

# Bootstrapping Cointegration Tests Under Structural Co-Breaks: A Robust Extended ECM Test

Miguel A. Arranz

*Bayes Forecast*

*Gran Vía 39, Madrid, Spain*

Álvaro Escribano\*

*Dept. of Economics, Universidad Carlos III de Madrid, Spain*

*and Dept. of Economics, Georgetown University, USA*

## Abstract

The aim of the paper is the analysis of ECM (Error Correction Model) bootstrap cointegration tests under structural breaks. Classical ECM tests depend on some nuisance parameters, which is an undesirable feature for empirical applications. This problem is overcome by using the bootstrap ECM test, which shows good size and power properties when there are no breaks. In this paper we study the small sample properties of alternative bootstrap ECM tests under different co-breaking situations. ECM test statistics are made robust to partial co-breaking by using extended error correction models or by imposing a common factor restriction.

**Key Words:** Bootstrap, structural breaks, cointegration testing, extended error correction model, co-breaks.

**AMS subject classification:** 62M10, 91B84, 62F40, 82C80, 62F03, 62P20.

## 1 Introduction

Cointegration tests are affected by nuisance parameters coming from having structural breaks in the series, from the timing of the break, and also from

---

\*Correspondence to: Álvaro Escribano. Department of Economics, Universidad Carlos III de Madrid, C./ Madrid, 126, 28903 Getafe (Madrid), Spain. E-mail: alvaroe@eco.uc3m.es

The second author acknowledges support from the following grants: Spanish MCyT BEC2002-00279, The European TMR-ERB-40618C97-0994 and the Secretaría de Estado de Universidades PR2002-0305

Received: September 2002; Accepted: June 2004

the size of the break, (see for example Campos et al., 1996; Kremers et al., 1992). However, it is common to make the assumption that the structural breaks affect both series so that the short-run relationship is not affected by any of the breaks, showing full co-breaking, as exposed in Clements and Hendry (1999).

Arranz and Escribano (2000a) study the small sample effects that structural breaks have on cointegration tests based on single equation error correction models (ECM cointegration test) under different levels of co-breaking. They provided further evidence that critical values of the test depended on the size of the jump and the contemporaneous correlation in the short-run conditional relationship even under simultaneous co-breaking, and that those dependencies increase when there is only partial co-breaking. They found that the ECM test is made robust to some standard breaks by using extended error correction models.

Computing intensive methods, such as the bootstrap,<sup>1</sup> based on extended error correction models, could therefore provide good small sample properties for ECM tests. The bootstrap technique has been traditionally considered in *iid* situations, but it can also be applied to time series (see Berkowitz and Kilian, 2000; Bühlmann, 1999; Kiviet, 1984; Li and Maddala, 1996; van Giersbergen and Kiviet, 1996). It has recently been stated (Mantalos and Shukur, 1998) that the bootstrap ECM test has good size and power properties. Our purpose is to check whether that conclusion is preserved in the case of co-breaks. In particular, we are concerned about the analysis of ECM cointegration tests under structural breaks when the critical values are obtained by bootstrapping the residuals of the short-run dynamic model. In this paper, we analyze several resampling schemes in order to obtain the most reliable bootstrap critical values for testing the null hypothesis of no cointegration vs. cointegration. We analyze both the size and the power of the ECM test under alternative cases of structural co-breaks: simultaneous co-breaking, co-breaking in levels but not in differences, and finally co-breaking in differences but not in levels.

The paper is organized as follows. The concepts of simultaneous co-breaking, co-breaking in levels and co-breaking in differences are briefly reviewed in Section 2. Section 3 discusses the implementation of the alternative bootstrap resampling techniques and bootstrap statistics for hypothesis

---

<sup>1</sup> See Davidson and Hinkley (1997); Hall (1992), Efron and Tibshirani (1993), Hall (1994), Shao and Tu (1995) and Horowitz (2000).

testing. We also make several remarks and guidelines about the different resampling schemes and bootstrap statistics applied in dynamic regression models (see Li and Maddala, 1996, 1997). Section 4 analyzes the results of our Monte Carlo simulation study, under no breaks, and Section 5 includes the Monte Carlo simulation results under partial co-breaking. Conclusions and directions of future research are given in Section 6.

## 2 Error correction models with co-breaks in levels and in differences

Consider the following bivariate error correction model (ECM)

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (2.1a)$$

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

$$c_t \equiv \Delta \mu_{y,t} - a\Delta \mu_{z,t} - b(\mu_{y,t-1} - \alpha \mu_{z,t-1}) \quad (2.1c)$$

$$\Delta \mu_{z,t} = sD_{j,t} \quad (2.1d)$$

where  $\mu_{y,t} = E(y_t)$ ,  $\mu_{z,t} = E(z_t)$  are, given fixed initial conditions,  $y_0$ , and  $z_0$ , respectively, the corresponding unconditional means, and therefore they include all possible deterministic components of  $y_t$  and the exogenous variable  $z_t$  like: constant terms, deterministic trends, dummy variables, segmented trends, outliers, etc. The stochastic errors  $u_{1t}$  and  $u_{2t}$  are jointly, and serially independent, zero mean, normally distributed white noise processes with constant variances  $\text{Var}(u_{1t}) = \sigma_1^2$  and  $\text{Var}(u_{2t}) = \sigma_2^2$ . In equation (2.1d)  $s$  is the parameter value 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. We define in Section 5.1 three different kinds of breaks ( $j = 1, 2, 3$ ). In this paper we investigate the effects of having alternative structural breaks in  $c_t$  when applying the ECM test for non-cointegration ( $b = 0$ ) by bootstrap methods. Three main possible co-breaking situations are considered in the following definitions:

**Definition 2.1.** *Let  $E(y_t) = \mu_{y,t}$  and  $E(z_t) = \mu_{z,t}$  given valid initial conditions, 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.** *Let  $E(y_t) = \mu_{y,t}$  and  $E(z_t) = \mu_{z,t}$  given valid initial conditions, 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.*

**Definition 2.3.** Let  $E(y_t) = \mu_{y,t}$  and  $E(z_t) = \mu_{z,t}$  given valid initial conditions, 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), there is also simultaneous co-breaking.

In the case of simultaneous co-breaking, the intercept  $c_t$  from (2.1c) is constant,  $c_t = c$  and the error correction model from (2.1a) becomes the standard conditional ECM model where the only deterministic regressor is a constant term,  $c$ .

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

Figure 1 displays an example of two time series showing simultaneous co-breaks. Notice that, although both series are affected by breaks in the trend, the value of  $c_t$  remains constant.

Several possible intermediate cases of interest in empirical applications are specially studied in our Monte Carlo experiment.

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

Co-break in levels implies  $\mu_{y,t} - \alpha\mu_{z,t} = c_t$ . Taking first differences, we have  $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$ , and equation (2.1a) becomes

$$\Delta y_t = -bc_t + (\alpha - a)sD_{j,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (2.3)$$

**Remark 2.1.** Hence, the breaks in the marginal process of  $\Delta z_t$  affect the error correction model, and therefore the ECM test, under the null hypothesis and under the alternative, unless the COMFAC restriction ( $a = \alpha$ ) is satisfied.

**Remark 2.2.** Notice that co-breaks in levels implies co-breaks in differences (and thus full co-breaking) when the COMFAC restriction ( $a - \alpha = 0$ ) is met (see Arranz and Escribano, 2000a).

Figure 2 displays two time series with co-breaks in levels, but no in differences. Notice that both series are affected by a double break in the trend and that this is also shown in the variable  $c_t$ .

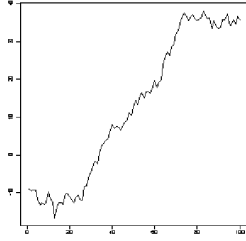


Figure 1(a):  $y$

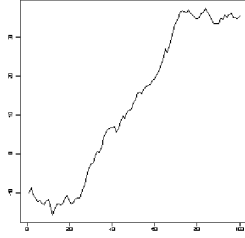


Figure 1(b):  $z$

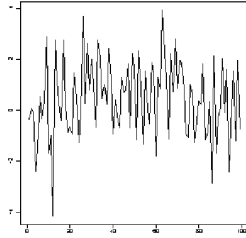


Figure 1(c):  $\Delta y$

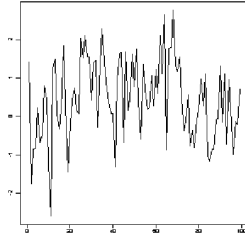


Figure 1(d):  $\Delta z$

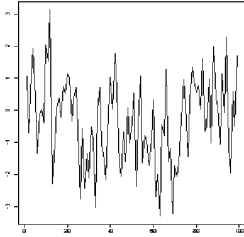


Figure 1(e):  $y_{t-1} - z_{t-1}$

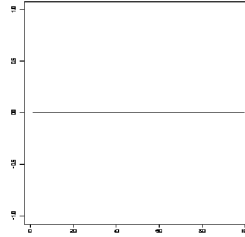


Figure 1(f):  $c_t$

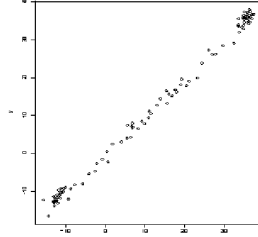


Figure 1(g): Cross-plot

Figure 1: Simultaneous co-breaks,  $N = 100$ ,  $a = 0$ ,  $s = 1$ ,  $b = -0.75$ , and  $D_{3,t}$  as defined in Section 5.1

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

*Co-break in differences implies that  $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d$ . From equation (2.1a)*

$$\Delta y_t = c_m + bc_d t - b(a - \alpha)\mu_{z,t-1} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (2.4a)$$

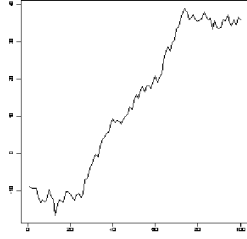


Figure 2(a): $y$

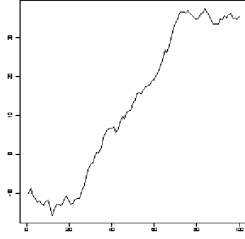


Figure 2(b): $z$

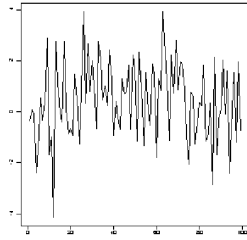


Figure 2(c): $\Delta y$

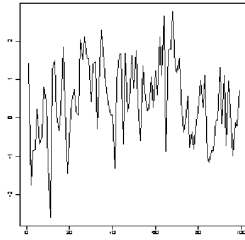


Figure 2(d): $\Delta z$

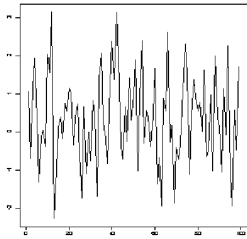


Figure 2(e): $y_{t-1} - z_{t-1}$

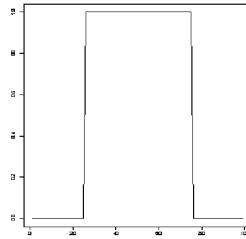


Figure 2(f): $c_t$

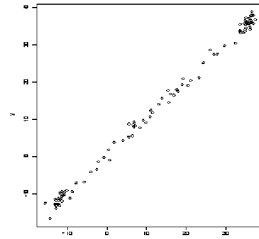


Figure 2(g):Cross-plot

Figure 2: Co-breaks in levels, not in differences,  $N = 100$ ,  $c_l = 0$ ,  $a = 0$ ,  $s = 1$ ,  $b = -0.75$ , and  $D_{3,t}$  as defined in Section 5.1

$$\mu_{z,t} = \mu_{z,0} + s \sum_{i=1}^t D_{j,i} \quad (2.4b)$$

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

Figure 3 shows two time series with co-breaks in differences, not in levels. Notice that  $c_t$  reflects the double break in the trend, but that  $c_t$  is very similar to  $y_{t-1} - z_{t-1}$ .

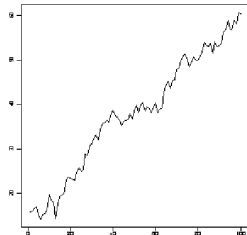


Figure 3(a):  $y$

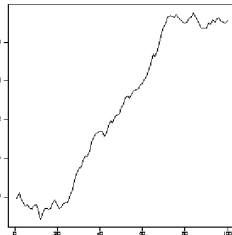


Figure 3(b):  $z$

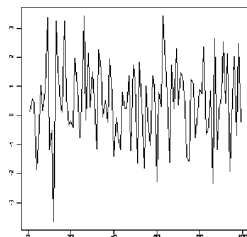


Figure 3(c):  $\Delta y$

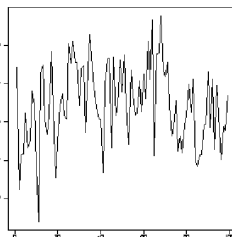


Figure 3(d):  $\Delta z$

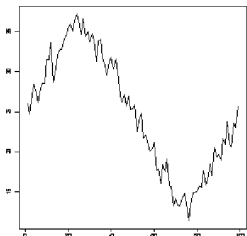


Figure 3(e):  $y_{t-1} - z_{t-1}$

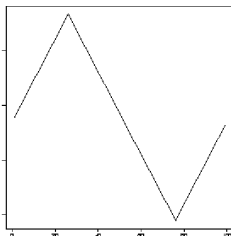


Figure 3(f):  $c_t$

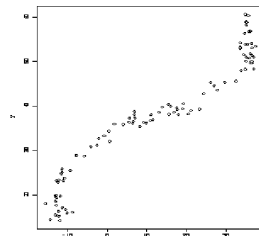


Figure 3(g): Cross-plot

Figure 3: Co-breaks in differences, not in levels,  $N = 100$ ,  $a = 0$ ,  $c_d = 0.5$ ,  $s = 1$ ,  $b = -0.75$ , and  $D_{\varepsilon,t}$  as defined in Section 5.1

**Remark 2.3.** *The break in the marginal process of  $z_t$  affects the error correction model only under the alternative hypothesis,  $H_1 : b < 0$ .*

**Remark 2.4.** Notice that co-break in differences would imply co-break in levels only in the case that the COMFAC (common factor) restriction ( $a - \alpha = 0$ ) is satisfied and  $c_d = 0$  (see Arranz and Escribano, 2000a).

Depending on the type of dummy variable considered,  $D_{j,t}$ , we could have segmented trends in the model (2.1a) with one or several breaking points, see Section 5 for more details.

The extended error correction model (extended ECM) is the following,

$$\Delta y_t = c + a\Delta z_t + b[y_{t-1} - \alpha z_{t-1}] + d[y_{t-2} - \alpha z_{t-2}] + \eta_t. \quad (2.5)$$

One important feature of the ECM-test analyzed by Arranz and Escribano (2000a), is that when using an extended error correction model (Dolado and Lütkepohl, 1996; Toda and Yamamoto, 1995), the critical values are robust (stable) in the sense that they do not depend so much on nuisance parameters, and the power of the test increases over the standard one when there is at least partial co-breaking. However, the results were clearly stated with sample sizes of at least 200 observations. That led us to search for an alternative testing procedure that can be used with smaller sample sizes and that do not depend so much on the type of co-breaking process considered, and bootstrap seems a good candidate.

### 3 Bootstrapping the ECM test statistic

The first two resampling schemes that we are using here were first proposed by van Giersbergen and Kiviet (1993, 1994) when discussing two rules suggested by Hall and Wilson (1991) in the context of hypothesis testing in dynamic regression models. To clarify our resampling schemes, we discuss them in the context of the ECM test that we will use in Sections 4 and 5.

Consider the following dynamic regression model in error correction form, where  $\alpha = 1$ ,

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + \epsilon_t. \quad (3.1)$$

We are interested in testing the null hypothesis  $H_0 : b = 0$  (no cointegration) vs.  $H_1 : b < 0$  (cointegration), with the  $t$ -ratio of  $b$ ,  $t(\hat{b})$ , statistic. Let  $\hat{a}$ ,  $\hat{b}$ , and  $\hat{c}$  be the OLS estimators of  $a$ ,  $b$ , and  $c$ , respectively, and let  $\epsilon_t^*$  be the bootstrap residuals obtained by resampling the OLS residuals,  $\hat{\epsilon}_t$ , or



the centered OLS residuals in the case that no constant term ( $c$ ) is included in the OLS regression.  $NB$  indicates the number of bootstrap resamples.

Consider two alternative sampling schemes  $S_1$ , and  $S_2$  for generating the bootstrap samples:

$$S_1 : \Delta y_t^* = \hat{c} + \hat{a}\Delta z_t + \hat{b}(y_{t-1}^* - z_{t-1}) + \epsilon_t^* \quad (3.2)$$

$$S_2 : \Delta y_t^* = \hat{c} + \hat{a}\Delta z_t + \epsilon_t^*, \quad (3.3)$$

Both procedures use  $\epsilon_t^*$  but differ in the way  $y_t^*$  is generated. For each sampling scheme consider two  $t$ -statistics

$$t_1 : t(\hat{b}) = (\hat{b}^* - \hat{b})/SE(\hat{b}^*) \quad (3.4)$$

$$t_2 : t(b_0) = \hat{b}^*/SE(\hat{b}^*) \quad (3.5)$$

where  $\hat{b}^*$  is the OLS estimator of  $b^*$  in

$$\Delta y_t^* = c^* + a^*\Delta z_t + b^*(y_{t-1}^* - z_{t-1}) + \epsilon_t^* \quad (3.6)$$

and  $SE(\hat{b}^*)$ , the standard error of  $\hat{b}^*$ , is calculated from the empirical distribution of  $\hat{b}^*$ , that is, after the  $NB$  resamplings are done or from every estimated regression with the usual textbook formula. The bootstrap critical value is obtained by looking at the 5% lower tail of the empirical distribution of  $t(\hat{b})$  or  $t(b_0)$ , when using  $t_1$  or  $t_2$ , respectively.

Thus four versions of the  $t$ -statistic can be defined. Hall and Wilson (1991) only consider sampling scheme  $S_1$  and suggest using  $t_1$  only. They do not consider the sampling scheme  $S_2$  which is the appropriate one for  $t_2$ . Van Giersbergen and Kiviet (1993) suggest, on the basis of a Monte Carlo study of an AR(1) model, the use of  $t_2$  under sampling scheme  $S_2$  instead of  $t_1$  under  $S_1$ . Their main conclusions are the following:

1. Inference based on  $t_2$  under  $S_1$  has low power, and its size is close to zero. Similarly,  $t_1$  under  $S_2$  should not be used. Therefore the resampling scheme should mimic the null distribution of the test statistic to be bootstrapped.
2. Using  $t_1$  under  $S_1$  and  $t_2$  under  $S_2$  are equivalent in non-dynamic models. However, in dynamic models this equivalence breaks down in finite samples. Van Giersbergen and Kiviet suggest using  $t_2$  under  $S_2$ .

3. The limiting distributions of  $t_1$  under  $S_1$  and  $t_2$  under  $S_2$  are identical even with dynamic models. The conclusion that  $t_2$  under  $S_2$  is suggested within this context is based on the small sample performance.

Van Giersbergen and Kiviet (1994), and Li and Maddala (1996, 1997) propose a different resampling scheme, based on the OLS residuals of the restricted model,  $\hat{\epsilon}_{0t} = \Delta y - \hat{c}^r + \hat{a}^r \Delta z_t$ , and  $\hat{a}^r$ ,  $\hat{b}^r$ , and  $\hat{c}^r$ , are the OLS estimates of the ECM model under  $H_0 : b = 0$ ,

$$\Delta y_t = c + a \Delta z_t + \epsilon_{0t} \quad (3.7)$$

$$S_3 : \Delta y_t^* = \hat{c}^r + \hat{a}^r \Delta z_t + \epsilon_{0t}^* \quad (3.8)$$

where  $\epsilon_{0t}^*$  is the bootstrap sample from  $\epsilon_{0t}$ . Notice that both  $S_1$  and  $S_2$  procedures use resampling based on the OLS residuals under  $H_1 : b < 0$ ,  $\hat{c}$ . If the null  $H_0 : b = 0$  is true but the OLS estimator of  $b$  is far from 0, the empirical distribution of the residuals will suffer from a poor approximation of the distribution of the errors under the null. The intuition behind the  $S_3$  procedure is the following: if the null hypothesis is true,  $\epsilon_{0t}$  is exactly the true distribution of the regression errors, and hypothesis testing based on this will give (approximately) the correct size of the test. If the null is not true, then  $\epsilon_{0t}$  is different from the true distribution of the errors. Hypothesis testing based on  $t_1$  or  $t_2$  will have good power depending on how far the null is away from the true value of  $b$ . Therefore, we would expect that  $t_2$  under  $S_3$  should work better than using  $t_1$  under  $S_1$  or  $t_2$  under  $S_2$ .

#### 4 Monte Carlo simulation experiment I. No breaks

Consider the following bivariate data generating process (DGP),

$$\Delta y_t = a \Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (4.1a)$$

$$\Delta z_t = u_{2t} \quad (4.1b)$$

where

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim IIN \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & s^2 \end{bmatrix} \right). \quad (4.1c)$$

Our experiment is a design with  $a = 0.0, 0.5, 1$  (contemporaneous correlation),  $b = 0.0$  (no cointegration),  $-0.05, -0.1, -0.25, -0.5, -0.75$  (cointegration),  $s = 1, 6, 16$ ,  $N = 25, 50, 100, 200, 500, 1000$  (sample size).

To obtain the empirical size of the test we simulate the  $y_t$  and  $z_t$  series following the DGP (4.1a)–(4.1c) with  $b = 0$  and we estimate equation (3.1). The lower 5% tail of the distribution of the bootstrap  $t(\hat{b})$  statistic is the empirical critical value. The empirical power of the test is calculated analogously by simulating the series with the other values of  $b$  ( $b \neq 0$ ), and obtaining the percentages of rejections. For simplicity and without loss of generality, we consider the cointegrating vector as known,  $(1, -\alpha) = (1, -1)$ .

The Monte Carlo results of the bootstrap procedures are based on  $M = 10000$  replications. In each Monte Carlo replication, the bootstrap critical value is calculated by generating  $NB = 200$  bootstrap replications of the test statistic. In all of the experiments, the first 50 observations of the simulated series are dropped to generate random initial conditions.

Table 1: Size of the bootstrap test. No breaks. No intercept. Model (4.1a)–(4.1c)

N	a	$S_1, t_1$			$S_2, t_2$			$S_3, t_2$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0.0	5.7	5.2	5.3	4.9	5.2	5.4	5.2	5.2	5.5
	0.5	6.4	5.2	5.3	5.0	5.1	5.3	5.1	5.1	5.3
	1.0	6.6	6.6	6.6	5.2	5.2	5.2	5.1	5.1	5.1
50	0.0	5.6	5.0	5.1	5.0	5.0	5.1	4.9	4.9	5.1
	0.5	6.2	5.1	5.0	5.1	5.0	5.1	5.1	4.8	5.0
	1.0	7.0	7.0	7.0	5.4	5.4	5.4	5.3	5.3	5.3
100	0.0	5.8	5.1	5.1	4.9	5.0	5.1	4.9	5.1	5.2
	0.5	6.3	5.1	5.1	4.9	4.9	5.1	4.9	4.9	5.2
	1.0	6.6	6.6	6.6	5.3	5.3	5.3	5.3	5.3	5.3
200	0.0	5.6	4.8	4.7	4.8	4.8	4.7	4.9	4.8	4.8
	0.5	5.9	5.0	4.7	4.6	4.9	4.7	4.5	4.9	4.7
	1.0	6.4	6.4	6.4	4.9	4.9	4.9	5.0	5.0	5.0

We start by estimating the same model as Mantalos and Shukur (1998), equation (4.1a), and the results, are included in Table 1. Notice that this model does not include an intercept in the equation, contrary to the model we estimate in the rest of the paper. In this case we must modify the resampling schemes explained in Section 3 by centering the estimated residuals before doing the bootstrap. Mantalos and Shukur (1998), using a parameter

space which is a subset of ours, obtained that the size of the test was around its nominal value with sample sizes as small as 20 observations the bootstrap  $t$ -statistic obtained from every regression with resampling scheme  $S_2 - t_2$ , with 10000 Monte Carlo replications and based on  $NB = 200$  bootstrap samples. As we can see in Table 1, our conclusion is similar for  $S_2, t_2$ , and  $S_3, t_2$ . However, for the resampling scheme  $S_1, t_1$  we need somewhat larger sample sizes to get around its nominal value (5%). We also find that the results are not sensitive to the choice of  $NB$ .

In most of the Monte Carlo simulations it is common to include a constant term in the regression in order to make the results insensitive to initial conditions of the variables. Also, in most of the empirical applications a constant term is included to take care of the mean values of the regression variables  $y_t$  and  $z_t$ . Does it make any difference to include an intercept in bootstrap analysis? To answer that question, we estimate the following ECM model,

$$\Delta y_t = c + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_t \quad (4.2)$$

and the simulation results are included in Table 2(a). As we can see, the inclusion of an intercept in the model creates important size distortions, especially for  $S_1, t_1$ , and  $S_2, t_2$ , getting sizes of the test values up to around 25%. In the case  $S_3, t_2$ , the largest departure from the 5% nominal size of the test is around 6% for sample size equal to 25. This result provides strong evidence in favor of the bootstrap method based on  $S_3, t_2$ .

We estimate the extended ECM model

$$\Delta y_t = c + a\Delta z_t + b[y_{t-1} - z_{t-1}] + d[y_{t-2} - z_{t-2}] + u_t, \quad (4.3)$$

to see whether the bootstrap ECM test is robust to the presence of a constant term in the model. The results are displayed in Table 2(b) and show that with all resampling schemes we get sizes closer to the nominal value. The largest size distortion is reduced from 24% to 11%, and once again,  $S_3, t_2$  clearly dominates. With  $S_3, t_2$  we never get a size of the test larger than 6.1% for sample sizes of 25 observations. This is an interesting result. We have shown that, even in a model without breaks, if we are uncertain about including an intercept in the model, we should use the extended ECM model (4.3) in order to make sure that the sizes of the test are close to the nominal values.

On the other hand, when we omit the intercept in the case that DGP includes it, the bootstrap test is not adequate, even with the extended ECM model, since the model is misspecified. Therefore in applied work

Table 2: Size of the bootstrap test. Simulated model (4.1a)–(4.1c). No breaks.

(a) Estimated model (4.2). ECM

N	a	$S_1, t_1$			$S_2, t_2$			$S_3, t_2$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0.0	12.1	5.4	4.9	15.6	5.7	5.0	5.3	5.1	4.9
	0.5	17.2	6.7	5.2	21.5	7.5	5.3	5.5	5.1	5.0
	1.0	21.2	21.2	21.2	26.3	26.3	26.3	6.2	6.2	6.2
50	0.0	11.8	5.4	5.1	14.8	5.6	5.0	5.3	5.0	5.0
	0.5	16.4	6.6	5.3	20.6	7.1	5.3	5.5	4.9	5.0
	1.0	20.4	20.4	20.4	25.5	25.5	25.5	5.9	5.9	5.9
100	0.0	12.1	5.5	4.9	14.0	5.4	4.9	5.0	4.8	4.9
	0.5	16.7	6.7	5.2	20.4	6.8	5.3	5.5	4.9	4.9
	1.0	19.9	19.9	19.9	23.9	23.9	23.9	6.0	6.0	6.0
200	0.0	12.2	5.3	4.7	13.4	5.3	4.6	5.0	4.8	4.6
	0.5	16.3	6.4	5.1	18.9	6.6	4.9	5.4	4.9	4.8
	1.0	20.5	20.5	20.5	23.4	23.4	23.4	5.9	5.9	5.9

(b) Estimated model (4.3). Extended ECM

N	a	$S_1, t_1$			$S_2, t_2$			$S_3, t_2$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0.0	8.9	5.1	5.0	8.1	5.0	4.8	5.2	5.1	5.0
	0.5	10.8	5.8	4.2	8.3	5.4	4.1	5.7	5.0	5.0
	1.0	13.5	13.5	13.4	11.1	11.3	11.2	6.1	6.1	6.1
50	0.0	7.1	4.8	4.7	6.9	4.7	4.6	4.8	5.3	5.1
	0.5	9.1	5.0	3.8	8.1	4.9	3.8	5.3	5.1	5.2
	1.0	10.6	10.6	10.6	9.8	9.8	9.8	5.7	5.7	5.7
100	0.0	6.3	5.0	4.9	6.1	4.9	5.0	4.8	4.7	4.8
	0.5	7.6	4.8	4.1	7.1	4.6	4.1	5.2	4.6	4.7
	1.0	8.5	8.5	8.5	8.1	8.2	8.2	5.2	5.2	5.2
200	0.0	6.0	5.2	5.2	5.9	5.2	5.2	4.9	4.8	4.9
	0.5	6.6	4.5	4.2	6.5	4.5	4.1	5.1	4.6	4.8
	1.0	7.1	7.1	7.1	7.0	7.0	7.0	5.2	5.2	5.2

we recommend to include an intercept and estimate ECM model (4.3) instead of (4.2).

Bias correction improves the results only if we calculate the  $t$ -stat using the empirical standard deviation of the bootstrap estimates instead of the

usual  $t$ -stat, but the gain is really small, and therefore we will not report the results (results are available upon request).

Table 3: Power of the bootstrap test,  $S_3, t_2$ . Simulated model (4.1a)–(4.1c).  $b = -0.2$ . No breaks.

Estimated model		ECM, eq.(4.2)			Extended ECM, eq. (4.3)		
N	$a$	s=1	s=6	s=16	s=1	s=6	s=16
25	0.0	100.0	100.0	100.0	23.6	99.4	100.0
	0.5	100.0	100.0	100.0	16.6	77.8	99.9
	1.0	100.0	100.0	100.0	13.5	13.5	13.5
50	0.0	100.0	100.0	100.0	49.0	100.0	100.0
	0.5	100.0	100.0	100.0	32.0	98.3	100.0
	1.0	100.0	100.0	100.0	25.4	25.4	25.4
100	0.0	100.0	100.0	100.0	80.9	100.0	100.0
	0.5	100.0	100.0	100.0	61.1	100.0	100.0
	1.0	100.0	100.0	100.0	50.5	50.5	50.5
200	0.0	100.0	100.0	100.0	98.2	100.0	100.0
	0.5	100.0	100.0	100.0	90.8	100.0	100.0
	1.0	100.0	100.0	100.0	83.5	83.5	83.5

Given our previous evidence, we will concentrate on the bootstrap procedure  $S_3, t_2$ , that considers the resampling residuals estimated under  $H_0 : b = 0$ . Our objective is to analyze the power of the test for models (4.2) and (4.3). The power of the test is obtained from the percentage of rejections of  $H_0$  with  $NB = 5000$  bootstrap samples. When there are no breaks, the results are displayed in Table 3. The highest power is obtained for model (4.2). When there is no common factor ( $a \neq 1$ ), model (4.3) has high power for moderate sample sizes ( $N > 50$ ). However, when the common factor restriction holds, we need more than 200 observations in order to get a power of 90% or above. Figure 4 shows the power of the bootstrap ECM test,  $t(b)$ , for the usual ECM model, equation (4.2), and the extended ECM model, equation (4.3), with  $a = 0.0$ ,  $s = 1$ , and different values of  $b$ . The power of the standard ECM test is never smaller than that of the extended ECM test, but for values of  $b$  lower than  $b = -0.1$ , both tests perform equally well (power close to 100%).

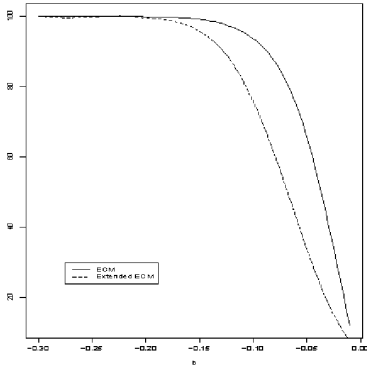


Figure 4(a):N=25

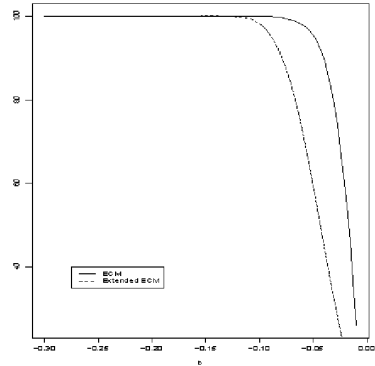


Figure 4(b):N=50

Figure 4: Power of the ECM and Extended ECM bootstrap test.

## 5 Monte Carlo simulation experiment II: Co-breaks

### 5.1 ECM with simultaneous co-breaks

Our data generating process (DGP) is based on Arranz and Escribano (2000a), which is an extension 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, with Granger causality in only one direction ( $z \rightarrow y$ ), and several structural breaks in the strongly exogenous variable,  $\Delta z_t$ , when the parameters of interest are those of equation (5.1a). Consider the following model,

$$\Delta y_t = a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1t} \quad (5.1a)$$

$$\Delta z_t = sD_{jt} + u_{2t} \quad (5.1b)$$

where

$$\begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \sim IIN \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & s^2 \end{bmatrix} \right). \quad (5.1c)$$

Notice that under simultaneous co-breaking there are breaks in the marginal process of  $\Delta z_t$ , equation (4.1b), but not in the conditional process,

equation (4.1a). We allow three kinds of dummy variables in order to simulate a single break in the deterministic trend (segmented trends), at two different points ( $N/4$  or  $N/2$ ) where  $N$  is the sample size,

$$D_{1t} = \begin{cases} 1 & t \geq N/4 \\ 0 & \text{otherwise} \end{cases} \quad D_{2t} = \begin{cases} 1 & t \geq N/2 \\ 0 & \text{otherwise} \end{cases}$$

and two breaks at points  $N/4$  and  $3N/4$ .

$$D_{3t} = \begin{cases} 1 & N/4 \leq t \leq 3N/4 \\ 0 & \text{otherwise} \end{cases}$$

Based on (5.1a)–(5.1c), the size of the test, see Table 4 ( $b = 0$ ), and the power ( $b \neq 0$ ) of the test are obtained by Monte Carlo simulation experiments. Without loss of generality, we take  $\sigma_1^2 = 1$  and  $\alpha = 1$ . Thus, the experimental design variables are the parameters  $a, b, s$  (notice that  $\sigma_2 = s$ ), where  $s$  is the coefficient of the dummy variable  $D_{j,t}$  of (5.1b), and the sample size,  $N$ .

The sizes of the bootstrap ECM test when the series exhibit simultaneous co-breaks and we estimate the usual ECM model (4.2) are large, especially when the COMFAC restriction holds for sample sizes of 200 observations or less, see Table 4(a). This problem is again caused by the inclusion of an intercept in our estimated regression. Once again, this size distortion problem is diminished by using the  $S_3, t_2$  resampling scheme<sup>2</sup>

For the extended ECM model (4.3), the sizes of the test are displayed in Table 4(b). We can see that the sizes of the test are around its nominal value in the case that the COMFAC condition ( $a = 1$ ) holds, i.e. there is a marginal improvement in the extended ECM model (4.3) over the ECM model (4.2).

In terms of power, under simultaneous co-breaks, model (4.2) performs better than model (4.3), see Table 5. Both procedures behave well under moderate sample sizes, say  $N \geq 100$ . As expected, the lowest power is obtained when  $a = 1$ , due to the common factor restriction. Recall that when  $a = 1$  the limiting distribution of  $t(\hat{b})$  is the Dickey–Fuller distribution and, it is thus not affected by the type of dummies ( $D_1, D_2, D_3$ ) analyzed

---

<sup>2</sup>We show only the tables corresponding to the  $S_3, t_2$  procedure. The full set of tables is available upon request



Table 4: Empirical size, bootstrap procedure  $S_3, t_2$ . Simultaneous co-breaks,  $t$ -stat.

(a) Estimated model (4.2). ECM model.

N	$\alpha$	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.4	5.1	5.1	5.3	5.2	5.1	5.6	5.2	5.2
	0.5	5.3	5.3	5.1	5.2	5.2	5.2	5.3	5.1	5.2
	1	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7	4.7
50	0	4.9	4.9	4.9	5.3	5.0	4.9	4.9	4.6	4.8
	0.5	5.2	4.9	4.9	4.9	5.1	4.9	5.2	4.6	4.7
	1	4.8	4.8	4.8	4.8	4.8	4.8	4.6	4.6	4.6
100	0	4.3	4.5	4.5	5.1	4.8	4.8	5.1	4.6	4.6
	0.5	4.8	4.4	4.5	4.9	4.8	4.8	4.8	4.7	4.6
	1	4.9	4.9	4.9	4.8	4.8	4.8	4.8	4.8	4.8
200	0	4.7	4.9	5.0	5.0	5.1	5.0	5.1	4.9	4.9
	0.5	5.1	4.9	4.9	4.8	5.0	5.0	4.6	4.9	4.8
	1	4.8	4.8	4.8	4.8	4.8	4.8	4.8	4.8	4.8

(b) Estimated model (4.3). Extended ECM Model

N	$\alpha$	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.4	5.1	5.1	5.2	5.2	5.2	5.2	5.1	4.9
	0.5	5.1	5.1	5.1	5.2	5.3	5.2	5.1	5.1	5.0
	1	6.2	6.2	6.2	6.2	6.2	6.2	6.1	6.1	6.1
50	0	4.9	4.9	5.1	4.7	5.2	5.2	4.8	5.0	5.1
	0.5	4.7	5.0	4.9	4.6	5.1	5.2	4.6	4.9	5.0
	1	5.8	5.8	5.8	5.8	5.8	5.8	5.9	5.9	5.9
100	0	4.5	4.6	4.8	4.7	4.9	4.9	4.4	4.7	4.7
	0.5	4.5	4.6	4.7	4.3	4.8	4.9	4.5	4.4	4.6
	1	5.1	5.1	5.1	5.3	5.3	5.3	5.3	5.3	5.3
200	0	4.6	4.8	4.8	4.8	4.7	4.9	4.8	4.7	4.8
	0.5	5.0	4.9	4.8	4.8	4.6	4.8	4.9	4.7	4.7
	1	5.4	5.4	5.4	5.4	5.4	5.4	5.3	5.3	5.3

Table 5: Power, bootstrap procedure  $S_3, t_2$ . Simultaneous co-breaks,  $t$ -stat,  $b = -0.2$

(a) Estimated model (4.2). ECM model

N	$\alpha$	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	61.7	99.9	100.0	60.8	100.0	100.0	56.2	99.9	100.0
	0.5	25.4	99.5	100.0	25.5	99.3	100.0	23.9	98.8	100.0
	1	14.6	14.6	14.6	15.1	15.1	15.1	14.8	14.8	14.8
50	0	96.9	100.0	100.0	97.8	100.0	100.0	96.8	100.0	100.0
	0.5	68.6	100.0	100.0	70.6	100.0	100.0	66.7	100.0	100.0
	1	34.8	34.8	34.8	34.4	34.4	34.4	34.9	34.9	34.9
100	0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	0.5	99.3	100.0	100.0	99.3	100.0	100.0	99.2	100.0	100.0
	1	86.2	86.2	86.2	85.7	85.7	85.7	85.5	85.5	85.5
200	0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	0.5	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
	1	100.0	100.0	100.0	99.9	99.9	99.9	100.0	100.0	100.0

(b) Estimated model (4.3). Extended ECM model.

N	$\alpha$	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	26.1	99.9	100.0	27.5	99.8	100.0	26.8	99.9	100.0
	0.5	16.5	83.3	100.0	17.4	84.4	100.0	16.9	85.6	100.0
	1	13.6	13.6	13.6	14.4	14.4	14.4	14.4	14.4	14.4
50	0	53.5	100.0	100.0	55.2	100.0	100.0	54.2	100.0	100.0
	0.5	33.0	99.2	100.0	35.2	99.4	100.0	33.3	99.4	100.0
	1	25.6	25.6	25.6	25.8	25.8	25.8	25.6	25.6	25.6
100	0	84.9	100.0	100.0	86.3	100.0	100.0	86.1	100.0	100.0
	0.5	64.6	100.0	100.0	65.4	100.0	100.0	65.9	100.0	100.0
	1	50.3	50.3	50.3	50.1	50.1	50.1	50.2	50.2	50.2
200	0	98.9	100.0	100.0	99.0	100.0	100.0	99.1	100.0	100.0
	0.5	91.8	100.0	100.0	92.4	100.0	100.0	92.1	100.0	100.0
	1	83.5	83.5	83.5	83.6	83.6	83.6	83.3	83.3	83.3

nor by the size of the jump under simultaneous co-breaking. When  $a \neq 1$ , the power of the test increases with the jump ( $s$ ), since it increases the signal to noise ratio  $q$ , where  $q = (1 - a)s/\sigma_1$ .

Finally, Figure 5 shows that the power of the bootstrap test is high (close to 100%) for both models, the ECM, equation (4.2), and the extended ECM, equation (4.3), when the parameter  $b$  takes values lower than  $b = -0.15$ , and there are simultaneous co-breaks.

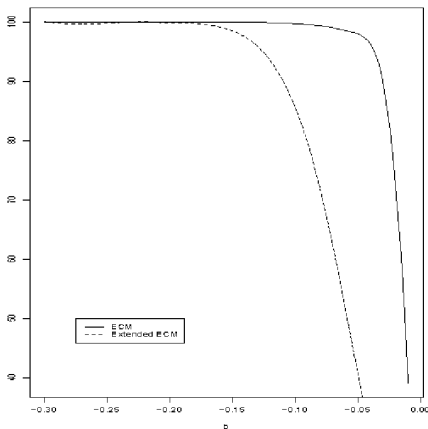


Figure 5: Power of the test with simultaneous co-breaks, for  $D3$ ,  $N = 25$ ,  $a = 0$ , and  $s = 6$ .

## 5.2 ECM under partial co-breaks

Partial co-breaking is characterized by having breaks in the marginal process  $\Delta z_t$ , equation (2.1d), that affects the conditional process, equation (2.1a), from the two alternative sources that affect  $c_t$ , equation (2.1c): either through  $\Delta\mu_{y,t} - a\Delta\mu_{z,t}$  (not co-breaking in differences) or through  $(\mu_{y,t-1} - \alpha\mu_{z,t-1})$  (not co-breaking in levels).

Our DGP is based on the following time varying error correction model,

$$\Delta y_t = c_t + a\Delta z_t + b(y_{t-1} - z_{t-1}) + u_{1t} \quad (5.2a)$$

$$\Delta z_t = \Delta\mu_{z,t} + u_{2t} \quad (5.2b)$$

$$c_t = \Delta\mu_{y,t} - a\Delta\mu_{z,t} - b(\mu_{y,t-1} - \alpha\mu_{z,t-1}) \quad (5.2c)$$

$$\Delta\mu_{z,t} = sD_{j,t} \quad (5.2d)$$

where  $D_{j,t}$  are the dummies used in our previous simulation exercise, see Section 5.1. To generate co-breaks in differences and not in levels, we make  $\Delta\mu_{y,t} - a\Delta\mu_{z,t} = c_d = 0.5$ . When we want to generate a set of series with co-breaks in levels and not in differences, we make  $\Delta\mu_{y,t} - \alpha\Delta\mu_{z,t} = 0$  (recall that we take  $\alpha = 1$ ).

### 5.2.1 Co-break in differences, but not in levels

In this case, equation (5.2a) can be written as

$$\Delta y_t = c_m + bc_d t - b(a - \alpha)s \sum_{i=1}^t D_{j,i} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (5.3)$$

The extended error correction model (4.3) approximates equation (5.3) well if

$$(y_{t-2} - \alpha z_{t-2}) \cong b_1 t + b_2 s \sum_{i=1}^t D_{j,i}, \quad (5.4)$$

and this is likely possible when there is no co-break in levels, as we could see in Figure 3 by comparing Figures 3(e) and 3(f).

The Monte Carlo results of the bootstrap ECM test based on  $S_3, t_2$  are reported in Table 6. Table 6(a) shows that there are no size distortions, but Table 6(b) shows important reductions in power relative to simultaneous co-breaking. The worst case occurs with the power close to zero under the COMFAC restriction ( $a = 1$ ). We can see that in some cases the power of the bootstrap ECM decreases when the sample size is increased.

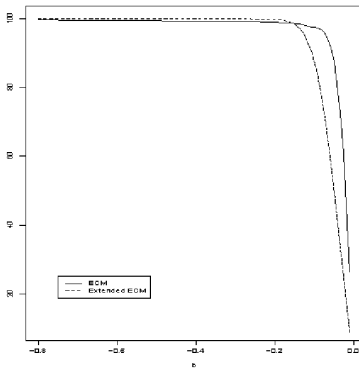


Figure 6(a):No break

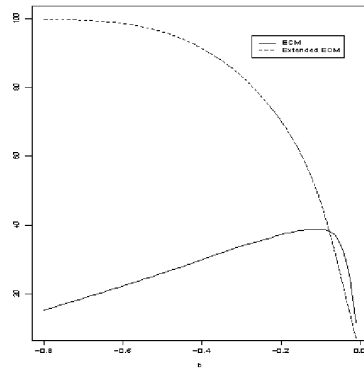


Figure 6(b):D3

Figure 6: Power of the ECM and Extended ECM bootstrap test. Co-breaks in differences, not in levels.  $N = 50$ .

Table 6: Bootstrap procedure  $S_3, t_2$ . Co-breaking in differences, not in levels. Estimated model (4.2), ECM model.

(a) Empirical size

N	$a$	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.5	5.2	5.1	5.6	5.3	5.3	5.2	5.1	4.9
	0.5	5.5	5.2	5.1	5.4	5.4	5.3	5.1	5.0	5.1
	1	4.6	4.6	4.6	4.8	4.8	4.8	4.7	4.7	4.7
50	0	4.6	4.8	5.1	4.9	5.3	5.2	4.6	5.1	5.1
	0.5	5.2	5.0	5.0	4.9	5.2	5.2	5.0	5.0	5.0
	1	5.1	5.1	5.1	5.1	5.1	5.1	4.9	4.9	4.9
100	0	4.5	4.6	4.8	4.8	5.0	4.9	4.6	4.7	4.7
	0.5	4.8	4.6	4.7	4.9	4.7	4.9	4.6	4.4	4.6
	1	4.6	4.6	4.6	4.7	4.7	4.7	4.7	4.7	4.7
200	0	4.7	4.8	4.8	4.8	4.7	4.9	4.8	4.7	4.8
	0.5	5.1	4.9	4.8	4.7	4.5	4.8	5.2	4.7	4.7
	1	4.9	4.9	4.9	4.9	4.9	4.9	5.0	5.0	5.0

(b) Power of the test,  $b = -0.2$

N	$a$	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	10.7	39.1	43.3	17.9	46.4	50.3	16.8	37.6	41.8
	0.5	15.6	28.4	40.9	6.6	35.6	48.5	8.8	28.7	40.0
	1	2.0	2.0	2.0	2.5	2.5	2.5	1.5	1.5	1.5
50	0	9.2	39.5	42.0	15.6	46.3	49.2	13.4	40.0	42.2
	0.5	21.3	31.8	40.3	2.5	38.7	47.8	7.4	33.2	41.3
	1	0.4	0.4	0.4	0.9	0.9	0.9	0.3	0.3	0.3
100	0	9.1	34.0	36.3	12.8	39.9	42.0	11.2	34.3	35.8
	0.5	9.2	28.1	34.8	0.9	34.0	41.2	2.5	29.5	35.5
	1	0.1	0.1	0.1	0.2	0.2	0.2	0.0	0.0	0.0
200	0	7.2	34.1	36.2	9.6	38.0	40.3	9.1	34.0	36.0
	0.5	3.6	28.8	35.2	0.8	31.8	39.3	0.9	28.7	34.9
	1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table 7: Bootstrap procedure  $S_3, t_2$ . Co-breaks in differences, not in levels. Estimated model (4.3). Extended ECM model.

(a) Empirical size

N	$a$	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.5	5.2	5.1	5.6	5.3	5.3	5.2	5.1	4.9
	0.5	5.5	5.2	5.1	5.4	5.4	5.3	5.1	5.0	5.1
	1	4.6	4.6	4.6	4.8	4.8	4.8	4.7	4.7	4.7
50	0	4.6	4.8	5.1	4.9	5.3	5.2	4.6	5.1	5.1
	0.5	5.2	5.0	5.0	4.9	5.2	5.2	5.0	5.0	5.0
	1	5.1	5.1	5.1	5.1	5.1	5.1	4.9	4.9	4.9
100	0	4.5	4.6	4.8	4.8	5.0	4.9	4.6	4.7	4.7
	0.5	4.8	4.6	4.7	4.9	4.7	4.9	4.6	4.4	4.6
	1	4.6	4.6	4.6	4.7	4.7	4.7	4.7	4.7	4.7
200	0	4.7	4.8	4.8	4.8	4.7	4.9	4.8	4.7	4.8
	0.5	5.1	4.9	4.8	4.7	4.5	4.8	5.2	4.7	4.7
	1	4.9	4.9	4.9	4.9	4.9	4.9	5.0	5.0	5.0

(b) Power of the test,  $b = -0.2$

N	$a$	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	10.5	48.1	64.7	15.4	48.5	62.1	14.4	52.3	67.3
	0.5	14.7	25.8	55.6	10.4	27.9	54.7	11.6	29.2	58.8
	1	6.6	6.6	6.6	6.9	6.9	6.9	5.9	5.9	5.9
50	0	17.0	68.8	79.4	21.9	65.8	76.2	19.4	69.4	78.1
	0.5	25.0	45.2	74.7	12.8	45.7	70.8	14.2	49.1	74.0
	1	9.2	9.2	9.2	9.3	9.3	9.3	8.9	8.9	8.9
100	0	31.5	89.8	93.4	33.2	86.2	91.2	33.1	85.8	89.4
	0.5	30.7	74.5	91.5	22.2	69.8	88.9	21.4	72.2	87.5
	1	18.0	18.0	18.0	18.3	18.3	18.3	17.5	17.5	17.5
200	0	56.5	98.8	99.3	57.5	98.2	99.4	56.7	97.0	98.4
	0.5	48.4	93.9	99.2	42.9	92.8	98.9	41.1	92.2	97.4
	1	34.7	34.7	34.7	34.2	34.2	34.2	34.4	34.4	34.4

As expected, the results obtained by the extended ECM test represent an important improvement, see Table 7. Table 7(a) shows the size of the bootstrap extended ECM test and they are always close to the nominal size. Table 7(b) shows important increases in power relative to the usual ECM test. The worst power results are obtained once again when there is a common factor restriction ( $a = 1$ ), but the power of the test always increases with the sample size.

Figure 6 shows the power of the bootstrap ECM and bootstrap extended ECM test for a sample size of 50 observations with no break and with a double break in the trend (dummy variable  $D_{3,t}$ ). In both cases, the power of the extended ECM test is higher than the power of the usual ECM test for values of  $b$  lower than  $b = -0.15$ . Figure 6(b) shows that the power of the of the ECM test decreases with  $|b|$ , while the power of the extended ECM test test display the usual behavior, increasing with  $|b|$ . These results show that the approximation (5.4) is working well.

### 5.2.2 Co-break in levels, but not in differences

In this case, equation (5.2a) becomes

$$\Delta y_t = -bc_t + (\alpha - a)sD_{j,t} + a\Delta z_t + b(y_{t-1} - \alpha z_{t-1}) + u_{1,t} \quad (5.5)$$

In the case of co-breaks in levels but not in differences, the extended error correction model might not be a good approximation, since

$$(y_{t-2} - \alpha z_{t-2}) \approx (\alpha - a)sD_{j,t},$$

as we could see in Figure 2 by comparing Figures 2(e) and 2(f).

Table 8 shows the important size distortions obtained with the bootstrap ECM test based on model (4.2). The oversize can reach levels near 47% when the nominal size of the test is 5%.

If we do the bootstrap  $t$ -test based on the extended ECM model, equation (4.3), the results are reported in Table 9. Table 9(a) shows that the largest size distortions are reduced form 47% to 18% but is is still oversized. This problem does not exist when the COMFAC restriction holds ( $a = 1$ ), because we have full co-breaking in that particular case, see Table 9(b). Figure 7 shows the power of the bootstrap extended ECM test as a function of the parameter  $b$ . It shows that for values of  $b$  lower than  $b = -0.15$  the

Table 8: Empirical size, bootstrap procedure.  $S_3, t_2$ . Co-breaking in levels, not in differences. Estimated model (4.2). ECM model.

N	a	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	13.1	31.4	32.3	18.0	37.2	37.9	10.8	21.5	22.5
	0.5	8.2	27.5	31.8	9.3	35.1	37.7	7.5	19.3	21.9
	1	6.4	6.4	6.4	6.6	6.6	6.6	6.4	6.4	6.4
50	0	21.5	38.7	39.4	30.0	42.7	43.1	16.0	28.5	29.2
	0.5	11.6	36.0	39.0	13.7	41.8	42.9	9.9	26.3	28.8
	1	5.7	5.7	5.7	5.9	5.9	5.9	5.8	5.8	5.8
100	0	32.4	43.5	44.1	40.4	46.2	46.4	23.4	35.1	36.0
	0.5	18.3	42.0	43.8	23.7	45.2	46.1	14.1	33.0	35.3
	1	6.0	6.0	6.0	6.1	6.1	6.1	6.0	6.0	6.0
200	0	40.1	45.7	45.9	45.6	47.2	47.3	32.3	40.5	41.0
	0.5	28.4	45.2	45.7	36.3	46.9	47.3	21.1	39.5	40.8
	1	5.8	5.8	5.8	5.7	5.7	5.7	5.9	5.9	5.9

power is close to 100%. It also shows that the power with the sample size ranging from 25 to 50 observations, the power is highly increased when  $b$  is larger than  $-0.15$ .

### A solution based on imposing the COMFAC restriction.

Following Arranz and Escribano (2000b), equation (5.5) becomes

$$(\Delta y_t - \alpha \Delta z_t) = -bc_l + b(y_{t-1} - z_{t-1}) + v_{1,t} \quad (5.6)$$

where  $v_{1,t} = (a - \alpha)(\Delta z_t - sD_{j,t}) + u_{1,t} = (a - \alpha)u_{2,t} + u_{1,t}$  and therefore the breaks in differences disappear from the testing equation. Notice that imposing the COMFAC restriction means that the kind of break we have ( $D_1, D_2, D_3$ ), is irrelevant.

We can therefore use the following two equations to perform the ECM test based on the usual ECM model or the extended ECM model:

$$(\Delta y_t - \alpha \Delta z_t) = c + b(y_{t-1} - z_{t-1}) + \epsilon_t \quad (5.7)$$

$$(\Delta y_t - \alpha \Delta z_t) = c + b(y_{t-1} - z_{t-1}) + d(y_{t-2} - z_{t-2}) + \epsilon_t \quad (5.8)$$



Table 9: Bootstrap procedure  $S_3, t_2$ . Co-breaks in levels, not in differences. Estimated model (4.3). Extended ECM model.

(a) Empirical size

N	a	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	6.4	11.5	12.6	7.7	17.2	18.8	7.3	12.7	13.2
	0.5	6.1	9.3	12.1	6.6	13.6	17.9	6.4	11.1	12.9
	1	6.2	6.2	6.2	6.2	6.2	6.2	6.1	6.1	6.1
50	0	6.3	12.8	14.3	7.9	16.7	18.5	7.5	12.0	12.7
	0.5	5.4	10.3	13.5	6.3	13.4	17.2	6.2	10.6	12.4
	1	5.8	5.8	5.8	5.8	5.8	5.8	5.9	5.9	5.9
100	0	6.8	12.9	14.6	7.8	15.8	17.6	7.2	11.5	12.1
	0.5	5.9	10.2	13.7	6.5	12.4	16.7	6.3	9.9	11.8
	1	5.1	5.1	5.1	5.3	5.3	5.3	5.3	5.3	5.3
200	0	7.0	13.1	14.7	7.8	15.1	16.8	7.0	11.2	11.8
	0.5	6.1	9.9	13.7	7.0	11.9	15.8	6.3	9.9	11.6
	1	5.4	5.4	5.4	5.4	5.4	5.4	5.3	5.3	5.3

(b) Power of the test,  $b = -0.2$

N	a	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	23.2	58.8	61.5	24.7	51.2	52.5	23.5	50.4	52.6
	0.5	16.5	48.4	59.7	17.8	45.4	52.0	17.6	44.5	52.1
	1	13.6	13.6	13.6	14.4	14.4	14.4	14.4	14.4	14.4
50	0	46.2	89.2	92.3	47.3	80.1	82.9	46.0	81.2	82.8
	0.5	33.3	80.3	90.8	35.5	74.4	81.8	34.2	73.4	82.1
	1	25.6	25.6	25.6	25.8	25.8	25.8	25.6	25.6	25.6
100	0	77.5	99.8	99.9	77.8	99.2	99.7	76.9	98.8	99.2
	0.5	62.2	98.4	99.9	62.5	96.9	99.5	62.7	96.6	99.1
	1	50.3	50.3	50.3	50.1	50.1	50.1	50.2	50.2	50.2
200	0	96.8	100.0	100.0	96.6	100.0	100.0	97.2	100.0	100.0
	0.5	90.5	100.0	100.0	90.4	100.0	100.0	90.7	100.0	100.0
	1	83.5	83.5	83.5	83.6	83.6	83.6	83.3	83.3	83.3

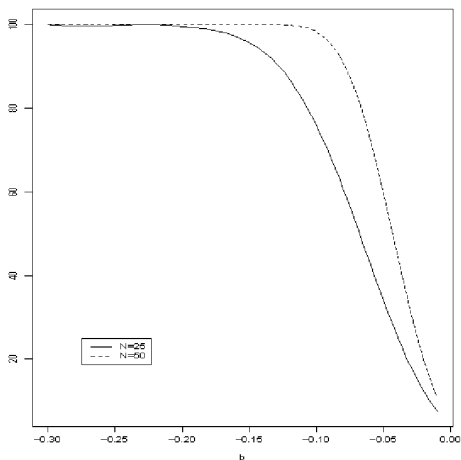


Figure 7: Power of the bootstrap test. Co-breaks in levels, not in differences,  $D3$ . Extended ECM model (4.3).

Table 10 summarizes the main results of the ECM test for the case of co-breaks in levels, but not in differences, and impose the COMFAC restriction. As we can see the size of the test when using (5.7) is bigger than 6% when the sample size is as small as 25 observations, and even when the sample size is 200 there are some cases when the size of the bootstrap test is close to 6%, see Table 10(a).

The size of the test is closer to the nominal size when we use equation (5.8). As we can see in Table 10(b), the size of the test is slightly larger than 6% only when the sample size is 25 and the COMFAC restriction holds. Taking samples of size equal or larger than 50 we get a size of the bootstrap test which is not significantly different from the nominal size.

Notice that this happens only in the case that there is co-breaks in levels, but not in differences. We also tried some other specifications such as simultaneous co-breaking, finding this approach inadequate (Results are available upon request).

Table 10: Bootstrap procedure  $S_3, t_2$ . Co-breaks in levels, not in differences, imposing COMFAC restriction.

(a) ECM model.

N	a	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.3	5.6	5.5	5.3	5.6	5.5	5.3	5.6	5.5
	0.5	5.4	5.6	5.6	5.4	5.6	5.6	5.4	5.6	5.6
	1	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3	6.3
50	0	5.7	5.5	5.6	5.7	5.5	5.6	5.7	5.5	5.6
	0.5	6.0	5.2	5.7	6.0	5.2	5.7	6.0	5.2	5.7
	1	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0
100	0	5.6	5.4	5.7	5.6	5.4	5.7	5.6	5.4	5.7
	0.5	5.8	5.5	5.5	5.8	5.5	5.5	5.8	5.5	5.5
	1	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9	5.9
200	0	5.9	5.7	5.4	5.9	5.7	5.4	5.9	5.7	5.4
	0.5	5.8	5.8	5.6	5.8	5.8	5.6	5.8	5.8	5.6
	1	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8	5.8

(b) Extended ECM model

N	a	$D_1$			$D_2$			$D_3$		
		s=1	s=6	s=16	s=1	s=6	s=16	s=1	s=6	s=16
25	0	5.2	5.8	5.8	5.2	5.8	5.8	5.2	5.8	5.8
	0.5	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7	5.7
	1	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0	6.0
50	0	5.3	5.5	5.4	5.3	5.5	5.4	5.3	5.5	5.4
	0.5	5.3	5.4	5.4	5.3	5.4	5.4	5.3	5.4	5.4
	1	5.6	5.6	5.6	5.6	5.6	5.6	5.6	5.6	5.6
100	0	5.1	5.3	5.3	5.1	5.3	5.3	5.1	5.3	5.3
	0.5	5.2	5.1	5.4	5.2	5.1	5.4	5.2	5.1	5.4
	1	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2	5.2
200	0	5.1	4.9	5.1	5.1	4.9	5.1	5.1	4.9	5.1
	0.5	5.4	5.1	5.1	5.4	5.1	5.1	5.4	5.1	5.1
	1	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3	5.3

## 6 Conclusions

We have analyzed the behavior of different bootstrap ECM-tests under structural breaks . In the cases of no breaks or with simultaneous co-breaks, we have found that the inclusion of a constant term in the regression test might lead to high size distortions in the usual bootstrap ECM test provided that there is not a constant term in the DGP. This problem is solved by using extended ECM models to calculate the bootstrap test. As a drawback, we have found that the only valid resampling scheme is what we call  $S_3, t_2$ . this procedure is based on resampling the estimated restricted model under the null hypothesis  $b = 0$ .

Furthermore, in the case of co-breaks in differences and not in levels, the usual bootstrap ECM test would render no power. This is also solved by using extended ECM models.

In the case of co-breaks in levels and not in differences, extended ECM models would bring the size of the bootstrap ECM test close to the nominal size, but keeping them still high. This problem is solved by imposing the COMFAC restriction.

## Acknowledgments

We are very grateful to an anonymous referee for all his insightful comments. Drafts of this paper were presented at the Bootstrap Workshop held at Aarhus University in 1999 and the International Symposium on Forecasting in Lisbon in 2000. The first author wishes to thank members of Dept. of Statistics and Econometrics, Universidad Carlos III de Madrid, and Dept. of Economics, University of Maryland at College Park.

## References

- ARRANZ, M. and ESCRIBANO, A. (2000a). Cointegration testing under structural breaks: A robust extended error correction model. *Oxford Bulletin of Economics and Statistics*, 61:23–52.

- ARRANZ, M. and ESCRIBANO, A. (2000b). COMFAC restriction in ECM tests under partial co-breaking. Working paper, Dept. of Statistics and Econometrics, Universidad Carlos III de Madrid.
- BERKOWITZ, J. and KILIAN, L. (2000). Recent developments in bootstrapping time series. *Econometric Reviews*, 19:1–48.
- BÜHLMANN, P. (1999). Bootstrap for time series. Research report, ETH, Zurich.
- CAMPOS, J., ERICSSON, N., and HENDRY, D. F. (1996). Cointegration tests in the presence of structural breaks. *Journal of Econometrics*, 70:187–220.
- CLEMENTS, M. P. and HENDRY, D. F. (1999). *Forecasting Non-Stationary Economic Time Series*. MIT Press, Cambridge.
- DAVIDSON, A. and HINKLEY, D. (1997). *Bootstrap Methods and Their Application*. Cambridge University Press, Cambridge.
- DOLADO, J. and LÜTKEPOHL, H. (1996). Making Wald tests work for cointegrated VAR systems. *Econometric Reviews*, 15:369–386.
- EFRON, B. and TIBSHIRANI, R. J. (1993). *An Introduction to the Bootstrap*, Vol. 57 of *Monographs on Statistics and Applied Probability*. Chapman and Hall, New York.
- HALL, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer Series in Statistics. Springer-Verlag, New York.
- HALL, P. (1994). Methodology and theory for the bootstrap. In R. Engle and D. McFadden, eds., *Handbook of Econometrics*, Vol. 4, Chap. 39, pp. 2340–281. North-Holland, Amsterdam.
- HALL, P. and WILSON, S. (1991). Two guidelines for bootstrap hypothesis testing. *Biometrika*, 47:757–762.
- HOROWITZ, J. L. (2000). The bootstrap. In J. J. Heckman and E. E. Leamer, eds., *Handbook of Econometrics*, Vol. 5, Chap. 52, pp. 3159–3228. North-Holland, Amsterdam.
- KIVIET, J. (1984). Bootstrap inference in lagged-dependent variable models. Working paper, University of Amsterdam.

- KREMERS, J., ERICSSON, N., and DOLADO, J. (1992). The power of cointegration tests. *Oxford Bulletin of Economics and Statistics*, 54:349–67.
- LI, H. and MADDALA, G. (1996). Bootstrapping time series models. *Econometric Reviews*, 15:115–158.
- LI, H. and MADDALA, G. S. (1997). Bootstrapping cointegrating regressions. *Journal of Econometrics*, 80:297–318.
- MANTALOS, P. and SHUKUR, G. (1998). Size and power of the error correction model cointegration test. A bootstrap approach. *Oxford Bulletin of Economics and Statistics*, 60:249–255.
- SHAO, J. and TU, D. (1995). *The Jackknife and the Bootstrap*. Springer Series in Statistics. Springer–Verlag, New York.
- TODA, H. and YAMAMOTO, T. (1995). Statistical inference in vector autor-regression with possibly integrated processes. *Journal of Econometrics*, 66:225–250.
- VAN GIERBERGEN, N. and KIVIET, J. (1993). A Monte Carlo comparison of asymptotic and various nonparametric inference procedures in first-order dynamic models. Discussion Paper TI 93–187, Tinbergen Institute, Amsterdam.
- VAN GIERBERGEN, N. and KIVIET, J. (1994). How to implement bootstrap hypothesis testing in static and dynamic regression models. Discussion Paper TI 94–130, Tinbergen Institute, Amsterdam.
- VAN GIERBERGEN, N. and KIVIET, J. (1996). Bootstrapping a stable AD model: Weak vs. strong exogeneity. *Oxford Bulletin of Economics and Statistics*, 48:631–656.
- ZIVOT, E. (2000). The power of single equation tests for cointegration when the cointegration vector is prespecified. *Econometric Theory*, 16:407–439.