because including the breaks does not significantly increase MSE but omitting them does, our results strongly suggest that breaks should be included if they are suspected. If we compare the one-step and two-step procedures, we see that they are, for practical purposes, equivalent. This is not surprising because, by the Frisch-Waugh theorem, a VAR in demeaned data and a VAR that includes the same intercept shifts used for the demeaning are asymptotically equivalent.

When the break dates are different, $T_{B_1} = 75$, $T_{B_2} = 150$, some differences in the procedures emerge. Consider the first and third panels of Table 2. Here, there is only a break to the second equation/series. For AO data, the 2SU1B procedure is often the best as would be expected in theory. When the data is IO, the 2SU1B procedure often has much higher MSE especially when there are two breaks. See the fourth panel. Interestingly, when the data is AO, the 2SU2B procedure is nearly as efficient as 2SU1B. The 2SU2B procedure remains robust to IO data and has MSE very close to the one-step procedures. The one-step procedures remain quite robust to AO data when the break dates are the same. But when the mean shifts are at different dates in AO data, the 1SV1B procedure can have higher MSE because there are two intercept shifts per equation in the VAR and only one is taken into account.

In accord with the theoretical discussion, when the break dates are known, the 1SV2B and 2SU2B procedures are the most robust when it is unknown whether the data is AO or IO. Our results do not indicate that one should be used over the other. The 2SU1B procedure is not recommended unless it is known that the data is of AO type or that the break dates are the same.

3.2 Unknown Break Dates

In practice, it is often the case that break dates are unknown and need to be estimated. Because we are focusing on mean shifts in stationary VARs, consistency of estimators of the break date ratios, $\lambda_i = T_{B_i}/T$, follows from results in Bai[15] and Bai and Perron.[16] Note that estimates of the break dates themselves are not consistent.

The break dates for the 1SU1B and 2SU1B procedures were estimated as follows. For each $T_{B_i} \in [0.15T, 0.85T]$, let $\mathrm{SSR}_i(T_{B_i})$ denote the sum of squared residuals from regression (13). Then

$$\hat{T}_{B_i} = \arg\min_{T_{B_i}} \mathrm{SSR}_i(T_{B_i})$$

The demeaned data $\hat{y}_{it}$ are then obtained from regression (13) using $\hat{T}_{B_i}$. Note that if $\hat{T}_{B_i} \neq \hat{T}_{B_j}$, then the VAR for the 1SU1B procedure is estimated using regression (10) rather than regression (12). This makes the 1SU1B procedure more robust to AO data than if regression (12) were always used.

For the 1SU2B and 2SU2B procedures, the two break dates are estimated jointly using (11). Following one of the methods proposed by Bai and Perron,[16] we estimate the break dates sequentially as follows. Letting $\mathrm{SSR}_i(T_{B_1})$ denote the sum of squared residuals from the ith regression in (13), then

$$\tilde{T}_{B_1} = \arg\min_{T_{B_1}} (\mathrm{SSR}_1(T_{B_1}) + \mathrm{SSR}_2(T_{B_1}))$$

is the estimator of the first break date. Note that $\tilde{T}_{B_1}$ is the least squares estimator of a single break date constrained to the be same for both series. $\tilde{T}_{B_1}$ is then used to define the dummy variable, $dU_{1t} = 1(t > \tilde{T}_{B_1})$ for regression (11). For each $T_{B_2} \in [0.15T, 0.85T]$, let $\mathrm{SSR}_i(T_{B_2}|\tilde{T}_{B_1})$ denote the sum of squared residuals from the ith regression in (11) using $\tilde{T}_{B_1}$ to define the dummy variable, dU_{1t}. Then,

$$\tilde{T}_{B_2} = \arg\min_{T_{B_2}} (\mathrm{SSR}_1(T_{B_2}|\tilde{T}_{B_1}) + \mathrm{SSR}_2(T_{B_2}|\tilde{T}_{B_1})$$

is the estimator of the second break date.

The break dates for the 1SV1B and 2SV2B procedures were estimated in exactly the same way as the U1B and U2B procedures except that regressions (12) and (10) were used instead of regressions (13) and (11) to construct the sum of squared residuals. Therefore, for the 1SV2B procedure, the two break dates are estimated jointly using both equations of the VAR.

The simulation results are reported in Tables 3 and 4. Table 3 gives results for break dates $T_{B_1} = T_{B_2} = 100$. As would be expected for the case where the break dates are the same, it matters little whether the data are generated according to the AO or IO models. Both the one-step and two-step procedures are fairly robust to the type of data. Because the break dates are the same, the procedures that have one break per series or equation tend to have the more accurate forecasts as expected. In most cases 2SU1B dominates 2SU2B, and 1SU1B generally dominates 1S2UB. On the other hand, 1SV2B dominates 1SV1B. This seems surprising at first since there is only one break per equation. But, recall that 1SV2B estimates the break dates sequentially using both equations of the VAR simultaneously and the first break date is estimated under the correct restriction that the break date is the same for both series. Thus, following Bai et al.[17] the estimate of the first break date is more efficient than estimates based on the individual equations as is done for 1SV1B. Even though 1SV2B includes a

Table 3. MSE Relative to 2SU2B; Unknown Break Dates: $TB_1 = 100$, $TB_2 = 100$; $a_{21} = 0.3$, $a_{22} = 0.6$

				One Step					Two Steps		
a_{11}	a_{12}	Γ_{12}	Γ_{21}	U2B	U1B	V1B	V2B	0B	U1B	U2B	0B
AO DGP							$\delta_1 = 0$, $\delta_2 = 1$				
0.60	0.30	0.00	0.00	1.01	1.01	0.94	0.94	0.92	1.00	1.00	0.92
0.60	0.00	0.00	0.00	1.00	1.00	1.00	0.97	0.99	0.99	1.00	0.99
0.30	0.30	0.00	0.00	1.01	1.00	0.96	0.95	0.97	0.99	1.00	0.97
0.30	0.00	0.00	0.00	1.00	0.98	0.98	0.97	1.01	0.98	1.00	1.01
0.60	0.30	0.00	0.30	1.01	1.00	0.96	0.95	1.00	0.99	1.00	1.00
0.60	0.00	0.00	0.30	1.00	1.00	1.01	0.97	1.00	0.99	1.00	1.00
0.30	0.30	0.00	0.30	1.00	0.99	0.97	0.96	1.02	0.98	1.00	1.02
0.30	0.00	0.00	0.30	1.00	0.98	0.98	0.97	1.01	0.97	1.00	1.01
0.60	0.30	0.30	0.00	1.00	1.01	0.99	0.97	0.99	0.99	1.00	0.99
0.60	0.00	0.30	0.00	1.00	1.00	1.00	0.97	1.03	1.00	1.00	1.03
0.30	0.30	0.30	0.00	1.00	0.99	0.99	0.97	1.01	0.98	1.00	1.02
0.30	0.00	0.30	0.00	1.00	0.98	0.98	0.97	1.04	0.99	1.00	1.04
0.60	0.30	0.30	0.30	1.00	1.00	1.00	0.97	1.00	0.99	1.00	1.00
0.60	0.00	0.30	0.30	1.00	1.00	1.01	1.00	0.98	1.02	1.00	0.98
0.30	0.30	0.30	0.30	1.00	0.99	0.98	0.97	1.02	0.98	1.00	1.02
0.30	0.00	0.30	0.30	1.00	0.98	1.00	0.99	1.01	1.00	1.00	1.01
							$\delta_1 = 1$, $\delta_2 = 1$				
0.60	0.30	0.00	0.00	1.01	0.99	0.99	0.99	0.90	0.98	1.00	0.90
0.60	0.00	0.00	0.00	1.00	0.96	0.99	0.96	1.01	0.97	1.00	1.01
0.30	0.30	0.00	0.00	1.01	0.98	1.00	0.95	0.98	0.97	1.00	0.98
0.30	0.00	0.00	0.00	1.00	0.96	1.00	0.96	1.06	0.97	1.00	1.06
0.60	0.30	0.00	0.30	1.00	0.98	1.01	0.96	1.01	0.97	1.00	1.01
0.60	0.00	0.00	0.30	1.00	0.97	0.99	0.97	1.03	0.96	1.00	1.03
0.30	0.30	0.00	0.30	1.01	0.96	0.99	0.96	1.04	0.96	1.00	1.04
0.30	0.00	0.00	0.30	1.00	0.96	1.00	0.96	1.06	0.96	1.00	1.06
0.60	0.30	0.30	0.00	1.00	0.97	1.00	0.96	1.00	0.97	1.00	1.00

(*continued*)

Table 3. Continued

				One Step					Two Steps		
a_{11}	a_{12}	Γ_{12}	Γ_{21}	U2B	U1B	V1B	V2B	0B	U1B	U2B	0B
0.60	0.00	0.30	0.00	1.00	0.97	1.00	0.97	1.07	0.97	1.00	1.08
0.30	0.30	0.30	0.00	1.01	0.97	1.01	0.96	1.05	0.97	1.00	1.05
0.30	0.00	0.30	0.00	1.00	0.97	1.01	0.97	1.09	0.97	1.00	1.10
0.60	0.30	0.30	0.30	1.00	0.97	1.01	0.98	1.01	0.97	1.00	1.02
0.60	0.00	0.30	0.30	1.00	0.97	1.01	0.98	1.01	0.98	1.00	1.02
0.30	0.30	0.30	0.30	1.00	0.97	1.01	0.97	1.04	0.97	1.00	1.05
0.30	0.00	0.30	0.30	1.00	0.97	1.02	0.97	1.05	0.97	1.00	1.05
IO DGP							$\delta_1=0,\ \delta_2=1$				
0.60	0.30	0.00	0.00	1.00	0.92	0.95	0.92	1.03	0.93	1.00	1.05
0.60	0.00	0.00	0.00	1.00	0.99	0.99	0.95	1.19	0.98	1.00	1.21
0.30	0.30	0.00	0.00	1.01	0.94	0.96	0.94	1.17	0.94	1.00	1.19
0.30	0.00	0.00	0.00	1.00	0.98	0.97	0.95	1.24	0.97	1.00	1.26
0.60	0.30	0.00	0.30	1.01	0.95	0.99	0.95	1.14	0.95	1.00	1.16
0.60	0.00	0.00	0.30	1.00	1.00	0.99	0.96	1.21	0.99	1.00	1.22
0.30	0.30	0.00	0.30	1.00	0.96	0.98	0.96	1.20	0.96	1.00	1.22
0.30	0.00	0.00	0.30	1.00	0.98	0.98	0.96	1.25	0.97	1.00	1.26
0.60	0.30	0.30	0.00	1.00	1.00	0.99	0.95	1.18	0.98	1.00	1.19
0.60	0.00	0.30	0.00	1.00	0.97	0.98	0.96	1.26	0.97	1.00	1.27
0.30	0.30	0.30	0.00	1.01	0.98	0.97	0.96	1.22	0.97	1.00	1.23

0.30	0.00	0.30	0.00	1.00	0.96	0.97	0.96	1.28	0.97	1.00	1.29
0.60	0.30	0.30	0.30	1.00	1.00	0.99	0.96	1.18	0.99	1.00	1.20
0.60	0.00	0.30	0.30	1.00	0.95	0.97	0.96	1.27	0.96	1.00	1.29
0.30	0.30	0.30	0.30	1.00	0.98	0.97	0.96	1.21	0.98	1.00	1.23
0.30	0.00	0.30	0.30	1.00	0.96	0.96	0.96	1.29	0.96	1.00	1.31
							$\delta_1 = 1,\ \delta_2 = 1$				
0.60	0.30	0.00	0.00	0.95	0.90	0.90	0.90	1.04	0.93	1.00	1.05
0.60	0.00	0.00	0.00	1.00	0.95	0.95	0.95	1.30	0.95	1.00	1.33
0.30	0.30	0.00	0.00	1.00	0.93	0.94	0.94	1.21	0.93	1.00	1.24
0.30	0.00	0.00	0.00	1.00	0.95	0.95	0.95	1.32	0.95	1.00	1.35
0.60	0.30	0.00	0.30	1.00	0.95	0.96	0.95	1.20	0.95	1.00	1.23
0.60	0.00	0.00	0.30	1.00	0.96	0.97	0.97	1.20	0.96	1.00	1.23
0.30	0.30	0.00	0.30	1.00	0.96	0.97	0.96	1.24	0.96	1.00	1.26
0.30	0.00	0.00	0.30	1.00	0.96	0.96	0.96	1.22	0.96	1.00	1.24
0.60	0.30	0.30	0.00	1.00	0.95	0.97	0.96	1.20	0.95	1.00	1.23
0.60	0.00	0.30	0.00	1.00	1.00	0.97	0.96	1.32	0.99	1.00	1.34
0.30	0.30	0.30	0.00	1.00	0.96	0.97	0.96	1.24	0.96	1.00	1.27
0.30	0.00	0.30	0.00	1.00	0.98	0.97	0.97	1.30	0.98	1.00	1.33
0.60	0.30	0.30	0.30	1.00	0.96	0.98	0.97	1.13	0.96	1.00	1.15
0.60	0.00	0.30	0.30	1.00	0.99	0.97	0.96	1.12	1.00	1.00	1.13
0.30	0.30	0.30	0.30	1.00	0.96	0.97	0.96	1.16	0.96	1.00	1.18
0.30	0.00	0.30	0.30	1.00	0.98	0.97	0.97	1.16	0.99	1.00	1.18

Table 4. MSE Relative to 2SU2B; Unknown Break Dates: $TB_1 = 75$, $TB_2 = 150$; $a_{21} = 0.3$, $a_{22} = 0.6$

				One Step					Two Steps		
a_{11}	a_{12}	Γ_{12}	Γ_{21}	U2B	U1B	V1B	V2B	0B	U1B	U2B	0B
AO DGP							$\delta_1 = 1,\ \delta_2 = 0$				
0.60	0.30	0.00	0.00	1.01	1.03	0.98	0.98	0.93	0.99	1.00	0.93
0.60	0.00	0.00	0.00	1.00	1.01	1.02	1.01	1.04	0.99	1.00	1.04
0.30	0.30	0.00	0.00	1.01	1.03	0.99	0.99	1.00	1.00	1.00	1.01
0.30	0.00	0.00	0.00	1.00	1.01	1.02	1.02	1.09	1.00	1.00	1.09
0.60	0.30	0.00	0.30	1.01	1.04	1.00	0.98	1.06	0.99	1.00	1.06
0.60	0.00	0.00	0.30	1.00	1.00	1.01	1.00	1.05	0.99	1.00	1.06
0.30	0.30	0.00	0.30	1.01	1.03	0.99	0.99	1.09	0.99	1.00	1.10
0.30	0.00	0.00	0.30	1.00	1.01	1.00	1.01	1.10	1.00	1.00	1.10
0.60	0.30	0.30	0.00	1.01	1.02	1.02	1.00	1.04	0.99	1.00	1.05
0.60	0.00	0.30	0.00	1.00	1.00	1.02	1.01	1.11	1.00	1.00	1.11
0.30	0.30	0.30	0.00	1.01	1.02	1.01	1.00	1.08	1.00	1.00	1.09
0.30	0.00	0.30	0.00	1.00	1.00	1.01	1.01	1.14	1.00	1.00	1.14
0.60	0.30	0.30	0.30	1.00	1.02	1.01	1.00	1.05	0.99	1.00	1.06
0.60	0.00	0.30	0.30	1.00	0.98	0.99	1.01	1.00	0.99	1.00	1.00
0.30	0.30	0.30	0.30	1.00	1.01	1.01	1.01	1.09	1.00	1.00	1.09
0.30	0.00	0.30	0.30	1.00	0.99	1.00	1.02	1.06	1.00	1.00	1.07
							$\delta_1 = 1,\ \delta_2 = 1$				
0.60	0.30	0.00	0.00	1.01	1.02	1.04	0.99	0.91	0.98	1.00	0.91
0.60	0.00	0.00	0.00	1.00	0.96	0.99	1.02	1.04	0.97	1.00	1.04
0.30	0.30	0.00	0.00	1.01	1.00	1.03	0.98	0.99	0.98	1.00	0.99
0.30	0.00	0.00	0.00	1.00	0.97	1.02	1.03	1.09	0.97	1.00	1.10
0.60	0.30	0.00	0.30	1.01	1.03	1.05	1.01	1.04	0.97	1.00	1.04
0.60	0.00	0.00	0.30	1.00	0.98	1.00	1.04	1.08	0.98	1.00	1.08
0.30	0.30	0.00	0.30	1.00	1.01	1.03	1.05	1.08	0.97	1.00	1.09
0.30	0.00	0.00	0.30	1.00	0.99	1.02	1.06	1.11	0.98	1.00	1.12
0.60	0.30	0.30	0.00	1.01	0.98	1.01	1.02	1.02	0.97	1.00	1.03

0.60	0.00	0.30	0.00	1.00	0.96	1.02	1.05	1.12	0.96	1.00	1.13
0.30	0.30	0.30	0.00	1.00	0.98	1.03	1.04	1.08	0.98	1.00	1.08
0.30	0.00	0.30	0.00	1.00	0.98	1.05	1.08	1.16	0.98	1.00	1.16
0.60	0.30	0.30	0.30	1.00	1.00	1.03	1.05	1.06	0.98	1.00	1.06
0.60	0.00	0.30	0.30	1.00	0.97	1.01	1.04	1.03	0.97	1.00	1.04
0.30	0.30	0.30	0.30	1.00	1.00	1.03	1.07	1.09	0.98	1.00	1.10
0.30	0.00	0.30	0.30	1.00	0.98	1.04	1.05	1.08	0.98	1.00	1.09
IO DGP							$\delta_1 = 1,\ \delta_2 = 0$				
0.60	0.30	0.00	0.00	0.98	0.97	0.97	0.96	1.19	1.00	1.00	1.23
0.60	0.00	0.00	0.00	1.00	1.01	1.01	0.98	1.52	1.00	1.00	1.54
0.30	0.30	0.00	0.00	1.00	1.00	1.00	0.98	1.47	0.99	1.00	1.51
0.30	0.00	0.00	0.00	1.00	1.00	1.00	0.98	1.63	0.99	1.00	1.66
0.60	0.30	0.00	0.30	1.00	1.00	1.01	1.00	1.42	0.99	1.00	1.45
0.60	0.00	0.00	0.30	1.00	1.01	1.01	1.00	1.55	1.00	1.00	1.57
0.30	0.30	0.00	0.30	1.00	1.00	1.00	1.00	1.55	0.99	1.00	1.58
0.30	0.00	0.00	0.30	1.00	1.00	1.00	1.00	1.65	0.99	1.00	1.68
0.60	0.30	0.30	0.00	1.00	1.03	1.00	0.99	1.48	1.00	1.00	1.50
0.60	0.00	0.30	0.00	1.00	0.99	1.00	0.99	1.63	0.99	1.00	1.64
0.30	0.30	0.30	0.00	1.00	1.01	1.00	1.00	1.58	0.99	1.00	1.60
0.30	0.00	0.30	0.00	1.00	0.99	1.00	0.99	1.69	0.99	1.00	1.71
0.60	0.30	0.30	0.30	1.00	1.02	1.01	0.99	1.49	1.00	1.00	1.51
0.60	0.00	0.30	0.30	1.00	0.98	0.99	0.99	1.71	0.98	1.00	1.74
0.30	0.30	0.30	0.30	1.00	1.01	1.00	1.00	1.56	0.99	1.00	1.59
0.30	0.00	0.30	0.30	1.00	0.98	0.99	1.00	1.75	0.99	1.00	1.77
							$\delta_1 = 1,\ \delta_2 = 1$				
0.60	0.30	0.00	0.00	0.96	1.04	0.91	1.05	1.10	1.29	1.00	1.13
0.60	0.00	0.00	0.00	1.00	1.10	0.97	1.21	1.59	1.10	1.00	1.65
0.30	0.30	0.00	0.00	0.99	1.04	0.98	1.30	1.42	1.18	1.00	1.49
0.30	0.00	0.00	0.00	1.00	1.03	0.99	1.35	1.71	1.04	1.00	1.76
0.60	0.30	0.00	0.30	1.00	0.99	0.99	1.26	1.40	1.17	1.00	1.44
0.60	0.00	0.00	0.30	1.00	0.98	0.99	1.30	1.50	0.97	1.00	1.54
0.30	0.30	0.00	0.30	1.00	0.99	1.00	1.37	1.55	1.07	1.00	1.59

(*continued*)

Table 4. Continued

				One Step					Two Steps		
a_{11}	a_{12}	Γ_{12}	Γ_{21}	U2B	U1B	V1B	V2B	0B	U1B	U2B	0B
0.30	0.00	0.00	0.30	1.00	0.98	0.99	1.41	1.60	0.98	1.00	1.63
0.60	0.30	0.30	0.00	1.00	1.11	1.00	1.16	1.44	1.10	1.00	1.49
0.60	0.00	0.30	0.00	1.00	1.13	1.01	1.24	1.67	1.23	1.00	1.70
0.30	0.30	0.30	0.00	1.00	1.05	1.01	1.29	1.56	1.04	1.00	1.61
0.30	0.00	0.30	0.00	1.00	1.05	1.00	1.33	1.71	1.13	1.00	1.75
0.60	0.30	0.30	0.30	1.00	1.01	1.02	1.25	1.39	0.98	1.00	1.42
0.60	0.00	0.30	0.30	1.00	1.00	1.00	1.41	1.47	1.22	1.00	1.49
0.30	0.30	0.30	0.30	1.00	1.00	1.01	1.34	1.48	0.98	1.00	1.50
0.30	0.00	0.30	0.30	1.00	1.00	1.02	1.46	1.58	1.12	1.00	1.60

superfluous break which reduces forecast precision, the greater efficiency of the first break date estimator outweighs this effect and results in more accurate forecasts. Clearly this does not occur for 2SU2B even though the break dates are estimated simultaneously using both series. In this case, the sampling variability of the estimate of the superfluous break coefficient is greater because the variance of z_t is larger than the variance of e_t. Thus while a more precise first break date estimate improves the forecast for 2SU2B, this improvement is more than o.set by the increase in forecast MSE from the extra break.

When the break dates are different, some different and interesting patterns emerge. These results are given in Table 4. When the data is AO and there are breaks in both series (see panel 2 of Table 4), the 2SU1B procedure dominates the other procedures as expected. But, the 2SU1B procedure cannot be recommended in practice if it is unknown whether the data is AO or IO because, under IO data with breaks in both equations, 2SU1B often does much worse than both 1SV1B and 2SU2B. Interestingly, under IO data, 2SU2B has forecasts that are nearly as efficient as the best of the one-step procedures. Thus, the 2SU2B procedure is quite robust when the break dates are unknown and is recommended in practice. Note that 1SU2B generates nearly equivalent forecasts to 2SU2B and is also recommended in practice.

On the other hand, the 1SV2B procedure performs badly when there are breaks in each series or equation. See panels 2 and 4. Things are especially bad when the data is IO. In this case, 1SV2B can deliver much worse forecasts than 1SV1B. This happens because the break dates are estimated sequentially. Consider the cases where each equation of the VAR has exactly one intercept shift ($\Gamma_{12} = \Gamma_{21} = 0$). Because the break dates are different, imposing the constraint that the break dates are the same when estimated the first break date results in a misspecified model and a break estimate that is far from either true break date. Thus, 1SV2B generates break date estimates far from the truth and performs similarly to the 1S0B procedure. More surprisingly is the fact that 1SV2B continues to do badly when $\Gamma_{12} \neq 0$ and/or $\Gamma_{21} \neq 0$ in which case both equations of the VAR have both intercept shifts.

The simulation results for the case of an unknown break date can be summarized as follows. The 1SU2B and 2SU2B procedures are the most robust methods and are recommended. Our results suggest that whether breaks are removed from the data prior to VAR analysis or the breaks are included directly into the VAR, the break dates should be obtained using estimators from the two-step procedures. The 1SV2B procedure is not recommended as it can lead to substantially less precise forecasts.

4 CONCLUSION

In this paper, we focus on two practical implications of omitted mean shifts in a multivariate setting. First, as is well known from simulation studies, omitted mean shifts can cause GC tests to over-reject. We formally show the conditions under which over-rejections occur. We illustrate the over-rejection problem in a bivariate VAR in a simple simulation study. Our second focus is on methods of estimating and modeling mean shifts in the context of forecasting. We consider a two-step method which estimates and removes the mean shifts. The VAR analysis proceeds using the demeaned data. We also consider a one-step method which estimates and includes intercept shifts directly in the VAR. When there is one break per series or equation at different breaks, we find that the most robust methods are those that allow for two breaks per series or equation. Fortunately for practitioners, there does not seem to be a big difference between removing the breaks in the first step or modeling them directly in the VAR. This is true whether the data is of the AO or IO type. However, whether we use the one or two step procedure, we recommend the break dates be estimated in the first step of the two-step procedure, and not directly from the VAR. Our results are based on comparisons of forecasting precision. Whether or not these conclusions continue to hold when other metrics are used awaits further analysis.

This paper should be viewed as a first step in understanding how to model unstable deterministic trends in multivariate models. There are several directions for future research. For example, we have taken the presence of mean shifts as given. In practice, preliminary tests are usually used to determine whether the mean is stable or not. For the two-step procedure, one would ideally want mean shift tests that allow multiple shifts and are robust to serial correlation and a unit root in the errors. Perron[7] and Vogelsang[2,4] robust to serial correlation/unit roots but not designed for multiple shifts. On the other hand, the multiple shift analysis of Bai and Perron[16] require stationarity of the errors. Hence no existing test is fully satisfactory. Much less research has been done on testing for intercept shifts in VARs. Bai et al.[17] is a natural starting point.

Additional topics worth investigation include trend stationary VARs with trend shifts, cointegrated VARs with unstable deterministic trends, and effects of trend breaks on lag length selection.

5 APPENDIX

Proof of Theorem 1

Let $\tilde{Y}_2$, $\widetilde{DU}$, $\tilde{d}$, and $\tilde{e}_2$ be $(T-1)\times 1$ vectors of demeaned y_{2t}, DU_t, d_t, and e_{2t} respectively. Let $\tilde{X}$ be the $(T-1)\times 2$ matrix of demeaned (y_{1t-1}, y_{2t-1}). The AO model can be written as

$$\tilde{y}_{2t} = \delta_2^{*\prime}\widetilde{DU} + \gamma_2^{*\prime}\tilde{d}_t + \tilde{B}_2' X_t + \tilde{e}_{2t} \tag{A1}$$

Using regression (7) and plugging in for $\tilde{Y}_2$ using (A1), we have

$$\begin{aligned}\hat{B}_2 &= (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y}_2 \\ &= B_2 + (\tilde{X}'\tilde{X})^{-1}\tilde{X}'(\widetilde{DU}\delta_2^* + \tilde{d}\gamma_2^* + \tilde{u}_2)\end{aligned}$$

We now consider the following limiting results:

$$\begin{aligned} T^{-1}\tilde{X}'\tilde{X} &= T^{-1}\sum \tilde{X}_t\tilde{X}_t' \\ &= T^{-1}\sum [\delta\tilde{D}U_{t-1} + \tilde{Z}_{t-1}][\delta\tilde{D}U_{t-1} + \tilde{Z}_{t-1}]' \\ &= T^{-1}\delta\Big[\sum \tilde{D}U_{t-1}\tilde{D}U_{t-1}'\Big]\delta' + T^{-1}\sum \tilde{Z}_{t-1}\tilde{Z}_{t-1}' \end{aligned}$$

$$\xrightarrow{p} \delta\Lambda\delta' + \Theta(0) \equiv C.$$

Because C is positive-definite for all values of δ_1 and δ_2, it follows that $(T^{-1}\tilde{X}'\tilde{X})^{-1} \xrightarrow{p} C^{-1}$. Furthermore, it is easy to see

$$T^{-1}\tilde{D}U_{it}{}^2 \xrightarrow{p} \lambda_i(1-\lambda_i),$$

$$T^{-1}\sum \tilde{D}U_{it}\tilde{D}U_{jt} \xrightarrow{p} (1-\lambda_i)\lambda_j$$

if $T_{B_i} < T_{B_j}$, and $(1-\lambda_j)\lambda_i$ otherwise. Thus,

$$\begin{aligned} T^{-1}\tilde{X}'\tilde{D}U\delta_i{}^A &= \Big[T^{-1}\sum X_t\tilde{D}U_t'\Big]\delta_i{}^A \\ &= \Big[T^{-1}\sum (\delta\tilde{D}U_{t-1} + \tilde{Z}_{t-1})\tilde{D}U_t'\Big]\delta_i{}^A \\ &= \Big[T^{-1}\sum \delta\tilde{D}U_{t-1}\tilde{D}U_t\Big]\delta_i{}^A + o_p(1) \xrightarrow{p} \delta\Lambda\delta_i{}^A \end{aligned}$$

where $\Lambda = 1 - \min(\lambda_i, \lambda_j) - (1-\lambda_i)(1-\lambda_j)$. Using these convergence results it directly follows that

$$p\lim(\hat{a}_2 - a_2) = p\lim[(T^{-1}\tilde{X}'\tilde{X})^{-1}T^{-1}\tilde{X}\widetilde{DU}\delta_2^*] = C^{-1}\delta\Lambda\delta_i{}^A$$

This establishes Theorem 1. Corollary 1 follows from the fact that the $n \times n$ matrix Λ becomes $\lambda(1-\lambda)\iota\iota'$ when the break dates coincide. The proof of Theorem 3 follows using similar arguments and is omitted.

Proof of Theorem 5
The F-test for the hypothesis that $b_{21} = 0$ is given by:

$$\mathrm{GC} = \frac{(\hat{b}_{21} - b_{21})}{s^2(\tilde{X}'\tilde{X})_{11}{}^{-1}} = \frac{T(\hat{b}_{21} - b_{21})^2}{s^2(T^{-1}\tilde{X}'\tilde{X})_{11}{}^{-1}} \tag{A2}$$

where X is the $(T-1)\times n$ matrix of values of $X_t = (y_{1t-1}, y_{2t-1})$, $\tilde{X}$ are the demeaned X's and $s^2 = (T-4)^{-1}\sum_{t=2}^{T}(\hat{u}_{2t}^*)^2$.

Consider the simple case when the break dates coincide. When $\delta_2 = 0$ (and $a_{21} = 0$), the AO model reduces to $y_{2t} = \mu_2^* + b_{22}y_{1t-1} + b_{22}y_{2t-1} + u_{2t}$. Therefore regression (7) is correctly specified and standard OLS results apply giving part 1 of the Theorem. When $\delta_2 \neq 0$, we have from Theorem 1 that $(\hat{b}_{21} - b_{21})^2 =$

$O_p(1) > 0$ and $(T^{-1}\tilde{X}'\tilde{X})_{11}{}^{-1} = O_p(1) > 0$. Let $M = I - \tilde{X}(\tilde{X}'\tilde{X})^{-1}\tilde{X}'$. The limit of s^2 follows from

$$\begin{aligned} s^2 = T^{-1}\tilde{Y}_2'M\tilde{Y}_2 &= T^{-1}(\widetilde{DU}\delta_2^* + \tilde{d}\gamma_2^* + \tilde{u}_2)'M(\widetilde{DU}\delta_2^* + \tilde{d}\gamma_2^* + \tilde{u}_2) \\ &= T^{-1}(\widetilde{DU}\delta_2^* + \tilde{u}_2)'M(\widetilde{DU}\delta_2^* + \tilde{e}_2) + o_p(1) \\ &= \delta_2^* T^{-1}\widetilde{DU}'M\widetilde{DU} + T^{-1}\tilde{e}_2'M\tilde{e}_2 + o_p(1) \\ &= (\delta_2^*)^2[T^{-1}\widetilde{DU}'\widetilde{DU} - T^{-1}\widetilde{DU}'\tilde{X}(T^{-1}\tilde{X}'\tilde{X})^{-1}T^{-1}\tilde{X}'\tilde{D}U] \\ &\quad + \sigma_2{}^2 + o_p(1) \\ &\xrightarrow{p} (\delta_2^*)^2[\lambda(1-\lambda) - \lambda(1-\lambda)\delta'C^{-1}\delta\lambda(1-\lambda)] + \sigma_2{}^2 \end{aligned}$$

Because $0 < p\lim(s^2) < \infty$, we have $0 < p\lim(1/s^2) < \infty$. Combining these results gives $T^{-1}\mathrm{GC} = (\hat{b}_{21} - b_{21})^2/[s^2(T^{-1}\tilde{X}'\tilde{X})_{11}{}^{-1}] = O_p(1) > 0$. When $\delta_1 = 0$, $b_{12} = 0$ and $\sigma_{12} = 0$, it is straightforward to show that $T(\hat{a}_{21} - a_{21})^2 = O_p(1)$, and since it is still true that $(T^{-1}\tilde{X}'\tilde{X})_{11}{}^{-1} = O_p(1)$ and $0 < p\lim(1/s^2) < \infty$, we have $\mathrm{GC} = O_p(1)$.

ACKNOWLEDGMENT

Ng acknowledges grants from the Social Science and Humanities Research Council of Canada (SSHRC), the Fonds de la Formation de Chercheurs et l'Aide à la Recherche du Québec (FCAR). Vogelsang thanks the Center for Analytic Economics at Cornell University. We thank the editor and a referee for helpful comments that helped improve the paper.

REFERENCES

1. Perron, P. The Great Crash, the Oil Price Shock and the Unit Root Hypothesis. Econometrica **1989**, *57*, 1361–1401.
2. Vogelsang, T.J. Wald-Type Tests for Detecting Shifts in the Trend Function of a Dynamic Time Series. Econometric Theory **1997**, *13*, 818–849.
3. Chu, C.-S.; White, H. A Direct Test for Changing Trend. Journal of Business and Economic Statistics **1992**, *10*, 289–300.
4. Vogelsang, T.J. Testing for a Shift in Mean Without Having to Estimate Serial Correlation Parameters. Journal of Business and Economic Statistics **1998**, *16*, 73–80.
5. Banerjee, A.; Lumsdaine, R.L.; Stock, J.H. Recursive and Sequential Tests of the Unit Root and Tread Break Hypothesis. Journal of Business and Economic Statistics **1992**, *10*, 271–287.
6. Ben-David, B.; Papell, D.H. The Great Wars, the Great Crash, and Steady State Growth: Some New Evidence about an Old Stylized Fact. Journal of Monetary Economics **1995**, *36*, 453–475.
7. Perron, P. A Test for Changes in a Polynomial Trend Function for a Dynamic Time Series. Manuscript, Princeton University, 1991.

8. Box, G.E.P.; Tiao, G.C. Intervention Analysis With Applications to Economics and Environmental Problems. Journal of the American Statistical Association **1975**, *70*, 70–79.
9. Lutkepohl, H. The Stability Assumption in Tests of Causality Between Money and Income. In *Econometrics of Structural Change*; Kramer, W., Ed.; Springer-Verlag: New York, 1989; 75–86.
10. Bianchi, M. Granger Causality in the Presence of Structural Changes, Mimeo, Bank of England, 1995.
11. Tiao, G.C. Autoregressive Moving Average Models, Intervention Problems and Outlier Detection in Time Series. *Handbook of Statistics 5*; North Holland: Amsterdam, 1985; 85–118.
12. Engle, R.F.; Kozicki, S. Testing for Common Features. Journal of Business and Economic Statistics **1993**, *11*, 369–379.
13. Blanchard, O.; Quah, D. The Dynamic Effects of Aggregate Demand and Supply. American Economic Review **1989**, *79*, 655–673.
14. Gambe, E.N.; Joutz, F. The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment. American Economic Review **1993**, *83*, 1387–1393.
15. Bai, J.S. Least Squares Estimation of a Shift in Linear Process. Journal of Time Series Analysis **1994**, *15*, 453–472.
16. Bai, J.S.; Perron, P. Estimating and Testing Linear Models with Multiple Structural Breaks. Econometrica **1998**, *66*, 47–78.
17. Bai, J.S.; Lumsdaine, R.L.; Stock, J.H. Testing for and Dating Breaks in Stationary and Nonstationary Time Series. Review of Economic Studies **1997**, *65*, 395–432.