COINTEGRATION TESTING UNDER STRUCTURAL BREAKS: A ROBUST EXTENDED ERROR CORRECTION MODEL

Miguel A. Arranz and Alvaro Escribano

I. INTRODUCTION

The properties of the cointegration tests based on single equation error correction models (ECM test) are well known. The dependence of critical values, and the power of the test on nuisance parameters are documented in Banerjee et al. (1986), Engle and Granger (1987), Kremers et al. (1992), Park and Phillips (1988, 1989), and Banerjee et al. (1993).

From the univariate point of view, the effects of having breaks when applying unit root test, like Dickey and Fuller (1979) test, are well known, and Perron (1989) is a good starting point to see those impacts. From Hendry (1996), a structural break essentially corresponds to an intermittent shock with a permanent effect on the series. If this shock is not explicitly taken into account, standard unit root tests would in general mistake the structural break for a unit root. Leybourne et al. (1998) indicate that the opposite can also happen if the break occurs at the beginning of the sample. The results of Hendry and Neale (1990) and Perron and Vogelsang (1992) indicate that a neglected shift in the mean also leads to spurious unit roots. Rappoport and Reichlin (1989) is probably the first reference to deal with the impact of having segmented trends as an alternative to a unit root model, and Andrés et al. (1990) extended the analysis to more that one break point in the trend.

The main drawback with this literature, that has expanded dramatically since then, is that we always have to add dummy variables to capture the structural breaks in order to correctly apply unit root tests. Therefore the critical values obtained depend on the size and on the timing of the break. Again, a vast literature emerged searching for unknown break points using recursive or sequential tests. See, for example, Banerjee et al. (1992), Zivot and Andrews (1992), Andrews (1993). Andrews et al. (1996), Bai (1997),

This paper is based on Arranz and Escribano (1998*b*), presented at the Econometric Society European Meeting in September 1998 and the CEPR Meeting on '*New Approaches to the Business Cycle*' held in Madrid in January 1998. This research has been partially financed with projects CICYT PB95–0298 and TMR ERB-4061 PL 97-0994.

Vogelsang (1997), Bai and Perron (1998), and Banerjee, Lazarova and Urga (1998).

Another class of unusual events are additive outliers. These are events with large, but temporary effects on the series. In certain cases, those effects dominate the remaining information contained in the series and biases unit root inference towards rejection of the unit root hypothesis even if the null hypothesis of a unit root is correct, as reported in Franses and Haldrup (1994) and Lucas (1995*a*,*b*).

With multivariate time series the situation could be worse, since we need to decide on the type of models that generate the anomalous observations (breaking trends, additive outliers, ...) taking into account that those irregularities need not occur simultaneously or on all of the variables. Therefore, the multivariate analysis is generally more difficult but in some cases could be easier as it will be explained later on.

In empirical applications it is more the rule than the exception to include dummy variables in order to obtain parameter constant ECM models. The effects of including dummy variables to capture structural breaks in ECM models have been analyzed in Kremers et al. (1992), and Campos et al. (1996).

The fact that critical values (C.V.) depend on the particular type of dummy variable included is a nuisance when doing applied work. However, we could avoid the use of dummy variables applying robust estimation techniques. This is the approach taken by Lucas (1995*a*,*b*) in the univariate case and by Lucas (1997) and Frances and Lucas (1997*a*,*b*) in a multivariate framework.

In this paper we follow a different route. We want to find robust modeling procedures to test for unit roots in the presence of structural breaks in an ECM context. Instead of including dummy variables in ECM methods, we allow to approximate those breaks by adding extra dynamic terms, as determined by the SBIC criterion, or by including some extra lags of the error correction term (extended ECM model). In particular, we look at the critical values obtained from the overparameterized models, we study the power and the size of the test under different MA(1) errors by Monte Carlo simulations.

The structure of the paper is as follows. In section II we analyze the effects of having deterministic elements (constant terms, deterministic trends, dummy variables, segmented trends, etc.) on alternative specifications of the ECM models, and in particular on the cointegrating errors. Three types of deterministic possibilities are studied in detail: simultaneous co-breaking, co-breaking in levels (not in differences) and co-breaking in differences (not in levels). The Monte Carlo experiments are introduced in Section III, and results of the usual ECM tests are analyzed in detail. Section IV shows the Monte Carlo results based on the extended ECM model. Section V presents some conclusions. The different co-breaking possibilities are analyzed in Appendix A.

II. ERROR CORRECTION MODELS WITH AND WITHOUT SIMULTANEOUS CO-BREAKING

Consider the following conditional error correction model (ECM)

$$\Delta(yt - \mu_{y,t}) = a\Delta(z_t - \mu_{z,t}) + b[(y_{t-1} - \mu_{y,t-1}) - \alpha(z_{t-1} - \mu_{z,t-1})] + u_{1t} \tag{2.1a}$$

$$\Delta(z_t - \mu_{z,t}) = u_{2t} \tag{2.1b}$$

Assume that $\ldots, y_{-1}, \ y_0 = 0$ and $\ldots, z_{-1}, \ z_0 = 0$, let $\mu_{y,t} = E(y_t)$, $\mu_{z,t} = E(z_t)$ be the corresponding unobserved unconditional means, based on the validity of the initial parameterization at t_0. Those means include all possible deterministic elements like: constant terms, deterministic trends, dummy variables, segmented trends, outliers, etc. Define B as the back-shift operator, $B^k y_t = y_{t-k}$, $\Delta = (1 - B)$ is the first differencing operator, and let $(1, -\alpha)$ be the cointegrating vector. The stochastic errors u_{1t} and u_{2t} are jointly, and serially uncorrelated with zero mean, and constant variances $var(u_{1t}) = \sigma_1^2$ and $var(u_{2t}) = \sigma_2^2$. Those conditions of the errors can be relaxed to allow for some serial dependence and joint dependence as long as Δz_t is weakly exogenous for the parameters of the model (2.1a). To make the analysis clear, we keep the same assumptions as those used in our Monte Carlo experiment.

Model (2.1a)–(2.1b) can be written in terms of the observable variables y_t and z_t as follows,

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \tag{2.2a}$$

$$\Delta z_t = \Delta\mu_{z,t} + u_{2t} \tag{2.2b}$$

$$c_t \equiv \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \alpha\mu_{z,t-1}) \tag{2.2c}$$

In this paper we investigate the effects of having alternative models for the intercept c_t on the ECM test for non-cointegration ($b = 0$) of (2.2a).

The cointegrating errors in terms of the observable variables are obtained from (2.2a)–(2.2c)

$$y_t - \alpha z_t = \frac{1}{1-(b+1)B} c_t + \frac{a-\alpha}{1-(b+1)B} \Delta\mu_{z,t} + \frac{a-\alpha}{1-(b+1)B} u_{2,t} + \frac{1}{1-(b+1)B} u_{1,t} \tag{2.3}$$

and we would need to include many deterministic regressors to approximate the first two elements of the RHS of (2.3), c_t and $\Delta\mu_{z,t}$. Therefore, the cointegrating regression requires the following set of regressors

$$y_t = \alpha z_t + \frac{1}{1-(b+1)B} c_t + \frac{a-\alpha}{1-(b+1)B} \Delta\mu_{z,t} + v_t. \qquad (2.4)$$

Notice that the problem of having to add arbitrary and influential deterministic regressors is reduced, but not solved by conditioning on Δz_t, $\Delta z_{t-1} \dots$

$$y_t = \alpha z_t + \frac{a-\alpha}{1-(b+1)B} \Delta z_t + \frac{1}{1-(b+1)B} c_t + w_t \qquad (2.5)$$

where $w_t = 1/[1-(b+1)B]u_{1,t}$, and we still have to approximate the dynamic deterministic effects of c_t.

Definition 2.1. *Given valid initial conditons, let* $E(y_t) = \mu_{y,t}$ *and* $E(z_t) = \mu_{z,t}$, *we say that the time series* y_t *and* z_t *have* co-breaks in differences *if* $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$, *where* c_d *is a finite constant parameter.*

Definition 2.2. *Given valid initial conditions, let* $E(y_t) = \mu_{y,t}$ *and* $E(z_t) = \mu_{z,t}$, *we say that the time series* y_t *and* z_t *have* co-breaks in levels *if* $\mu_{y,t} - \alpha\mu_{z,t} = c_l$, *where* c_l *is a finite constant parameter.*

Several possible intermediate cases are of interest in empirical applications and will be considered in the the simulation experiments later on.

Case 2.1. *Co-break in differences but not in levels.*

Co-break in differences: $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$ implies that $\Delta\mu_{y,t} - \alpha\Delta\mu_{y,t} = (a-\alpha)\Delta\mu_{z,t} + c_d$. From recursive substitution $\mu_{y,t} - \alpha\mu_{z,t} = (\mu_{y0} - \alpha\mu_{z0}) + c_d t + (a-\alpha)\mu_{z,t}$, and c_t becomes

$$c_t = c_d - b(\mu_{y,0} - \alpha\mu_{z,0}) - bc_d(t-1) - b(a-\alpha)\mu_{z_{t-1}} \qquad (2.6)$$

and equation (2.2a) becomes

$$\Delta y_t = c_m + bc_d t - b(a-\alpha)\mu_{z,t-1} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t}, \qquad (2.7)$$

where c_m is a constant equal to $c_m = c_d - b(\mu_{y,0} - \alpha\mu_{z,0}) + bc_d$.

Remark: Assuming that $\mu_{y,0} - \alpha\mu_{z,0} = constant$, co-break in differences $\Rightarrow$ co-break in levels if $a = \alpha$ (COMFAC) and $c_d = 0$.

Case 2.2. *Co-break in levels but not in differences.*

Co-break in levels ($\mu_{y,t} - \alpha\mu_{z,t} = c_l$). Taking first differences, we have $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$. But from equation (2.2c)

$$c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - bc_l = (\alpha - a)\Delta\mu_{z,t} - bc_l, \qquad (2.8)$$

and equation (2.2a) becomes

$$\Delta y_t = -bc_l + (\alpha - a)\Delta\mu_{z,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \qquad (2.9)$$

Remark: Co-break in levels $\Rightarrow$ co-break in differences if $a = \alpha$ (COMFAC restriction).

Definition 2.3 *Given valid initial conditions, let* $E(y_t) = \mu_{y,t}$ *and* $E(z_t) = \mu_{z,t}$, *we say that the time series* y_t *and* z_t *have* simultaneous co-breaks *if* $\Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t} - \alpha\mu_{z,t}) = c_s$, *where* c_s *is a finite constant parameter.*

It is clear that when y_t and z_t have *co-breaks in levels and in differences (full co-break),* this is a particular case of simultaneous co-breaking. In the case of simultaneous co-breaking, the intercept c_t from (2.2c) is constant, $c_t = c$ and the error correction model from (2.2a) becomes the standard conditional ECM model where the only deterministic regressor is the constant term, c.

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \tag{2.10}$$

On the other hand, even if $c_t = c$, the cointegration regression (2.4) has a constant term, say $\tilde{c} = 1/[1 - (b+1)B]c$, and also lags of $\Delta\mu_{z,t}$

$$y_t = \tilde{c} + \alpha z_t + \frac{a - \alpha}{1 - (b+1)B}\Delta\mu_{z,t} + v_t. \tag{2.11}$$

Therefore, we would have to add a complicated dynamic structure of dummy variables when $\Delta\mu_{z,t}$ has structural breaks. This problem is solved now by conditioning on $\Delta z_t, \Delta z_{t-1}, \ldots,$ since (2.11) becomes

$$y_t = \tilde{c} + \alpha z_t + \frac{a - \alpha}{1 - (b+1)B}\Delta z_t + w_t. \tag{2.12}$$

This regression is simplified if there is a common factor restriction, $a = \alpha$ (COMFAC restriction), since then $\Delta\mu_{z,t}$ has no effect on the cointegrating regression (2.11). From equation (2.4), it is clear that to have $a = \alpha$ (COMFAC restriction) is not a universal solution because the cointegration regression takes the form,

$$y_t = \alpha z_t + \frac{1}{1 - (b+1)B}c_t + u_t \tag{2.13}$$

and we have a strange cointegrating regression with a complicated structure through the lagged deterministic elements of c_t.

In general, without having any co-break in levels or in differences, the most parisimonious representation is the conditional ECM model (2.1a), and in terms of observable variables is representation (2.2a), because it only requires to add the deterministic regressors coming from the contemporaneous values of c_t. Clearly, if we are interested in estimating the parameters a, α and b, it is much easier and more parsimonious to estimate them by 1-step procedures (OLS or NLS) in ECM representations (2.1a) or (2.2a) than in any other of the representations discussed. However, to do that in (2.1a) we need to know or to estimate first the unconditional means $\mu_{y,t}$ and $\mu_{z,t}$ and this generally incorporates arbitrary information about unknown events.

Error correction models with simultaneous co-breaking

From equations (2.2a)–(2.2c) and the analysis of Escribano (1987) and Andrés et al. (1990), it is clear that any error correction model in terms of the observable variables should account for the joint effects of the following elements: $\Delta\mu_{y,t}$, $\Delta\mu_{z,t}$, $\mu_{y,t-1}$ and $\mu_{z,t-1}$.

Previous error correction models with co-breaks have been treated in Campos et al. (1996) and Hendry (1996). In this section we consider deterministic segmented trends in y_t and z_t that have *simultaneous co-breaks* (see Definition 2.3), where $c_s = 0$. The segmented trend in z_t is generated by $\Delta\mu_{z,t} = sD_{j,t}$, where s is a parameter that measures the size of the break, and $D_{j,t}$ is a dummy variable that takes the value 0 before the break and the value 1 at the break and after the break, see section 3.1 for details. In this case, (2.2a) and (2.2b) can be simplified to

$$\Delta y_t = a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.14a)$$

$$\Delta z_t = \Delta\mu_{z,t} + u_{2,t} \quad (2.14b)$$

$$\Delta\mu_{z,t} = sD_{j,t} \quad (2.14c)$$

where (2.14a) has the form of the usual single equation error correction without a constant term since $c_t = 0$.

Consider the DGP given by (2.14a) and (2.14b) with $\Delta\mu_{z,t} = 0$. The distribution of the t-ratio of the parameter b under the null hypothesis that $b = 0$ (no cointegration) was analyzed by Banerjee et al. (1986), Kremers et al. (1992), Park and Phillips (1988, 1989) and Banerjee et al. (1993).

Assuming that α is known and equal to 1, $\alpha = 1$ one could consistently estimate the parameters of equation (2.14a) by OLS,

$$\Delta y_t = \hat{a}\Delta z_t + \hat{b}(y_{t-1} - z_{t-1}) + \hat{u}_{1,t} \quad (2.15)$$

and obtain the asymptotic distribution of the t-ratio of $\hat{b}$ under $H_0 : b = 0$. $t_{\hat{b}} = (\hat{b}/\hat{\sigma}_{\hat{b}})$ is the ECM test and

$$t_{\hat{b}} \Rightarrow \frac{(a-1)\int B_{u_2} dB_{u_1} + r^{-1}\int B_{u_1} dB_{u_1}}{\sqrt{(a-1)^2\int B_{u_2}^2 + 2(a-1)r^{-1}\int B_{u_2}B_{u_1} + r^{-2}\int B_{u_1}^2}} \quad (2.16)$$

where '$\Rightarrow$' denotes weak convergence, B_{u_1} and B_{u_2} are independent Brownian motions and $r = \sigma_1/\sigma_2$. In terms of the 'signal to noise ratio', $a = -(a-1)r = [-(a-1)\sigma_1]/\sigma_2$,

$$t_{\hat{b}} \Rightarrow \frac{\int B_{u_2} dB_{u_1} + \frac{1}{q}\int dB_{u_1}\,\mathrm{d}B_{u_1}}{\sqrt{\int B_{u_2}^2 + 2\left(\frac{1}{q}\right)\int B_{u_2} b_{u_1} + \left(\frac{1}{q}\right)^2 \int B_{u_1}^2}}. \tag{2.17}$$

Notice that when $q = 0$ (or $a = 1$, COMFAC restriction) (2.16) is reduced to

$$t_{\hat{b}} \Rightarrow \frac{\int B_{u_1} dB_{u_1}}{\sqrt{\int B_{u_1}^2}} \equiv DF \tag{2.18}$$

where DF is the Dickey–Fuller distribution, (see Dickey and Fuller, 1979), of the t-ratio of $\hat{b}$ from the OLS regression (2.20), which is the non-cointegration D–F test of Engle and Granger (1987) when the cointegration vector is known. From (2.17) and for large q

$$t_{\hat{b}} \Rightarrow \frac{\int B_{u_2} dB_{u_1}}{\sqrt{\int B_{u_2}^2}} + O_p(q^{-1}) \tag{2.19}$$

Since B_{u_1} and B_{u_2} are independent Brownian motions, the leading term in the right hand side follows a standard Normal distribution (Park and Phillips, 1988).

When $a = \alpha$, and α is known, $\alpha = 1$, equation (2.14a) can be written as

$$\Delta(y_t - z_t) = b(y_{t-1} - z_{t-1}) + u_{1t} \tag{2.20}$$

which is a standard Dickey–Fuller equation. If we estimate the unrestricted equation (2.14a) with $\alpha = 1$, then the t-ratio $t_{\hat{b}} \to DF$ distribution, see equation (2.13).

When $\Delta\mu_{zt} = 0$, the distribution of the t-ratio of the parameter b in (2.15) under a local alternative hypothesis, $b = h/T$, $h < 0$ was derived by Kremers et al. (1992) following Phillips (1987),

$$t_{\hat{b}} \Rightarrow h(1+q^2)^{1/2}\left(\int K_e^2\right)^{1/2}$$
$$+ \frac{(a-1)\int K_{u_2} dB_{u_1} + r^{-1}\int K_{u_1} dB_{u_1}}{\sqrt{(a-1)^2\int K_{u_2}^2 + 2(a-1)r^{-1}\int K_{u_2}K_{u_1} + r^{-2}\int K_{u_1}^2}} \tag{2.21}$$

where $e_t = (a-1)\Delta z_t + u_{1t}$ and K_e is a diffusion process. Notice that for

$h = 0$, $K_i = B_i$ (a Brownian motion), and (2.21) coincides with (2.16). Therefore, the power of $t_{\hat{b}}$ should increase with h for a given T, or increase with the absolute value of b, since the distribution of $t_{\hat{b}}$ under H_1 is shifted to the left of $t_{\hat{b}}$ under H_0.

When $a = 1$, COMFAC restriction, the t-ratio of b in (2.15)

$$t_{\hat{b}} \Rightarrow c\left(\int K_e^2\right)^{1/2} + \frac{\int K_{u_1} dB_{u_1}}{\sqrt{\int K_{u_1}^2}} \tag{2.22}$$

which is the distribution of the DF statistic under the local alternative.

For $a \neq 1$ and a large q, (2.21) is approximately Normal conditional on u_{2t},

$$t_{\hat{b}} \Rightarrow N\left[(c(1+q^2))^{1/2}\left(\int K_{u_2}^2\right)^{1/2}, 1\right] + O_p(q^{-1}) \tag{2.23}$$

The unconditional mean of $t_{\hat{b}}$ is approximately $E(t_{\hat{b}}) \approx c(1 + q)^2)^{1/2}(1/\sqrt{2})$. Therefore, increasing a, since c is negative, the distribution of $t_{\hat{b}}$ under the local alternative can arbitrarily be shifted towards the left and hence the power of the test can be made arbitrarily close to 100%.

However, in small samples the power of $t_{\hat{b}}$ in (2.14a) with $\alpha = 1$ can be lower than the power of $t_{\hat{b}}$ in (2.20) since the estimation of the unrestricted model (2.14a) is less efficient, therefore generating smaller t-ratios than the asymptotic ones in absolute values.

From equations (2.14a)–(2.14c) it is clear that $y_t \sim I(1)$, and $z_t \sim I(1)$ with segmented trends, and they are cointegrated with cointegration vector equal to $(1, -\alpha)$. Furthermore, the segmented trend in the 'exogenous' variable z_t simultaneously co-breaks ($c_s = 0$) with the endogenous variables y_t and therefore the cointegrating relationship can explicitly be written as in equation (2.5) with $c_t = 0$. Therefore, from model (2.14a)–(2.14c), it is clear that if $-2 < b < 0$[1] the variables are cointegrated, that is, $(y_t - \alpha z_t) \sim I(0)$ but with nonstationary errors due to the structural breaks in Δz_t, and if $b = 0$, the variables are not cointegrated, $(y_t - \alpha z_t) \sim I(1)$ with segmented trends.

Substituting equations (2.14b) and (2.14c) in (2.3) and under simultaneous co-breaking with $c_t = 0$, we get an interesting relationship expression for the cointegrating errors,

[1]It is not uncommon to find the cointegration condition to be $-1 < b < 0$. See for example Kremers et al. (1992) and Campos et al. (1996).

$$y_t - \alpha z_t = \frac{a-\alpha}{1-(b+1)B} s D_{j,t} + \frac{a-\alpha}{1-(b+1)B} u_{2t} + \frac{1}{1-(b+1)B} u_{1t}. \tag{2.24}$$

It is clear that to estimate the cointegrating parameter α in (2.24) we need to include lags of the dummy variables $D_{j,t}$ unless the common factor restriction, $a = \alpha$, holds. Hence in general it is better to estimate (2.5) without using any dummy variable. These conclusions are valid for most cointegrating static regression models that use parametric or non–parametric procedures to estimate α in the cointegrating regression $y_t = \alpha z_t + \epsilon_t$, such as OLS (Engle and Granger, 1987), FM-OLS (Phillips and Hansen, 1990) or canonical cointegration (Park, 1992). The solution is simple when $\Delta\mu_{z,t}$ is a constant or a trend, but it is not that simple, see equation (2.24) when $\Delta\mu_{z,t}$ has level shifts, segmented trends or other types of unknown structural breaks that occur with economic data.

In the simultaneous co-breaking case, there is a clear advantage with the conditional dynamic model (2.14a) because the cointegrating parameter can efficiently be estimated by OLS without needing to use any deterministic explanatory variables.

In summary, dynamic conditional error correction models based on economic variables that have simultaneous co-breaks do not require the use of dummy variables. On the other hand, static cointegrating regressions need to explicitly incorporate dummy variables when the contemporaneous short run parameters differ from the cointegrating parameter $a \neq \alpha$ (no COMFAC). Notice that $a \neq \alpha$ is the most common case in empirical applications.

Error correction models without simultaneous co-breaking

In the previous section we have discussed the of ECM formulations in the presence of simultaneous co-breaking. Our purpose now is to discuss several interesting alternative cases.

Case 2.1. *Co-breaking in differences, but not co-breaking in levels.*

From equation (2.7), (2.14b) and (2.14c), we have that

$$\Delta y_t = c_m + bc_d t - b(a-\alpha)\mu_{z,t-1} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \tag{2.25a}$$

$$\mu_{z,t} = \mu_{z,0} + s\sum_{i=1}^{t} D_{j,i} \tag{2.25b}$$

$$c_m = c_d - b(\mu_{y0} - \alpha\mu_{z0}) + bc_d \tag{2.25c}$$

Therefore, depending on the type of dummy variable, $D_{j,t}$ we could have segmented trends with one or several breaking points in $\mu_{z,t}$ see section 3. Under $H_0: b = 0$, equation (2.25a) becomes

$$\Delta y_t = c_d + a\Delta z_t + u_{1,t}. \tag{2.26}$$

Cases 2.2. *Co-breaking in levels, but not co-breaking in differences.*
From equations (2.9), (2.14b) and (2.14c) we have

$$\Delta y_t = -bc_l + (\alpha - a)sD_{j,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \tag{2.27}$$

Therefore the breaks in the marginal process of Δz_t, (2.14b), affects the error correction model unless the COMFAC restriction is satisfied $(a - \alpha)$. Under $H_0 : b = 0$, equation (2.27) becomes

$$\Delta y_t = (\alpha - a)sD_{j,t} + a\Delta z_t + u_{1t} \tag{2.28}$$

Case 2.3. *No co-breaking in differences nor in levels.*
This likely empirical situation is the result of joining equations (2.27) and (2.25a)–(2.25c) and it is a particular case of the general equation (2.2a).

However, for the purpose of the next section it is enough to show that for cases (2.27) anbd (2.25a) the situations one has to face in practice are complicated in the presence of structural breaks. In particular we analyze the impact of having different structural breaks in terms of the empirical critical values (C.V.) and on the size and the power of the ECM test for $b = 0$ obtained from (2.10).

III. MONTE CARLO SIMULATION EXPERIMENT

The data generating process (DGP) is based on several extensions of the one used by Kremers et al. (1992) and Campos et al. (1996). It is a linear first–order vector autoregression with Normal disturbances, Granger causality in one direction $(z \longrightarrow y)$, and structural breaks in the strongly exogenous variables (Δz_t) for the parameters of interest, a, b, and α.

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \tag{3.1a}$$

$$\Delta z_t = \Delta\mu_{z,t} + u_{2t} \tag{3.1b}$$

$$c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \alpha\mu_{z,t-1}) \tag{3.1c}$$

$$\Delta\mu_{z,t} = sD_{j,t} \tag{3.1d}$$

where

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim IIN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}\right). \tag{3.1e}$$

We alllow three kinds of dummy variables $(D_{j,t}, j = 1, 2, 3)$ in order to simulate either a *single break* in the deterministic trend (segmented trends), at two *different break points* (T/4 or T/2 where T is the sample size),

$$D_{1t} = \begin{cases} 1 & t \geqslant \mathrm{T}/4 \\ 0 & \text{otherwise} \end{cases} \qquad D_{2t} = \begin{cases} 1 & t \geqslant T/2 \\ 0 & \text{otherwise} \end{cases}$$

and a *double break* at points $T/4$ and $3T/4$,

$$D_{3t} = \begin{cases} 1 & T/4 \leqslant t \leqslant 3T/4 \\ 0 & \text{otherwise.} \end{cases}$$

Without loss of generality, we take $\sigma_1^2 = 1$, and $\alpha = 1$. Thus, the experimental design variables are the parameters a, b, s where $\sigma_2 = s$, and the sample size T.

The experiment is a full factorial design with:

$a = 0.0.0.5, 1$ (contemporaneous correlation in first differences)
$b = 0.0$ (no integration), -0.05, -0.1, -0.25, -0.5, -0.75 (cointegration)
$s = 1, 6, 16$ (size of the breaks)
$T = 25, 50, 100, 200, 500, 1000$ (sample size)[2]

and allowing the possibilities of no breaks (NO), single breaks at T/4 or T/2 (D_1 and D_2) or a double-break at T/4 and 3T/4 (D_3) where all of the breaks considered are jumps in the slope of $\mu_{z,t}$ of size s. This represents 216 experiments for each value of b. Remember that when $a = 1$ there is a common (COMFAC) restriction in the error correction model since $\alpha = 1$.

The Monte Carlo experiments are based on 2000 replications of each experiment where the first 50 observations of the simulated series are dropped to consider random initial conditions.

To obtain the empirical critical values we generated the y_t and z_t series following the DGP (3.1a)–(3.1e) under $H_0: b = 0$ and we estimated the equations

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \tag{3.2}$$

$$\phi(B)\Delta y_t = c + \theta(B)\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \tag{3.3}$$

where we have imposed $\alpha = 1$. The orders of the polynomials $\phi(B)$ and $\theta(B)$ are chosen following the SBIC criterion.[3] The lower 5% tail of the empirical distribution of the t-ratio, $t(\hat{b})$ statistic under H_0 is the critical value. The empirical size of the test is analyzed by adding a MA(1) to the errors $u_{1,t}$, i.e., $u_{1,t} + \theta\mu_{1,t-1} = v_t$ with parameter values ($\theta = \pm 0.5$). The empirical power of the test is calculated analogously by simulating the DGP (3.1a)–(3.1e) under $H_1 : b \neq 0$, and computing the percentages of rejec-

[2]Full set of tables are available in Arranz and Escribano (1998*b*).

[3]In order to choose the maximum number of lags of each variable included in the regression, we allow from 0 to 10 lags of each variable and search all possible 121 possibilities, except in the case $T = 25$, in which we allow no more than 4 lags of each variable.

tions obtained from models (3.2) and (3.3) using the previous empirical critical values (size-corrected critical values).

Simultaneous co-breaking is imposed by making $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t} - \alpha\mu_{z,t}) = 0$. In order to get only co-breaks in differences, we impose that $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d = 0.5$. On the other hand, to simulate a set of series with only co-breaks in levels, we impose $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$, see Appendix for a detailed derivation.

Notice that if we had set $c_d = 0$, the critical values would be the ones obtained for the simultaneous co-breaking case, since under $H_0: b = 0$ we would have simultaneous co-breaking with $c_s = 0$. Furthermore, when $a = \alpha = 1$, co-breaks in differences would imply co-breaks in levels (full co-breaking). Under the COMFAC restriction, co-breakinbg in levels imply co-breaks in differences (full co-break), see Appendix.

Monte Carlo simulation experiment: ECM with simultaneous co-breaking
Critical values of the ECM test with simultaneous co-breaking.
The 5 percent critical values from the left tail of the empirical distribution of the t-ratio of $\hat{b}$ are given in Table 1.1. Table 1.1 is generated by making σ_2^2, see (3.1e), equal to the jump size (s) for $s = 1, 6, 16$. We also fixed the variance of $u_{2,t}(\sigma_2^2 = 1)$ and changed $s = 1$, 6, and 16 to make the jump in Δz_t to be more pronounced, and the obtained empirical C.V. were very similar.[4]

Several comments are worth making:

Remark 1. When there is no COMFAC ($a \neq 1$), the distribution of the t-ratio ($t_{\hat{b}}$) is shifted to the right as the jump size (s) increases. Therefore, the large the jump in Δz_t, the more likely that we under-reject the null hypothesis of non-cointegration with the usual C.V. for non-cointegration tests (too many unit roots in the cointegrating errors). However, those shifts of the distribution are not very pronounced since the critical values are similar for different types of jumps (D_1, D_2, and D_3). For example, for $T = 100$ and $a = 0.5$ the 5 percent C.V.'s are between -1.6 ($s = 16$) and -2.2 ($s = 1$). For larger samples ($T = 1000$), the 5 percent C.V.'s are stable around -1.7. Therefore, increasing s is like increasing q in (2.17) and $t_{\hat{b}}$ is approaching the standard normal distribution (2.19).

The main impact on the empirical distribution is obtained while changing the short-run parameter $a(a = 0, 0.5, 1)$. This is not a surprise as it is clear from equations (2.16)–(2.18) where the limiting distribution is given for different parameter values. Equation (2.16) makes explicit that the asymptotic critical values of the t-ratio, $t_{\hat{b}}$, depend on the short run parameter a.

Remark 2. When $a = 1$ (COMFAC restriction) $q = -(a - 1)\sigma_1/\sigma_2$ is zero and the limit distribution coincides with the one obtained by Dickey–Fuller,

[4]Results are available upon request.

see equation (2.18). Therefore, one should expect that going from $a = 0$ to $a = 1$, the empirical distribution be shifted to the left, creating higher critical values in absolute values, see Table 1.1.[5] When there is a COMFAC ($a = 1$) restriction, the C.V. obtained are those of the DF distribution when the DGP is a pure random walk and the regression is estimated with a constant term. Notice that the C.V.'s. do not depend the break size ($s = 1, 6, 16$) or on the type of the break (D_1, D_2, D_3) either. The 5 percent C.V. is around -3.0 for $T = 25$ and for $T = 100$ and $T = 1000$ is around -2.9, which is a good result for empirical applications, since we are not imposing the COMFAC restriction in the regression test.

Remark 3. We also obtained the Critical Values allowing for uncertainty in the dynamics of the model, unknown lags. The maximum lags were estimated by using the SBIC order selection criterion. Our conclusions remain unchanged for sample sizes 100 and 1000. However, as expected, for $T = 25$ the distribution is shifted towards the left with the 5 percent C.V. between -3.2 and -4.9.

Empirical Size of the ECM test with Simulatneous Co-breaking.
In order to assess the validity of our test we analyzed the empirical size of the test by using the previous critical values obtained under the null but with longer dynamics generated by a MA(1) process on the errors. In particular, we simulated our data with the following DGP: $\Delta y_t = c_t + a\Delta z_t + u_{1t} + \theta u_{1,t-1}$, where θ is equal to 0.5 and -0.5. We found a dramatic size distortion depending on the sign of θ when we do not include the relevant dynamic terms, see model (3.2). For positive MA(1) parameters, $\theta = 0.5$, the largest size distortions are for $a \neq 1$, and for negative MA(1) parameters, $\theta = -0.5$, the largest distortions are for $a = 1$. This problem is mitigated if we add dynamic terms, model (3.3), as selected by SBIC criterion, but the size of the test still depends on the sign of θ, see Arranz and Escribano (1998*b*). Therefore, the empirical critical values obtained are not reliable under MA(1) errors and especially when $\theta = -0.5$. More reliable critical values could be obtained by using bootstrap techniques, see Arranz and Escribano (1998*a*), but this is out of the scope of this paper.

Power of the ECM test with Simulatneous Co-breaking.
The power of the ECM-test ($t_{\hat{b}}$) is analyzed by generating data from the DGP under H_1 for values of b that satisfy $-2 < b < 0$. Several parameter values for b are considered, $b = -0.05, -0.1, -0.25, -0.5$, and -0.75. Remember from equation (2.4) that the cointegrating error has an autoregressive representation that depends on the parameter b. If we call ρ_1 the first order autoregressive parameter, $\rho_1 = b + 1$, then $\rho_1 = 0.95, 0.9, 0.75,$

[5]Notice that when $a = 0$ the critical values for NO, D_1, D_2, D_3 do not coincide since in the regression under H_1 we are including Δz_t, which depends on $D_{j,t}$ and that marginally affects the critical values obtained.

TABLE 1: CRITIAL VALUES OF $t(\hat{b})$

Estimated model $\Delta y_t = \hat{c} + \hat{a}\Delta z_t + \hat{b}(y_{t-1} - z_{t-1}) + u_t$

Simultaneous Cobreaking ($c_t = 0$).

		$a = 0$			$a = 0.5$			$a = 1$		
T	DUM	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$
25	NO	−2.678	−1.912	−1.814	−2.831	−2.134	−1.877	−3.027	−2.994	−2.991
	D_1	−2.118	−1.833	−1.716	−2.648	−1.828	−1.801	−3.026	−3.088	−2.985
	D_2	−2.284	−1.828	−1.813	−2.723	−2.010	−1.817	−3.069	−2.950	−3.095
	D_3	−2.213	−1.818	−1.760	−2.662	−1.835	−1.720	−2.976	−3.071	−3.012
100	NO	−2.590	−1.921	−1.795	−2.811	−2.156	−1.796	−2.854	−2.852	−2.898
	D_1	−1.187	−1.649	−1.633	−2.048	−1.632	−1.649	−2.892	−2.996	−2.925
	D_2	−1.942	−1.723	−1.700	−2.268	−1.772	−1.801	−2.907	−2.954	−2.956
	D_3	−1.795	−1.738	−1.641	−2.252	−1.788	−1.730	−2.954	−2.937	−2.889
1000	NO	−2.632	−1.880	−1.745	−2.797	−2.150	−1.856	−2.794	−2.872	−2.857
	D_1	−1.696	−1.730	−1.653	−1.759	−1.609	−1.700	−2.901	−2.903	−2.984
	D_2	−1.741	−1.639	−1.577	−1.771	−1.651	−1.735	−2.887	−2.887	−2.908
	D_3	−1.722	−1.740	−1.602	−1.854	−1.696	−1.654	−2.906	−2.983	−2.900

Cobreaking in Differences, Not in Levels ($\Delta\mu_{y,t} - a\Delta\mu_{z,t} = 0.5$).

		$a = 0$			$a = 0.5$			$a = 1$		
T	DUM	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$
25	NO	−2.447	−1.913	−1.825	−2.483	−2.046	−1.882	−2.560	−2.391	−2.480
	D_1	−2.478	−1.837	−1.713	−2.836	−1.850	−1.804	−2.460	−2.430	−2.469
	D_2	−2.551	−1.824	−1.807	−2.879	−2.051	−1.836	−2.581	−2.582	−2.493
	D_3	−2.550	−1.848	−1.754	−2.802	−1.934	−1.714	−2.538	−2.488	−2.410

100	NO	−2.012	−1.900	−1.824	−2.060	−2.066	−1.851	−2.039	−2.034	−1.964
	D_1	−2.141	−1.649	−1.632	−2.662	−1.668	−1.644	−2.023	−2.060	−1.994
	D_2	−2.189	−1.742	−1.677	−2.369	−1.814	−1.779	−1.997	−2.040	−2.053
	D_3	−2.292	−1.739	−1.639	−2.567	−1.833	−1.737	−1.954	−2.018	−2.022
1000	NO	−1.796	−1.681	−1.757	−1.685	−1.691	−1.793	−1.664	−1.711	−1.775
	D_1	−1.777	−1.737	−1.659	−1.969	−1.637	−1.689	−1.661	−1.735	−1.747
	D_2	−1.965	−1.652	−1.588	−1.862	−1.683	−1.744	−1.741	−1.769	−1.746
	D_3	−1.833	−1.754	−1.598	−1.911	−1.678	−1.645	−1.802	−1.704	−1.736

Cobreaking in Levels, Not in Differences $\mu_{y,t} - \mu_{z,t} = 0$).

		$a = 0$			$a = 0.5$			$a = 1$		
T	*DUM*	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$
25	NO	−2.768	−1.912	−1.814	−2.831	−2.134	−1.877	−3.027	−2.994	−2.991
	D_1	−3.397	−5.771	−6.252	−3.163	−4.458	−6.466	−3.026	−3.088	−2.985
	D_2	−3.694	−6.894	−7.923	−3.375	−5.662	−7.692	−3.069	−2.950	−3.095
	D_3	−3.299	−4.424	−4.583	−3.063	−4.044	−4.575	−2.976	−3.071	−3.012
100	NO	−2.590	−1.921	−1.795	−2.811	−2.156	−1.796	−2.854	−2.852	−2.898
	D_1	−4.704	−11.035	−12.270	−3.558	−8.148	−11.497	−2.892	−2.996	−2.925
	D_2	−4.998	−13.150	−15.340	−3.725	−9.686	−13.783	−2.907	−2.954	−2.956
	D_3	−4.096	−7.876	−8.296	−3.391	−6.668	−8.135	−2.954	−2.937	−2.889
1000	NO	−2.632	−1.880	−1.745	−2.797	−2.150	−1.856	−2.794	−2.872	−2.857
	D_1	−10.546	−32.748	−37.437	−6.432	−23.787	−34.546	−2.091	−2.903	−2.984
	D_2	−12.429	−38.894	−43.772	−7.134	−28.103	−42.559	−2.887	−2.887	−2.908
	D_3	−9.038	−23.042	−23.910	−5.894	−18.265	−23.635	−2.906	−2.983	−2.900

The DGP is generated under $H_0 : \Delta y_t = c_t + a\Delta z_t + u_{1_t}, \Delta z_t = sD_{j_t} + u_{2_t}$, where $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t}, \sigma_1^2 = var(u_{1_t}) = 1$, and $\sigma_2^2 = var(u_{2_t}) = s^2$.

0.5, and 0.25, corresponding to the previous values of the parameter b. Therefore, for $b = -0.05$ we would expect to have low power against stationary alternatives, since the AR(1) parameter, 0.95, is close to the unit root. This intuition can be explained by using (2.21). When b increases is like increasing h relative to T. Therefore, from equation (2.22) we should expect to obtain a reduction in power when the COMFAC restriction ($a = 1$) is satisfied.

In general, under simultaneous co-breaking, the size-adjusted power of the test given in Table 2.1 is high for all possible jump sizes ($s = 1, 6, 16$) and for all sample sizes. The lowest power of the ECM test occurs for values of $a = 1$ (COMFAC restriction) and especially for small sample sizes ($T = 25$) with small absolute values of b. Remember that $a = 1$ corresponds to $q = 0$, see equation (2.18), and in that case the limiting distribution of the test is the Dickey and Fuller distribution. This fact motivated Kremers et al. (1992), Hansen (1995), and Banerjee, Dolado and Mestre (1998) to suggest the addition of variables like Δz_t in the test regression equation to increase the power of the test for non-cointegration, $b = 0$, or for a unit in the univariate context.

In summary, under structural simultaneous co-breaks, the approach based on testing for non-cointegration ($t_{\hat{b}}$) in an error correction model is remarkably robust when there is no COMFAC ($a \neq 1$) and when σ_2 is large relative to σ_1. When $a = 1$ the power is low for $T = 25$ and 100 and for values $-0.5 < b \leqslant 0$, but increases when $T = 1000$. Similar results are obtained for $a \neq 1$ and $s = 1$. Yet, the power increases with the sample size and the size of the breaks, $s = 6, 16$, see Table 2.1. The problem remains when the variables are not co-breaking in levels and/or in differences, and the analysis of this question is the main purpose of the following section.

Monte Carlo simulation experiment: ECM without simultaneous co-breaking.
Critical Values of the ECM test.
Critical values are obtained for different breaks, when we ignore that those breaks have occurred and we run the ECM-test on the usual error correction equation (2.2a) assuming that c_t is constant (misspecified model).

Remark 1. *Co-breaks in differences but not in levels*: In this case, under the null of $b = 0$ there is a constant term, c_d, in the DGP, see equation (2.26), and therefore, since in the estimation equations (3.2)–(3.3) there is no trend, the t-ratio, $t_{\hat{b}}$, has a Normal limiting distribution, see Table 1.2. Notice that for $T = 1000$, the 5 percent C.V. is stable around -1.7, and robust to s, the type of break (D_1, D_2, D_3), and the value of the parameter a.

Remark 2. *Co-breaks in levels but not in differences:* When $a = 1$ co-breaking in levels implies co-breaking in differences, and therefore the comments made for simultaneous co-break case apply. However, the results change dramatically for other values of a, $a \neq 1$. Under $H_0: b = 0$ the dummies, D_1, D_2 and D_3 affect the DGP, see equation (2.28) and therfore the C.V.'s are very unstable. For example, when $T = 1000$, $a = 0.5$ and D_3, the C.V. is -5.9 when $s = 1$, and -23.6 when $s = 16$. Therefore, this type of misspecification creates the most unstable critical values.

This result complicates testing for cointegration from an empirical point of view since we have to generate C.V.'s for every particular break and for every particular jump size (s), as well as for any value of a.

Remark 3. Our comments made on the C.V. remain valid even when we try to approximate the misspecification of the break by adding dynamic terms to the model, see equation (3.3). This conclusion based in misspecified ECM models make the empirical analysis even more dependable on the use of correct C.V.'s, which in fact depend on the particular type of level shift that occurred.

Empirical Size of the ECM test without Simultaneous Co-breaking

We obtained that the ECM test has wrong empirical size under dynamic misspecification, especially when there is no co-break in levels, and not co-breaking in differences and the parameter of the MA(1) is $\theta = -0.5$, see Arranz and Escribano (1998*b*). However, those negative conclusions are tempered when we add dynamic terms because now the size distortion is not affected so much be the sign of the MA(1) parameter, θ, although it is far from the desired level of 5 percent.

Power of the ECM test without Simultaneous Co-breaking

Table 2.2 presents the results of the power of the ECM-test ($t_{\hat{b}}$) when there is only co-breaking in differences but not in levels and we ignore them by proceeding as if no breaks occurred in the dynamic ECM model. Similar situations could be analyzed by introducing dummy variables for the breaks at the wrong unknown date. The results indicate that the ECM-test based on an equation that misspecifies the deterministic breaks in levels has no power for any parameter value analyzed or for any sample sizes considered.

Higher power of the ECM test is obtained for the alternative extreme case. Consider that there is only co-breaking in levels but not in differences.[6] Then, the size-adjusted power of the test is higher than before, but it is still low when there is a break for $a \neq 1$ and if $-0.5 < b < 0$, see Table 2.3. In this case, the size-adjusted power of the test is highly affected by the jump size. The power is reasonably good for parameter values of b larger or equal to 0.5 in absolute value. Our conclusions remain when included extra dynamic terms in our model, see equation (3.3).

In summary, these results are not satisfactory for applied work since we are never certain whether there is no cointegration or whether there is

[6]Remember that when $a = 1$ co-breaking in levels implies co-breaking in differences, and that is the reason why the power of the ECM test improves so much in Table 2.3.

TABLE 2: SIZE ADJUSTMENT POWER OF TEST

Estimated model $\Delta y_t = \hat{c} + \hat{a}\Delta z_t + \hat{b}(y_{t-1} - z_{t-1}) + u_t.$

Simultaneous cobreaking

			NO					D_3				
T	a	b	−0.05	−0.1	−0.25	−0.5	−0.75	−0.05	−0.1	−0.25	−0.5	−0.75
25	0.5	$s=1$	7.70	10.70	26.05	69.40	95.20	11.20	17.55	43.10	82.10	98.25
		$s=6$	32.55	64.55	98.60	100.00	100.00	79.65	96.35	99.90	100.00	100.00
		$s=16$	83.30	98.70	100.00	100.00	100.00	99.25	100.00	100.00	100.00	100.00
	1.0	$s=1$	4.65	7.45	14.85	48.05	83.00	6.70	8.90	16.80	50.20	86.00
		$s=6$	4.90	8.00	14.25	48.45	86.05	5.30	7.70	14.10	48.15	82.60
		$s=16$	4.75	7.70	15.80	48.55	84.75	6.10	7.75	16.05	48.55	85.15
100	0.5	$s=1$	18.35	47.45	99.60	100.00	100.00	71.50	95.40	100.00	100.00	100.00
		$s=6$	95.25	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
		$s=16$	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	1.0	$s=1$	13.90	35.85	97.65	100.00	100.00	12.80	31.55	96.05	100.00	100.00
		$s=6$	13.40	33.80	97.70	100.00	100.00	12.40	34.15	96.25	100.00	100.00
		$s=16$	11.65	31.20	97.45	100.00	100.00	14.05	34.15	97.75	100.00	100.00

Cobreaking in Differences, not in Levels

			NO					D_3				
T	a	b	−0.05	−0.1	−0.25	−0.5	−0.75	−0.05	−0.1	−0.25	−0.5	−0.75
25	0.5	$s=1$	4.80	3.55	1.60	1.35	0.95	5.90	7.70	11.20	17.35	28.55
		$s=6$	28.00	48.50	78.80	87.75	91.60	12/50	22.65	31.50	30.90	26.20
		$s=16$	80.20	97.20	100.00	100.00	100.00	32.65	40.05	39.85	34.00	29.85
	1.0	$s=1$	3.50	2.60	1.05	0.20	0.05	3.50	2.50	0.60	0.15	0.00
		$s=6$	6.10	3.80	1.20	0.10	0.10	4.60	3.50	0.35	0.45	0.05
		$s=16$	3.85	2.65	1.00	0.20	0.25	4.75	3.30	1.10	0.15	0.10

100	0.5	$s=1$	1.50	1.05	0.20	0.10	0.00	3.70	2.85	1.45	1.75	1.45
		$s=6$	41.80	50.05	53.75	51.60	48.90	26.90	31.40	29.80	20.00	15.50
		$s=16$	99.20	99.75	99.80	99.70	99.80	43.05	41.95	34.55	25.75	18.55
	1.0	$s=1$	0.75	0.10	0.00	0.00	0.00	1.25	0.20	0.00	0.00	0.00
		$s=6$	1.20	0.00	0.00	0.00	0.00	1.55	0.05	0.00	0.00	0.00
		$s=16$	1.60	0.20	0.00	0.00	0.00	0.85	0.05	0.00	0.00	0.00
1000	0.5	$s=1$	0.15	0.20	0.00	0.10	0.00	0.45	0.30	0.10	0.00	0.00
		$s=6$	39.10	37.55	33.85	23.15	17.40	28.60	28.15	22.35	16.90	10.45
		$s=16$	69.50	66.45	61.80	54.55	49.40	39.65	36.65	28.70	19.25	11.70
	1.0	$s=1$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		$s=6$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		$s=16$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Cobreaking in Levels, Not in Differences

			NO					*D*$_3$				
T	*a*	*b*	−0.05	−0.1	−0.25	−0.5	−0.75	−0.05	−0.1	−0.25	−0.5	−0.75
25	0.5	$s=1$	7.70	10.70	26.05	69.40	95.20	7.80	10.15	20.30	55.90	89.25
		$s=6$	32.55	64.55	98.60	100.00	100.00	9.25	16.25	35.65	72.80	96.55
		$s=16$	83.30	98.70	100.00	100.00	100.00	10.05	16.85	32.95	71.30	97.15
	1.0	$s=1$	4.65	7.45	14.85	48.05	83.00	6.70	8.90	16.80	50.20	86.00
		$s=6$	4.90	8.00	14.25	48.85	86.05	5.30	7.70	14.10	48.15	82.60
		$s=16$	4.75	7.70	15.80	48.55	84.75	6.10	7.75	16.05	48.55	85.15
100	0.5	$s=1$	18.35	47.45	99.60	100.00	100.00	13.75	28.00	85.90	100.00	100.00
		$s=6$	95.25	100.00	100.00	100.00	100.00	13.40	22.05	52.90	99.60	100.00
		$s=16$	100.00	100.00	100.00	100.00	100.00	12.55	18.70	45.80	98.35	100.00
	1.0	$s=1$	13.90	35.85	97.65	100.00	100.00	12.80	31.55	96.05	100.00	100.00
		$s=6$	13.40	33.80	97.70	100.00	100.00	12.40	34.15	96.25	100.00	100.00
		$s=16$	11.65	31.20	97.45	100.00	100.00	14.05	34.15	97.75	100.00	100.00

The DGP is generated under $H_1 : b<0$, $\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1,t}$, $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \mu_{z,t-1})$, $\Delta z_t = sD_{j_t} + u_{2,t}$, where $\sigma_1^2 = var(u_{1,t}) = 1$, and $\sigma_2^2 = var(u_{2,t}) = s^2$.

cointegration but without co-breaking in levels or in differences. Since the critical values and the size-adjusted power of the ECM-test depend on the type of structural break considered, there are many possible alternative combintions of breaks that could change completely the results on cointegration testing. In the following section we analyze whether this problem could be solved or reduced by using extended ECM models.

IV. EXTENDED ERROR CORRECTION MODELS

This section extends the implications of Toda and Yamamoto (1995) and Dolado and Lutkepohl (1996). They suggested to add an extra lag of the error correction term in order to get standard inference results even with nonstationary variables. In model (3.2), we just have to include an extra lag of the error correction term in the regression to obtain the following extended ECM model,

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + d(y_{t-2} - \alpha z_{t-2}) + u_t. \qquad (4.1)$$

The t-ratio of the parameter b in (4.1), $t_{\hat{b}}$ does not converge to the limiting distribution given in (2.16) or (2.17) but to the standard Normal distribution. This result not only simplifies the empirical work but also reduces, as we will see in this section, the unstability of the critical values due to structural breaks.

The intuition is the following. Consider the error correction model introduced in equations (2.2a) (2.2c) and compare the formulation with equation (4.1). Both are equivalent if the following condition is satisfied

$$c_t = c + d(y_{t-2} - \alpha z_{t-2}) \qquad (4.2)$$

or if

$$\Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \alpha\mu_{z,t-1}) = c + d(y_{t-2} - \alpha z_{t-2}). \qquad (4.3)$$

In general, this condition is not satisfied when there is no co-break in levels neither in differences. However, in the case of co-break in differences (not in levels), the term $(y_{t-2} - \alpha z_{t-2})$ will also have a break in levels that will approximate the ommited term $(\mu_{y,t-1} - \alpha\mu_{z,t-1})$. That is,

$$(\mu_{y,t-1} - -\alpha\mu_{z,t-1}) \approx c + d(y_{t-2} - \alpha z_{t-2}). \qquad (4.4)$$

On the other hand, if there is co-break in levels (not in differences), the term $(y_{t-2} - \alpha z_{t-2})$ has a break that can approximate the effect of $(\Delta\mu_{y,t} - a\Delta\mu_{z,t})$. Therefore,

$$\Delta\mu_{y,t} - a\Delta\mu_{z,t} \approx c + d(y_{t-2} - \alpha z_{t-2}). \qquad (4.5)$$

When there are extra dynamic regressors in the overparameterized model, as in (3.3), the extended equation of the ECM test becomes

$$\phi(B)\Delta y_t = c + \theta(B)\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + d(y_{t-k-2} - \alpha z_{t-k-2}) + u_t, \tag{4.6}$$

where the polynomials in B, $\phi(B)$ and $\theta(B)$ are

$$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_{p^*} B^{p^*}$$

$$\theta(B) = a - \theta_1 B - \theta_2 B^2 - \ldots - \theta_{q^*} B^{q^*}$$

and where $k = \max\{p^*, q^*\}$ is selected by using the SBIC selection criterion.

In this section we analyze the behavior of the critical values and the power of the ECM test ($t_{\hat{b}}$) with and without simultaneous co-breaking when using the extended ECM models, of equation (4.1) and (4.6).

As can be seen from Table 3, when we estimate Model (4.1), we get very stable critical values in all three co-breaking cases analyzed. It is important to notice that the critical values are not close to those obtained by Dickey and Fuller, but close to -1.8, which is similar to the Gaussian 5 percent critical value. The largest difference is found in the case of co-break in levels but not in differences, where critical values range from -1.67 to -2.5.

As usual, the critical values of Table 3 depend on the sample size. Only for sample size $T = 25$ and $T = 100$ the C.V.'s increase in absolute value when the common factor restriction holds ($a = 1$). For example, for $T = 100$, $a = 0.5$, $s = 16$, and D_3, the critical values is -1.7 and when $a = 1$ critical value is -2.04.

In summary, for reasonable sample sizes ($T > 100$) we could use standard $N(0, 1)$ critical values for any value of the parameter a, for any jump size (s), and for any type of segmented trend (D_1, D_2, D_3). We think that this is a very useful (robust) result for applied econometricians. Furthermore the power of the test is very good in all the three cases analyzed: simultaneous co-breaking, co-breaking in differences but not in levels, and co-breaking in levels but not in differences.

Remark 1. As expected, in the situation when there is no co-break in differences nor in levels the extended ECM model is not robust to structural changes. But even in this extreme case, the extended error correction critical values are more robust than those obtained with the usual error correction model.

Remark 2. The behaviour of the test when we include extra dynamic terms in the estimated model by means of the SBIC criterion is similar in terms of the critical values. However, the size-adjusted power of the test is reduced, especially in the case of co-breaks in differences but not in levels, see Arranz and Escribano (1998*b*).

TABLE 3: CRITICAL VALUES OF $t(\hat{b})$

Estimated model $\Delta y_t = \hat{c} + \hat{a}\Delta z_t + \hat{b}(y_{t-1} - z_{t-1})\hat{d}(y_{t-2} - z_{t-2}) + u_t$.

Simultaneous Cobreaking ($c_t = 0$).

		$a = 0$			$a = 0.5$			$a = 1$		
T	*DUM*	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$
25	NO	−2.500	−1.869	−1.806	−2.425	−2.076	−1.787	−2.614	−2.661	−2.635
	D_1	−1.953	−1.731	−1.703	−2.372	−1.814	−1.764	−2.658	−2.698	−2.607
	D_2	−1.948	−1.792	−1.761	−2.187	−1.796	−1.782	−2.594	−2.599	−2.517
	D_3	−2.000	−1.877	−1.735	−2.190	−1.797	−1.734	−2.660	−2.566	−2.719
100	NO	−2.006	−1.737	−1.582	−2.094	−1.800	−1.729	−2.133	−2.057	−2.107
	D_1	−1.796	−1.722	−1.727	−1.783	−1.623	−1.603	−2.105	−2.114	−2.150
	D_2	−1.737	−1.620	−1.704	−1.757	−1.735	−1.730	−2.069	−2.100	−2.069
	D_3	−1.704	−1.722	−1.762	−1.823	−1.701	−1.697	−2.162	−2.063	−2.046
1000	NO	−1.691	−1.672	−1.636	−1.871	−1.704	−1.757	−1.733	−1.673	−1.838
	D_1	−1.696	−1.664	−1.720	−1.701	−1.574	−1.672	−1.813	−1.749	−1.777
	D_2	−1.650	−1.624	−1.557	−1.735	−1.619	−1.675	−1.809	−1.822	−1.739
	D_3	−1.796	−1.608	−1.634	−1.727	−1.673	−1.637	−1.834	−1.810	−1.827

Cobreaking in differences, not cobreaking in levels ($\Delta\mu_{y,t} - a\Delta\mu_{z,t} = 0.5$).

		$a = 0$			$a = 0.5$			$a = 1$		
T	*DUM*	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$
25	NO	−2.230	−1.843	−1.817	−2.215	−1.993	−1.775	−2.113	−2.187	−2.119
	D_1	−2.173	−1.726	−1.705	−2.670	−1.829	−1.762	−2.217	−2.121	−2.121
	D_2	−2.135	−1.790	−1.762	−2.347	−1.799	−1.806	−2.176	−2.112	−2.043
	D_3	−2.135	−1.881	−1.735	−2.386	−1.787	−1.735	−2.203	−2.205	−2.342

100	NO	−1.848	−1.752	−1.584	−1.855	−1.787	−1.731	−1.848	−1.776	−1.830
	D_1	−1.810	−1.726	−1.723	−2.064	−1.628	−1.598	−1.831	−1.857	−1.808
	D_2	−1.787	−1.620	−1.700	−1.936	−1.732	−1.725	−1.830	−1.826	−1.799
	D_3	−1.715	−1.726	−1.764	−1.873	−1.711	−1.700	−1.873	−1.810	−1.780
1000	NO	−1.639	−1.669	−1.625	−1.785	−1.707	−1.728	−1.670	−1.659	−1.742
	D_1	−1.682	−1.662	−1.722	−1.774	−1.570	−1.672	−1.764	−1.664	−1.686
	D_2	−1.673	−1.635	−1.558	−1.784	−1.609	−1.670	−1.715	−1.737	−1.677
	D_3	−1.778	−1.608	−1.634	−1.734	−1.672	−1.635	−1.756	−1.665	−1.748

Cobreaking in levels, not cobreaking in differences ($\mu_{y,t} - \mu_{z,t} = 0.5$).

		$a = 0$			$a = 0.5$			$a = 1$		
T	*DUM*	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$	$s = 1$	$s = 6$	$s = 16$
25	NO	−2.500	−1.869	−1.806	−2.425	−2.076	−1.787	−2.614	−2.661	−2.635
	D_1	−2.474	−2.716	−2.852	−2.695	−2.470	−2.784	−2.658	−2.698	−2.607
	D_2	−2.535	−3.054	−2.380	−2.582	−2.770	−3.080	−2.594	−2.599	−2.517
	D_3	−2.555	−2.801	−2.873	−2.537	−2.602	−2.737	−2.660	−2.566	−2.719
100	NO	−2.006	−1.737	−1.582	−2.094	−1.800	−1.729	−2.133	−2.057	−2.107
	D_1	−2.165	−2.397	−2.519	−2.112	−2.181	−2.453	−2.105	−2.114	−2.150
	D_2	−2.081	−2.625	−2.849	−2.198	−2.373	−2.602	−2.069	−2.100	−2.069
	D_3	−2.045	−2.402	−2.403	−2.081	−2.250	−2.371	−2.162	−2.063	−2.046
1000	NO	−1.691	−1.672	−1.636	−1.871	−1.704	−1.757	−1.733	−1.673	−1.838
	D_1	−1.826	−2.225	−2.412	−1.792	−2.133	−2.363	−1.813	−1.749	−1.777
	D_2	−1.907	−2.388	−2.510	−1.912	−2.169	−2.591	−1.809	−1.822	−1.739
	D_3	−1.819	−2.112	−2.297	−1.789	−2.059	−2.088	−1.834	−1.810	−1.827

The DGP is generated under $H_0 : \Delta y_t = c_t + a\Delta z_t + u_{1_t}$, $\Delta z_t = sD_{j_t} + u_{2_t}$, where $c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t}$, $\sigma_1^2 = var(u_{1_t}) = 1$, and $\sigma_2^2 = var(u_{2_t}) = s^2$.

TABLE 4: SIZE ADJUSTMENT POWER OF THE TEST.
Estimated model $\Delta y_t = \hat{c} + \hat{a}\Delta z_t + \hat{b}(y_{t-1} - z_{t-1})d(y_{t-2} - z_{t-2}) + u_t$.

Simultaneous cobreaking

			NO					D_3				
T	*a*	*b*	*−0.05*	*−0.1*	*−0.25*	*−0.5*	*−0.75*	*−0.05*	*−0.1*	*−0.25*	*−0.5*	*−0.75*
25	0.5	$s=1$	7.25	9.95	22.15	62.50	92.75	9.80	12.05	30.95	73.45	97.15
		$s=6$	12.60	29.90	91.60	100.00	100.00	20.70	46.25	96.50	100.00	100.00
		$s=16$	50.85	93.25	100.00	100.00	100.00	62.10	97.15	100.00	100.00	100.00
	1.0	$s=1$	5.80	7.80	15.85	48.75	81.30	6.00	7.60	15.70	46.15	82.90
		$s=6$	4.05	6.65	13.35	45.50	82.35	6.40	9.20	16.60	52.40	84.60
		$s=16$	5.05	6.55	14.35	45.60	82.20	4.40	5.35	14.25	42.95	81.50
100	0.5	$s=1$	9.45	21.65	78.25	99.95	100.00	13.55	29.35	86.65	100.00	100.00
		$s=6$	43.50	90.70	100.00	100.00	100.00	50.75	94.00	100.00	100.00	100.00
		$s=16$	98.25	100.00	100.00	100.00	100.00	99.50	100.00	100.00	100.00	100.00
	1.0	$s=1$	8.80	16.90	67.85	99.90	100.00	9.60	17.60	68.40	99.80	100.00
		$s=6$	9.95	19.10	70.20	99.85	100.00	10.10	20.75	70.95	99.75	100.00
		$s=16$	8.45	17.30	68.45	100.00	100.00	10.75	20.20	72.75	99.85	100.0

Cobreaking in Differences, Not in Levels

			NO					D_3				
T	*a*	*b*	*−0.05*	*−0.1*	*−0.25*	*−0.5*	*−0.75*	*−0.05*	*−0.1*	*−0.25*	*−0.5*	*−0.75*
25	0.5	$s=1$	4.75	5.20	6.95	19.10	44.05	6.30	7.60	14.35	31.80	59.50
		$s=6$	12.80	24.40	67.70	93.25	99.05	10.25	17.45	42.20	74.25	90.90
		$s=16$	49.50	89.70	100.00	100.00	100.00	23.80	43.50	67.15	86.75	95.30
	1.0	$s=1$	5.30	6.25	7.30	17.85	39.25	5.00	4.80	7.50	16.45	34.15
		$s=6$	4.65	4.75	6.55	16.30	34.80	4.90	4.60	6.40	15.55	35.25
		$s=16$	6.20	5.40	7.35	17.95	37.25	3.15	3.05	5.55	12.60	30.80

100	0.5	$s=1$	6.45	9.65	28.95	84.20	99.65	8.05	11.65	34.00	84.25	99.60
		$s=6$	19.45	40.15	94.00	100.00	100.00	20.00	38.05	84.70	99.75	100.00
		$s=16$	88.00	99.70	100.00	100.00	100.00	48.55	67.85	93.75	99.95	100.00
	1.0	$s=1$	5.70	8.20	26.30	74.55	99.10	5.85	8.70	26.50	75.00	99.20
		$s=6$	6.45	9.65	28.50	78.95	99.05	6.45	9.70	26.65	76.55	98.65
		$s=16$	5.70	8.60	24.80	75.60	98.95	7.30	10.20	29.40	79.40	99.45

Cobreaking in Levels, Not in Differences

			NO					D_3				
T	*a*	*b*	*−0.05*	*−0.1*	*−0.25*	*−0.5*	*−0.75*	*−0.05*	*−0.1*	*−0.25*	*−0.5*	*−0.75*
25	0.5	$s=1$	7.25	9.95	22.15	62.50	92.75	8.10	9.15	19.80	56.95	90.60
		$s=6$	12.60	29.90	91.60	100.00	100.00	9.25	15.70	42.65	90.30	99.65
		$s=16$	50.85	93.35	100.00	100.00	100.00	11.10	17.75	44.70	94.50	100.00
	1.0	$s=1$	5.80	7.80	15.85	48.75	81.30	6.00	7.60	15.70	46.15	82.90
		$s=6$	4.05	6.65	13.35	45.50	82.35	6.40	9.20	16.60	52.40	84.60
		$s=16$	5.05	6.55	14.35	45.60	82.20	4.40	5.35	14.25	42.95	81.50
100	0.5	$s=1$	9.45	21.65	78.25	99.95	100.00	10.70	21.90	78.55	99.95	100.00
		$s=6$	43.50	90.70	100.00	100.00	100.00	17.55	41.10	98.95	100.00	100.00
		$s=16$	98.25	100.00	100.00	100.00	100.00	19.75	44.15	100.00	100.00	100.00
	1.0	$s=1$	8.80	16.90	67.85	99.90	100.00	9.60	17.60	68.40	99.80	100.00
		$s=6$	9.95	19.10	70.20	99.85	100.00	10.10	20.75	70.95	99.75	100.00
		$s=16$	8.45	17.30	68.45	100.00	100.00	10.75	20.20	72.25	99.85	100.00

The DGP is generated under $H_1 : b<0$, $\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t}$, $c_t = \Delta\mu_{z,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \mu_{z,t-1})$, $\Delta z_t = sD_{j_t} + u_{2,t}$, where $\sigma_2^2 = var(u_{1,t}) = 1$, and $\sigma_2^2 = var(u_{2,t}) = s^2$.

We tried some other specifications, such as

$$\phi(B)\Delta y_t = c + \theta(B)\Delta z_t + b(y_{t-1} - z_{t-1}) + d(y_{t-2} - z_{t-2}) + u_t$$

$$\phi(B)\Delta y_t = c + a\Delta z_t + \theta(B)(y_{t-1} - z_{t-1}) + u_t$$

$$\Delta y_t = c + \theta(B)\Delta z_t + \phi(B)(y_{t-1} - z_{t-1}) + u_t$$

but we did not get better results. Therefore, with overparameterized models like (4.6), we found that there is a trade-off between critical values stability and power of the test.

Final remark. We also analyzed the ECM model of the trend components of y_t, and z_t, where the trend components were obtained by applying the Hodrick and Prescott (1980, 1997), and the Baxter and King (1995) filters. The experimental results are similar to the ones obtained with the observed series, but the power of the test is lower, see Arranz and Escribano (1998*b*).

V. CONCLUSIONS

We have analyzed the effects of different structural breaks on the ECM test for non-cointegration when no dummy variables are included in the model. It is well known that critical values (C.V.) depend on the type of break and other nuisance parameters (constant, trends, etc.). In particular, we have analyzed the dependence of the critical values on the different timing of the break, (D_1, D_2), different sizes of the break, ($s = 1, 6, 16$), and even allow for the possibility of having two breaks (D_3) on the first difference of the mean (segmented trends in levels).

The fact that critical values depend on nuisance parameters and that the usual ECM test has very low power under misspecification of the co-breaks in the level of the series, opens the possibility of improving the robustness of the results by using extended ECM models. The most simple extended model is the one that adds an extra lag of the ECM term, $(y_{t-2} - \alpha z_{t-2})$, to the usual error correction model. Our Monte Carlo experiments show that the critical values of the extended ECM test are very stable (robust), follow a standard distribution and, when no extra dynamic terms are included, the power of the extended ECM test is excellent under any partial co-breaking circumstances. Clearly those results are not robust when there is no co-breaks in levels nor in differences either. Therefore, extended ECM models represent a simple robust testing technique in the presence of structural co-breaks.

Universidad Carlos III de Madrid

APPENDIX A. CO-BREAKS IN LEVELS AND DIFFERENCES

Consider the DGP given by (2.2a)–(2.2c). There are four cases of interest

1. $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$ and $\mu_{y,t} - \alpha\mu_{z,t-1} \neq constant$, co-break in differences.
2. $\Delta\mu_{y,t} - a\Delta\mu_{z,t} \neq constant$ and $\mu_{y,t-1} - \alpha\mu_{z,t-1} = c_l$, co-break in levels.
3. $\Delta\mu_{y,t} - a\Delta\mu_{z,t} \neq constant$ and $\mu_{y,t-1} - \alpha\mu_{z,t-1} \neq constant$, no co-break.
4. $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$ and $\mu_{y,t-1} - \alpha\mu_{z,t-1} = c_l$, co-break in levels and differences.

Remark 1. When the slopes of the deterministic trends are constant, say $\Delta\mu_z = g_z$ and $\Delta\mu_y = g_y$, then

$$\mu_{y,t-1} - \alpha\mu_{z,t-1} = \mu_y^0 - \alpha\mu_z^0 + (g_y - \alpha g_z)(t-1) \tag{A.1a}$$

$$c_t = (g_y - ag_z) - b(\mu_y^0 - \alpha\mu_z^0) - b(g_y - \alpha g_z)(t-1) \tag{A.1b}$$

For co-breaking in levels, the necessary condition is $g_y - \alpha g_z = 0$, but that implies that $\Delta\mu_{y,t} - a\Delta\mu_{z,t} - (\alpha - a)g_z$, which is constant, and the co-break in differences condition is met. Therefore, in the case of constant slopes of the trend (no breaks), co-break in levels implies co-break in differences.

On the other hand, in the case of co-break in differences, it must be $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = g_y - ag_z = c_d$ and hence

$$\mu_{y,t-1} - \alpha\mu_{z,t-1} = \mu_{y,0} - \alpha\mu_{z,0} + c_d(t-1) + g_z(a - \alpha)(t-1).$$

Thus, co-break in differences implies co-break in levels only when $c_d = 0$ and $a = \alpha$ (COMFAC restriction).

Remark 2. Assume that there is a break, so that $\Delta\mu_{z,t} = g_z + s_z D_t$ and $\Delta\mu_y = g_y + s_y D_t$, then

$$\begin{aligned}\mu_{y,t-1} - \alpha\mu_{z,t-1} &= \mu_y^0 - \alpha\mu_z^0 + (g_y - \alpha g_z)(t-1)\\ &\quad + (s_y - \alpha s_z)(t-1-t_1)D_{t-1}\end{aligned} \tag{A.2a}$$

$$\begin{aligned}c_t &= (g_y - ag_z) + (s_y - as_z)D_t - b(\mu_y^0 - \alpha\mu_z^0)\\ &\quad - b(g_y - \alpha g_z)(t-1) - b(s_y - \alpha s_z)(t-1-t_1)D_{t-1}\end{aligned} \tag{A.2b}$$

The necessary conditions to have co-breaks in differences are $g_y - ag_z = c_d$ and $s_y - as_z = 0$. In that case

$$\mu_{y,t-1} - \alpha\mu_{z,t-1} = \mu_y^0 - \alpha\mu_Z^0 + c_d(t-1) + (a-\alpha)g_z(t-1)$$
$$+ (a-\alpha)s_z(t-1-t_1)D_{t-1}$$

and co-breaks in differences implies co-break in levels only when $a = \alpha$ (COMFAC restriction) and $c_d = 0$.

Conversely, the necessary conditions to have co-breaks in levels, from equation (A.2a), are $g_y - ag_z = 0$ and $s_y - \alpha s_z = 0$. Therefore, when $a = \alpha$, these conditions are the ones required for co-break in differences taking $c_d = 0$. In effect, under co-breaks in levels

$$\Delta\mu_{y,t} - a\Delta\mu_{z,t} = (a-\alpha)g_z + (a-\alpha)s_z D_t,$$

and in the case that $a = \alpha$ (COMFAC restriction), there would be co-break in differences too.

REFERENCES

Andrés, J., Escribano, A., Molinas, C. and Taguas, D. (1990). *La inversión en España: Econometría con restricciones de equilibrio*, A. Bosch.

Andrews, D. (1993). 'Tests of parameter instability and structural change with unknown break point', *Econometrica* Vol. 61, pp. 821–56.

Andrews, D., Lee, W. and Ploberger, W. (1996). 'Optimal change points tests for normal linear regression', *Journal of Econometrics* Vol. 70, pp. 9–38.

Arranz, M. and Escribano, A. (1998*a*). Cointegration testing under transitory breaks: Robust ECM testing procedures, Working Paper 98–82, Universidad Carlos III de Madrid.

Arranz, M. and Escribano, A. (1998*b*). Detrending procedures and cointegration testing: ECM tests under structural breaks, Working Paper 98–15, Universidad Carlos III de Madrid.

Arranz, M. and Escribano, A. (1998*c*). Is unit root testing on trend components robust against additive outliers?, Working Paper 98–79, Universidad Carlos III de Madrid.

Bai, J. (1997). 'Estimating multiple breaks one at a time', *Econometrics Theory* Vol. 13, pp. 315–52.

Bai, J. and Perron, P. (1998). 'Estimating and testing linear models with multiple structural changes', *Econometrica* Vol. 66, pp. 47–78.

Banerjee, A., Dolado, J., Hendry, D. F. and Smith, G. (1986). 'Exploring equilibrium relationships in econometrics through static models: Some Monte Carlo evidence', BULLETIN Vol. 48, pp. 853–77.

Banerjee, A., Dolado, J. and Mestre, F. (1998). 'ECM test for cointegration in a single equation framework', *Journal of Time Series Analysis* Vol. 19, pp. 267–83.

Banerjee, A., Dolado, J., Galbraith, J. W. and Hendry, D. F. (1993). *Cointegration, error-correction, and the econometric analysis of non-stationary data*, Oxford University Press.

Banerjee, A., Lumsdaine, R. and Stock, J. (1992). 'Recursive and sequential tests

of unit root and trend break hypothesis: Theory and international evidence', *Journal of Business and Economic Statistics* Vol. 10, pp. 271–87.
Banerjee, A., Lazarova, S. and Urga, G. (1998). Bootstrapping sequential tests for multiple structural breaks. Discussion Paper 17–98, Center for Economic Forecasting, London Business School.
Baxter, M. and King, R. (1995). Measuring business cycles: Approximate band-pass filters for economic time series, Working Paper 5022, NBER.
Campos, J., Ericsson, N. and Hendry, D. (1996). 'Cointegration tests in the presence of structural breaks', *Journal of Econometrics* Vol. 70, pp. 187–220.
Dickey, D. and Fuller, W. (1979). 'Distribution of the estimators for autoregressive time series with a unit root', *Journal of the American Statistical Association* Vol. 74, pp. 427–31.
Dolado, J. and Lutkepohl, H. (1996). 'Making Wald tests work for cointegrated VAR systems', *Econometrica Reviews* Vol. 15, pp. 369–86.
Engle, R. F. and Granger, C. W. J. (1987) 'Cointegration and error-correction: Representation, estimation, and testing', *Econometrica* Vol. 55, pp. 251–76.
Escribano, A. (1987). Co-integration, time co-trends and error-correction systems: an alternative approach, Discussion Paper 8715, CORE.
Franses, P. and Lucas, A. (1997*a*). Outlier robust cointegration analysis of Dutch interest rates, *in* 'Proceedings of the 1996 Meeting of the American Statistical Association', ASA, pp. 106–9.
Franses, P. and Lucas, A. (1997*b*). Outlier robust cointegration tests, VU Research Memorandum 97–45, Vrije Universiteit Amsterdam.
Franses, P. and Haldrup, N. (1994). 'The effects of additive outliers on unit roots and cointegration', *Journal of Business and Economic Statistics* Vol. 12, pp. 471–78.
Hansen, B. (1995). 'Rethinking the univariate approach to unit root testing: Using covariates to increase power', *Econometric Theory* Vol. 11, pp. 1148–1172.
Hendry, D. (1996). A theory of co-breaking. Paper presented at the Time Series Meeting organized by NBER-NSF.
Hendry, D. and Neale, A. (1990). The impact of structural breaks on unit root tests, *in* P. Hackl and A. Westlung, eds, 'Economic Structural Change: Analysis and Forecasting', Berlin: IIASA and Springer-Verlag, pp. 95–119.
Hodrick, R. and Prescott, E. (1980). Post-war U.S. business cycle: An empirical investigation, Technical report, Carnegie- Mellon University.
Hodrick, R. and Prescott, E. (1997). 'Post-war U.S. business cycle: An empirical investigation', *Journal of Money, Credit and Banking* Vol. 29, pp. 1–16.
Kremers, J., Ericsson, N. and Dolado, J. (1992). 'The power of cointegration tests', BULLETIN Vol. 54, pp. 349–67.
Leybourne, S., Mills, T. and Newbold, P. (1998). 'Spurious rejections by Dickey–Fuller tests in the presence of a break under the null', *Journal of Econometrics* Vol. 87, pp. 191–203.
Lucas, A. (1995*a*). 'An outlier robust unit root tess with an application to the extended Nelson-Plosser data', *Journal of Econometrics* Vol. 66, pp. 153–73.
Lucas, A. (1995*b*). 'Unit roots tests based on M estimators', *Econometric Theory* Vol. 11, pp. 331–46.
Lucas, A. (1997). 'Cointegration testing using pseudo likelihood-ratio tests', *Econometric Theory* Vol. 13, pp. 149–69.

Park, J. and Phillips, P. (1988). 'Statistical inference in regressions with integrated processes: Part I', *Econometric Theory* Vol. 4, pp. 458–97.
Park, J. and Phillips, P. (1989). 'Statistical inference in regressions with integrated processes: Part II', *Econometric Theory* Vol. 5, pp. 95–131.
Park, J. Y. (1992). 'Canonical cointegrating regression: *Econometrica* Vol. 60, pp. 119–43.
Perron, P. (1989). 'The great crash, the oil price shock and the unit root hypothesis', *Econometrica* Vol. 57, pp. 1361–1401.
Perron, P. and Vogelsang, T. (1992). 'Nonstationarity and level shifts with an application to purchasing power parity', *Journal of Business and Economic Statistics* Vol. 10, pp. 301–20.
Phillips, P. C. B. (1987). 'Towards a unified asymptotic theory for autoregression', *Biometrika* Vol. 74, pp. 535–47.
Phillips, P. C. B. and Hansen, B. (1990). 'Statistical inference in instrumental variables regression with I(1) processes', *Review of Economic Studies* Vol. 57, pp. 99–125.
Rappoport, P. and Reichlin, L. (1989). 'Segmented trends and nonstationary time series', *Economic Journal* Vol. 99, pp. 168–77.
Toda, H. and Yamamoto, T. (1995). 'Statistical inference in vector autoregression with possibly integrated processes', *Journal of Econometrics* Vol. 66, pp. 225–50.
Vogelsang, T. (1997). 'Wald-type tests for detecting breaks in the trend of a dynamic time series', *Econometric Theory* Vol. 13, pp. 818–49.
Zivot, E. and Andrews, D. (1992). 'Further evidence on the great crash, the oil price shock and the unit root hypothesis', *Journal of Business and Econometric Statistics* Vol. 10, pp. 251–87.