THE EFFECT OF STRUCTURAL BREAKS ON THE ENGLE-GRANGER TEST FOR COINTEGRATION *

Antonio E. Noriega
Banco de México and Universidad de Guanajuato

Daniel Ventosa-Santaulària
Universidad de Guanajuato

Resumen: Este trabajo extiende los resultados de Gonzalo y Lee (1998) mediante el estudio del comportamiento asintótico y en muestras finitas de la prueba Engle-Granger de cointegración, cuando la función de tendencia está mal especificada y omite cambios estructurales. Consideramos quiebres de nivel o de tendencia en las variables dependiente y explicativa. Se estudian también las interacciones entre estos procesos y procesos I(1) sin quiebres. Bajo circunstancias específicas los quiebres sesgan hacia el rechazo de una relación cointegrada verdadera y el no rechazo de una relación cointegrada inexistente. Se presenta una ilustración empírica de los resultados.

Abstract: This paper extends Gonzalo and Lee's (1998) results by studying the asymptotic and finite sample behavior of the Engle-Granger test for cointegration, under misspecification of the trend function in the form of neglected structural breaks. We allow breaks in level and slope of trend in both dependent and explanatory variables. We also allow these processes to interact with I(1) processes without breaks. In some cases, breaks bias the EG test towards both rejecting a true cointegration relation, and not rejecting a non-existent one. Using real data, we present an empirical illustration of the theoretical results.

Clasificación JEL/JEL Classification: C12 ,C13, C22

Palabras clave/keywords: cointegración, quiebres estructurales, procesos integrados, prueba Engle-Granger, cointegration, structural breaks, integrated processes, Engle-Granger test

Fecha de recepción: 13 XI 2010 *Fecha de aceptación: 14 IV 2012*

* The opinions in this paper correspond to the authors and do not necessarily reflect the point of view of Banco de México. anoriega@banxico.org.mx, daniel@ventosa-santaularia.com

1. Introduction

Since the seminal contribution of Engle and Granger (1987), many economic theories involving long-run relationships have been analyzed through the concept and techniques of cointegration. These theories include money demand relations, consumption functions, the unbiased forward-market hypothesis, and purchasing power parity (see for instance Maddala and Kim, 1998, and Enders, 2004). The basic idea is that even though economic time series may wander in a nonstationary way, it is possible that a linear combination of them could be stationary. If this is the case, then there should be some long-run equilibrium relation tying the individual variables together: this is what Engle and Granger (1987) call cointegration. Assume that we are interested in testing whether two time series, x_t and y_t, are cointegrated. A preliminary requirement for cointegration is that each series is individually $I(1)$ nonstationary, that is, each has a unit root. If that is the case, cointegration among them implies that a linear combination will be stationary, that is $I(0)$. The Engle-Granger (EG) test proceeds in two steps. The first step involves the following static OLS regression

$$y_t = \hat{\alpha} + \hat{\delta} x_t + \hat{u}_t \tag{1}$$

which captures any potential long-run relationship among the variables. In the second step the residuals, $\hat{u}_t$, are used in the following Dickey-Fuller (DF) regression:

$$\Delta \hat{u}_t = \hat{\gamma} \hat{u}_{t-1} + \hat{\varepsilon}_t \tag{2}$$

If we cannot reject the hypothesis $\gamma = 0$ then there will be a unit root in the residuals, and therefore, the series x_t and y_t will not be cointegrated. On the other hand, when the t-statistic for testing the hypothesis $\gamma = 0$ $(t_{\hat{\gamma}})$ is smaller than the corresponding critical value, the residuals will be stationary, thus indicating cointegration between y_t and x_t.[1] As argued above, the EG residual-based DF t-test

[1] Critical values for this test can be found in Phillips and Oularis (1990) and MacKinnon (1991).

for cointegration assumes that both variables have a unit root, i.e., they are each $I(1)$. Gonzalo and Lee (1998) study the robustness of this test when the variables deviate from pure $I(1)$ processes. In particular, they find that the test is robust (i.e. suffers almost no size problems) to the following miss-specifications: (*a*) AR roots larger than unity; (*b*) stochastic *AR* roots; (*c*) $I(2)$ processes, and (*d*) $I(1)$ processes with deterministic linear trends.

This paper extends Gonzalo and Lee's (1998) results by studying the robustness of the EG cointegration test to structural breaks in the deterministic trend function of the processes generating y_t and x_t in model (1). Starting with Perron (1989), there is vast evidence of breaks in the long-run deterministic component of macroeconomic time series, including output, exchange rates, inflation and interest rates. Such breaks can bias the result of a unit root test towards not rejecting, when the data was generated according to a stationary process around a broken (but otherwise linear) trend. In general, miss-specification of the trend function biases unit root tests towards non-rejection. See also Perron (2003).

Since establishing the order of integration of an economic time series is still an open question, the effects of potential miss-specification on second-round tests, such as the EG test, becomes an important research topic.[2] The objective of this paper is thus to document the effects of breaks in the trend function of the variables on the Engle-Granger test for cointegration. In order to achieve this objective, we study, using asymptotic theory and simulation experiments, the behavior of the EG test for cointegration (*CI*) when the underlying processes seem to be $I(1)$ nonstationary, when in fact the type of the nonstationarity in the data is due to the presence of breaks, which induce permanent shocks in the variables, but of a deterministic nature, instead of a stochastic one. We believe this is important since the researcher may conclude in favor of unit roots in the data due to neglected structural breaks in the trend function (as shown by Perron, 1989, and others) and thus continue with *CI* tests. In particular, one of the questions we seek to answer is the following: What is the effect of the variables being not really $I(1)$ but $I(0)$ with breaks on inference based on the EG test? Even though the EG test was not intended to detect *CI* among $I(0)$ variables with breaks, but to detect *CI* among "pure" $I(1)$ variables, we believe this is a relevant question, since it could be easy in practice to confound a broken trend with a unit root.

[2] For instance, Elliot (1998) finds over-rejections of the null hypothesis of no cointegration due to a root close to but less than one in the autoregressive representation of individual variables.

Furthermore, the paper goes beyond that by studying the behavior of the EG test under combinations of $I(1)$ processes and $I(0)$ processes with breaks, which have not been analyzed before in the literature.

Leybourne and Newbold (2003) present simulated evidence of the effect on the EG test when the variables follow random walk processes with breaks in the trend function. Their findings indicate that the EG test suffers severe size distortions whenever there is an early break in y_t. Our asymptotic analysis provides a theoretical explanation of their results, and is the first to present the relevant asymptotic theory on this subject. There is some research which documents evidence on the failure of classical cointegration tests in the presence of structural breaks, but this evidence remains circumscribed to Monte Carlo experiments. See for instance Campos, Ericsson and Hendry (1996) and Gregory, Nason and Watt (1996). Related references include Gregory and Hansen (1996a, 1996b), Inoue (1999) and Arai and Kurozumi (2007). Also using simulations, Kellard (2006) finds that the Engle-Granger test tends to find substantial spurious cointegration when assessing market efficiency.

The relevance of our analysis stems from the well-known difficulty in distinguishing pure $I(1)$ processes from stationary linear trend models with structural breaks. Based on asymptotic and finite sample results, we show that, in some cases of empirical relevance, breaks have the effect of biasing the EG test towards both rejecting a true cointegration relation, and not rejecting a non-existent one. The ability of the EG test to correctly identify a cointegrated relationship depends on the break affecting the dependent variable, or the regressor, and on the position of the break.

Contrary to the robustness results found in Gonzalo and Lee (1998), our findings point to a warning on the use of this test under possible trend breaks. This arises because equation (1) is miss-specified in the presence of breaks. The obvious solution would be to correctly specify equation (1), by adding a dummy variable which captures the break, under the assumption that the change point is known.

In particular, we show that when y_t and x_t in model (1)are $I(1)$ and cointegrated, and there is a break in the level or trend of x_t, the EG test will (correctly) indicate cointegration as the sample size grows. However, when the break is in y_t, the test does not function properly. In some cases it rejects cointegration; and in general, results depend on the value of the parameters in the Data Generating Process (DGP). We also show that when y_t and x_t in model (1) are independently generated from each other, the EG test statistic will diverge, if the variables follow linear trends with breaks.[3] This occurs under $I(0)$ and $I(1)$ structures for y_t and x_t. One implication of these results is that, if divergence is towards minus infinity, the null hypothesis of no cointegration will be spuriously rejected, and this size distortion will increase with the sample size, approaching one asymptotically.

Using large sample arguments, section 2 presents results of assuming that variables are either cointegrated or independently generated from each other on the behavior of the EG -DF t-statistic under a variety of DGPs. It shows that breaks do affect the test performance, since the limit expression of the test statistic is not pivotal. Section 3 presents results from simulation experiments designed to understand the small sample behavior of the test statistic. Section 4 presents an empirical illustration of the results, by testing for cointegration between mortality and marriages (the Yule, 1926, data), variables which, on *a priori* grounds, should bear no long-run relationship. However, our empirical results suggest, according to our theory, that the test rejects the null hypothesis of no cointegration, and this seems to be the result of breaks in the variables. Last section concludes.

2. Structural breaks and the Engle-Granger test: Asymptotics

This section studies the asymptotic effects on the EG test of miss-specification due to neglected breaks in the trend function, both in the level and in the slope of the trend. The study distinguishes between a break in the dependent variable and a break in the regressor. In some cases, it turns out that the effect of a break on the EG test depends on whether it affects the dependent variable, or the regressor.

The complication with the EG test, as indeed with many other tests of cointegration, is the pre-testing problem, which arises when identifying the order of integration of the variables. This is indeed a problem, since there is ample evidence on the difficulty in differentiating between broken trend stationary models and $I(1)$ processes in economic time series. This is why we study the asymptotic behavior of the EG test under four different DGPs, widely used in applied work in economics, including $I(1)$ processes, and broken-trend stationary

[3] A similar result is reported in Noriega and Ventosa-Santaulària (2007) in the context of a spurious regression with independent processes, i.e., the t-statistic of the slope parameter in model (1) diverges when variables follow linear trends with breaks.

processes. Table 1 summarizes the DGPs considered below for both the dependent and the explanatory variables in model (1).

Table 1
DGPs for $z_t = y_t, x_t$

	DGP	*Model*
1	MS + breaks	$z_t = \mu_z + \theta_z DU_{zt} + u_{zt}$
2	TS + breaks	$z_t = \mu_z + \theta_z DU_{zt} + \beta_z t + \gamma_z DT_{zt} + u_{zt}$
3	$I(1)$ + drift	$z_t = \mu_z + \beta_z t + S_{zt}$
4	$I(1)$ + breaks	$z_t = \mu_z + \beta_z t + \gamma_z DT_{zt} + S_{zt}$

MS and TS stand for *Mean Stationary* and *Trend Stationary*, respectively, $DT_{zt} = \sum_{(i=1)}^{t} DU_{zi}$ and DU_{zt} are dummy variables allowing changes in the trend's slope and level respectively, that is, $DT_{zt} = (t - T_{b_z})1(t > T_{b_z})$ and $DU_{zt} = 1(t > T_{b_z})$, where $1(\cdot)$ is the indicator function, and T_{b_z} is the unknown date of the break in z, and we define the fraction $\lambda_z = T_{b_z}/T \in (0,1)$. Finally, $S_{zt} = \sum_{(i=1)}^{t} u_{zi}$ is a partial sum process that obeys the following assumptions (see Phillips 1986: 313):

(a) $E(S_{zt}) = 0$ for all t;

(b) $Sup_t\ E|S_{zt}|^{\beta+\varepsilon} < \infty$ for some $\beta > 2$ and $\varepsilon > 0$;

(c) $\sum =\ lim_{T\to\infty} T^{-1} E(S_{zT} S'_{zT})$ exists and is positive definite;

(d) $\{S_{zT}\}_1^{\infty}$ is strong mixing with mixing numbers α_m satisfying $\sum_1^{\infty} \alpha_m^{1-2/\beta} < \infty$.

These assumptions are quite weak, and allow S_{zT} to be general integrated processes, which include ARIMA (p, 1, q) models under very general conditions on the underlying errors (see Phillips, 1986, for further details).

The DGPs include both deterministic as well as stochastic trending mechanisms, with 16 possible combinations of them among the dependent and the explanatory variables. These combinations have practical importance, given the empirical relevance of structural breaks in the time series properties of many macro variables. DGP 1 is used to model (broken-) mean stationary variables, such as real exchange rates, unemployment rates, inflation rates, great ratios (i.e. output-capital ratio), and the current account. DGPs 2-4 are widely used to model growing variables, real and nominal, such as output, consumption, money, and prices.

2.1. *The case of independent variables*

We start by studying the behavior of the EG test under the assumption that the variables are not *CI*, i.e. they are independent of each other. The case of cointegrating variables is analyzed in the next subsection.[4]

Theorem 1 below shows that, when the DGP of at least one variable includes structural breaks, the EG test does not possess a limiting distribution, but diverges with probability approaching one asymptotically; on the other hand, in the absence of breaks, the test does not diverge.

THEOREM 1. *Let y_t and x_t be independently generated from each other under all possible combinations of DGPs in table 1. If the estimated residuals from equation (1) are used in regression model (2), then, as the sample size $T \to \infty$, the order in probability of $t_{\hat{\gamma}}$ in model (2) depends on the combination of DGPs for y_t and x_t in table 1, as follows*:

a) $t_{\hat{\gamma}} = O_p(T^{1/2})$ for all combinations except for the case when both y_t and x_t are generated by DGP 3;

b) $t_{\hat{\gamma}} = O_p(1)$ for combination 3-3.

The proof is in the appendix.

[4] All the analytic (asymptotic) results in the paper have been verified via simulations (using Matlab).

Remarks:

1. Part *a*) shows that the t-statistic diverges at rate $\sqrt{T}$, whether there are structural breaks in both variables, or just in one of them.[5] It is important to note that, as we discuss below, even under a suitable ($T^{1/2}$) normalization, which would lead to an appropriate limiting distribution, the t-statistic's distribution would depend on several nuisance parameters, and therefore, it would not be practical to simulate it in order to derive critical values. We discuss in section 2.3. the non-pivotal nature of the asymptotic distribution, and how it can be very sensitive to different parameter values.

2. Given that the Engle-Granger DF-based test for cointegration is a left tail test, result *a*) in theorem 1 is not enough to establish the presence of spurious cointegration; for this the t-statistic has to diverge to minus infinity, since divergence in the opposite direction would imply nonrejection asymptotically. Divergence towards minus infinity would imply that the size of the test approaches one asymptotically. As in the case of cointegrated variables, discussed below, the limiting expression of $t_{\hat{\gamma}}$ depends on a number of unknown parameters in the DGP (trends, location and size of breaks, among others), which makes it difficult to establish the direction of divergence. We present in section 3 some simulation experiments on the behavior of the test under different parameter values.

3. Part *b*) shows that when both y_t and x_t are independently generated according to DGP 3 in table 1, the t-statistic does not diverge.

2.2. *The case of cointegrated variables*

We will first analyze the case of two $I(1)$ cointegrated variables x_t and y_t, where there is a break in the (otherwise) linear trend of x_t. We also study by simulation the case of multiple breaks, and find that results remain qualitatively the same as the ones we present in this section. We therefore concentrate on the single break case, which allows us to economize on notation and space. We use the following DGP for x_t and y_t:

[5] An extreme miss-specification would occur if the researcher erroneously considers both dependent and explanatory variables to be $I(1)$, when in fact they are stationary around a linear trend without breaks. In this unlikely event, it can be shown that the t-statistic will also diverge, that is $t_{\hat{\gamma}}=O_p(T^{1/2})$.

$$x_t = x_0 + \mu_x t + \theta_x DT_{xt} + S_{xt} \tag{3}$$

$$y_t = \mu_y + \beta_y x_t + u_{yt} \tag{4}$$

In this case, deviations of x_t from a linear trend with a break in slope are nonstationary, i.e. they follow a unit root, which is, in turn, also part of y_t. Since u_{yt} is stationary, x_t and y_t are cointegrated, *CI*(1,1). In order to study the asymptotic behavior of the EG test for cointegration under this DGP, the following OLS regressions are considered,

$$y_t = \alpha_1 + \beta_1 x_t + u_{1t} \tag{5}$$

$$x_t = \alpha_2 + \beta_2 y_t + u_{2t} \tag{6}$$

and residuals $\hat{u}_{i,t}, i = 1, 2$ are used in the following OLS regression,

$$\Delta\hat{u}_{it} = \gamma_i \hat{u}_{it-1} + \varepsilon_{it} \tag{7}$$

As shown in theorem 2, the EG test will diverge to minus infinity, thus rejecting the null hypothesis $H_0 : \gamma = 0$, and (correctly) indicating the presence of cointegration. In other words, the trend break has no effect on the large sample performance of the test.

THEOREM 2. *Let x_t and y_t be generated according to DGPs (3) and (4), respectively. If the estimated residuals from equations (5) and (6) are used in regression model (7), then, as the sample size $T \to \infty$, for $i = 1, 2$:*

a) $\hat{\gamma}_i \xrightarrow{p} -1$

b) $T^{-1/2} t_{\hat{\gamma}_i} \xrightarrow{p} -1$

The proof is in the appendix.

As shown in part *b*), the EG-DF t-statistic will diverge to minus infinity, thus (correctly) rejecting the null hypothesis of no cointegration. The same results as those of theorem 2 are obtained when the break in x_t is in level, instead of in slope (details available upon request). When the break is in y_t instead of in x_t, however, results become less clear cut. Consider the following DGP:

$$x_t = x_0 + \mu_x t + S_{xt} \tag{8}$$

$$y_t = \mu_y + \beta_y x_t + \gamma_y DT_{yt} + u_{yt} \tag{9}$$

In this case, deviations of x_t from a linear trend are nonstationary, i.e. they follow a unit root, which is, again, also part of y_t. Since u_{yt} is stationary, x_t and y_t are cointegrated, *CI*(1,1).[6] As shown in theorem 3, the EG test statistic will diverge.

THEOREM 3. *Let x_t and y_t be generated according to DGP (8) and (9), respectively with $\lambda_y = (T_{b_y}/T) \in (0,1)$*

1. *If the estimated residuals from (5) are used in regression model (7), then, as $T \to \infty$,*

a) $\hat{\gamma}_1 = O_p(T^{-1})$

b) $t_{\hat{\gamma}_1} = O_p(T^{1/2})$

2. *If the estimated residuals from (6) are used in regression model (7) then, as $T \to \infty$,*

a) $\hat{\gamma}_2 = O_p(T^{-1})$

b) $t_{\hat{\gamma}_2} = O_p(T^{1/2})$

The proof is in the appendix.

---

[6] Note that, when x_t and y_t are generated by DGPs (8) and (9), respectively, both are cointegrated since they share a common stochastic trend, (S_{xt}). Also note that the difference $y_t - \mu_y - \beta_y x_t = \gamma_y DT_{yt} + u_{yt}$, although it does not contain a unit root, is still non-stationary due to the presence of a broken trend (and thus can be called broken-trend stationary).

Remarks:

1. Parts *1.a*) and *2.a*) show that the estimated coefficient converges in probability to a (non-zero) constant as the sample grows large (at rate T).

2. Parts *1.b*) and *2.b*) show that the t-statistic will asymptotically diverge (at rate $\sqrt{T}$). Should the t-statistic diverge to minus infinity, the null hypotheses $H_0 : \gamma_i = 0, i = 1, 2$ in regression model (7) will be correctly rejected in large samples (see below).

As a corollary of theorem 3, consider the classical formula of the t-ratio, $t_{\hat{\gamma}_1} = \hat{\gamma}_1 / \sqrt{\sigma^2_{\gamma_1}}$, where $\sigma^2_{\gamma_1}$ is the estimated variance of the parameter. Note that the sign of the t-statistic is the same as that of the numerator (the estimated parameter), since the denominator is always positive. As shown at the end of the proof of theorem 3 in the appendix, the asymptotic value of the estimated parameter $\hat{\gamma}_1$ is:

$$T_{\hat{\gamma}_1} \xrightarrow{p} 3 \frac{\frac{1}{2} - \lambda_y}{\lambda_y(\lambda_y - 1)}$$

As can be seen, the denominator of this limit expression is always negative, since $0 < \lambda_y < 1$, if $\lambda_y > 1/2(\lambda_y < 1/2)$, the numerator and hence the t-ratio, are positive (negative). Hence, if we define $t^a_{\hat{\gamma}_1}$ as the value of $T^{-1/2}\ t_{\hat{\gamma}_1}$ when $T \to \infty$, then $t^a_{\hat{\gamma}_1} < 0$ if $\lambda_y < 1/2$ and $t^a_{\hat{\gamma}_1} > 0$ if $\lambda_y > 1/2$. This implies that, asymptotically, the test correctly identifies cointegration when the break is early in the sample. A similar result is obtained in Montañes and Reyes (1998) in the context of unit root testing under breaking trend functions. A break in the second half of the sample, however, will induce the test to (erroneously) indicate no cointegration asymptotically.

The expression for the asymptotic t-statistic $t_{\hat{\gamma}2}$ is much more complicated and, in general, depends on other parameters in the DGP. Thus, the sign of the limit cannot be established analytically. Simulation evidence on the asymptotic and small sample behavior of this t-statistic is presented below.

Finally, as one might expect, if in the cointegrating equations (5) or (6) we include a trend break dummy variable (assuming of course we know the timing of the break), then the Engle-Granger t-statistic for testing cointegration will diverge to minus infinity, thereby correctly rejecting the null of no cointegration.[7]

2.3. *Numerical calculations*

The above asymptotic results show that, in some cases, the test will (1) fail to indicate cointegration, when in fact the variables are cointegrated, and (2) indicate spuriously the presence of cointegration for independent time series. Let us resort to numerical calculations, based on the asymptotic results, to study first the possibility of the test failing to indicate the presence of cointegration when in fact there is cointegration among the variables. We consider the case of $CI(1,1)$ variables with a trend-break in y_t, and concentrate on the (asymptotic) sign of the t-statistic $t_{\hat{\gamma}_2}$ from theorem 3, when regression equation (6) is estimated.[8] Assume, then, that the parameters in the DGP (8) and (9) are as follows:

$$x_t = 0.4t + S_{xt}$$

$$y_t = 0.7 + 0.2\ x_t + 0.1\ DT_{yt} + u_{yt}$$

with $\lambda_y = 0.7$. With these parameter values, and using the asymptotic expression for the t-statistic (not shown), it is easy to calculate that, as $T \to \infty$, $T^{-1/2}\ t_{\hat{\gamma}_2} \overset{a}{=} 0.0062$. On the other hand, when $\lambda_y = 0.3$, then $T^{-1/2}\ t_{\hat{\gamma}_2} \overset{a}{=} -0.0141$. Furthermore, letting $\lambda_y = 0.3$, with $\mu_x = -1.1$, $\beta_y = 0.8$, and $\gamma_y = 0.9$, we obtain that $T^{-1/2}\ t_{\hat{\gamma}_2} \overset{a}{=} 0.0835$. Therefore, the t-statistic can be asymptotically negative or positive, depending on parameter values. In other words, the test will not be pivotal, leading to potentially incorrect inference in the limit.

[7] In this case, the practitioner would estimate $y_t = \alpha_3 + \beta_3 x_t + \gamma_3 DT_{yt} + u_{3t}$. The cointegration equation is correctly specified, which implies that $\hat{\alpha}_3 \xrightarrow{p} \mu_y$, $\hat{\beta}_3 \xrightarrow{p} \beta_y$ and $\hat{\gamma}_3 \xrightarrow{p} \gamma_y$, in DGP (9). This in turn means that u_{3t} is asymptotically equivalent to u_{yt}. The *Mathematica* code that proves this is available at: http://dl.dropbox.com/u/1307356/EstEc_SpCo/MathCodeEstEc.zip.

[8] We do not study the behavior of $t_{\hat{\gamma}_1}$ as it is shown following theorem 3 that its asymptotic sign depends only on the location of the break.

Next consider the asymptotic behavior of the test for the case of independent (not cointegrated) time series. As argued above, the limiting expression of the test statistic depends on a number of unknown parameters in the DGP. In order to evaluate numerically the direction of divergence, several exercises were carried out for a range of parameter values and combinations of DGPs, using the asymptotic expression of the t-statistic (normalized by $T^{1/2}$, and a sample of size $T = 400\ 000$. The normalized t-statistic can assume both positive and negative values, depending on parameter values and combinations of DGPs. For example, assume that both variable have been independently generated by DGP 4 using the following parameters: $\sigma_x = \sigma_y = 1.00, \mu_x = 2.00, \mu_y = -7.00, \theta_x = 1.00, \theta_y = -0.50$, $\lambda_x = 0.30$, and $\lambda_y = 0.70$. In this case, $T^{-1/2} t_{\hat{\gamma}_2} \overset{a}{=} -0.0766$. When the parameter values are the same, except for $\lambda_x = 0.70$, and $\lambda_y = 0.30$, then $T^{-1/2} t_{\hat{\gamma}_2} \overset{a}{=} 0.1667$. But, if we also change to $\mu_x = 4.00$, $\mu_y = 3.00$, $\theta_x = 0.75$, $\theta_y = 1.50$, $\lambda_x = 0.70$, and $\lambda_y = 0.30$, then $T^{-1/2} t_{\hat{\gamma}_2} \overset{a}{=} 0.1440$. This simple example shows that the break locations are not the only parameter values that define the sign of the t-statistic.

3. Small sample results

To learn about the behavior of the t-statistic in finite samples, we present the results of a Monte Carlo experiment, whose design allows a number of combinations of parameter values.[9] Results show that, depending on such combinations, divergence occurs in either direction. We begin with the case of independent (not cointegrated) variables. We simulate four combinations of the DGPs introduced in table 1, and generate graphs of the EG test t-statistic, which reveal its behavior under different parameter values and sample sizes. Table 2 presents the combinations used for the simulations. For instance, combination 3-2 involves regressing a unit root process with drift against a TS model with a break in level and slope of trend. The values of the parameters were inspired by real data from Perron and Zhu (2005), comprising historical real *per capita* GDP series for industrialized economies. Figure 1 shows the behavior of the t-statistic for testing the null hypothesis of no cointegration for each of the four combinations of DGP and a sample of size $T = 100$ (Figure 2 depicts

[9] All calculations were carried out in *Matlab R.12*. Codes available from the authors upon request.

results for $T = 500$). Graphs *a, b, c*, and *d* correspond to combinations 3-2 (an $I(1)$ process against a TS with break process), 3-1 (an $I(1)$ process against an MS with break process), 2-2 (a TS with break process against a TS with break process), and 2-4 (a TS with break process against an $I(1)$ with break process), respectively. The graphs show that the t-statistic takes only negative values.[10] This implies that the possibility of divergence towards minus infinity cannot be ruled out. Accordingly, the possibility of spurious cointegration among independent series with breaks is prevalent in finite samples, and seems to grow with the sample size.

Table 2

DGPs	*Panel a* 3 $y\sim I(1)$ 2$x\sim TS$+break	*Panel b* 3 $y\sim I(1)$ 1 $x\sim MS$+break	*Panel c* 2 $y\sim TS$+break 2 $x\sim TS$+break	*Panel d* 2 $y\sim TS$+break 4 $x\sim I(1)$+break
μ_y	0.7	0.7	0.7	0.7
μ_x	0.4	0.4	0.4	0.4
β_y	0.04	0.04	0.04	0.04
β_x	[**-0.01,0.1**]		[**-0.01,0.1**]	0.07
θ_y			0.04	0.04
θ_x	0.07	[**-0.1,0.1**]	0.07	0.07
γ_y			[**-0.05,0.05**]	0.07
γ_x	[**-0.05,0.05**]		0.02	0.02
λ_y		0.3	0.3	[**0,1**]
λ_x	0.7	[**0,1**]	0.7	[**0,1**]
ρ_y	0	0.7	0	0
ρ_x	0.7	0.7	0	0

From graphs *c* and *d* in figure 1 it is interesting to note that the value of the t-statistic is uniformly below the critical value at the 1% level (-4) for any value of the location of breaks, contrary to the findings of Leybourne and Newbold (2003), who report high rejection rates for the EG test only when there is an early break in the dependent variable (i.e. $\lambda_y < 0.3$). The source of the difference

[10] Note, however, that the t-statistic is not always smaller than the critical values at, say, the 1% level, that is, -4.008 (T=100) and -3.92 (T=500).

is that the DGP they used is not exactly the same as the one used here (but see below). Using parameter values from combination 2-4 in table 2 (with $\lambda_x = 0.3$ and $\lambda_y = 0.7$) we computed rejection rates based on simulated data for various sample sizes and combination of DGPs. Rejection rates of the t-statistic for testing $\gamma = 0$ in equation (2) were computed using the critical values reported in Enders (2004) at the 1% level. The number of replications is 10 000. Results are presented in table 3.

Figure 1

Graphs of $t_{\hat{\gamma}}$. Parameter values from table 2: graph (a)-panel a; graph (b)-panel b; graph (c)-panel c; graph (d)-panel d; T=100

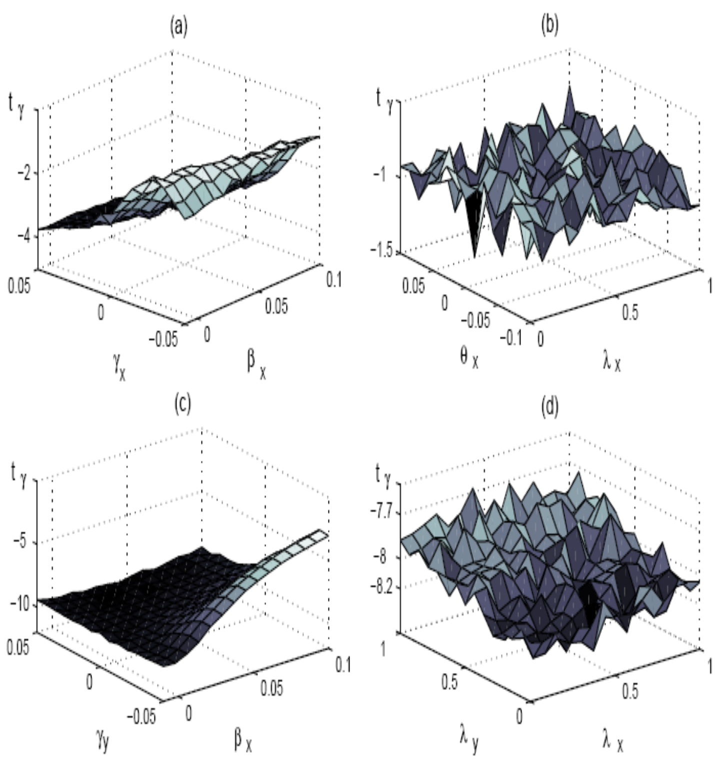
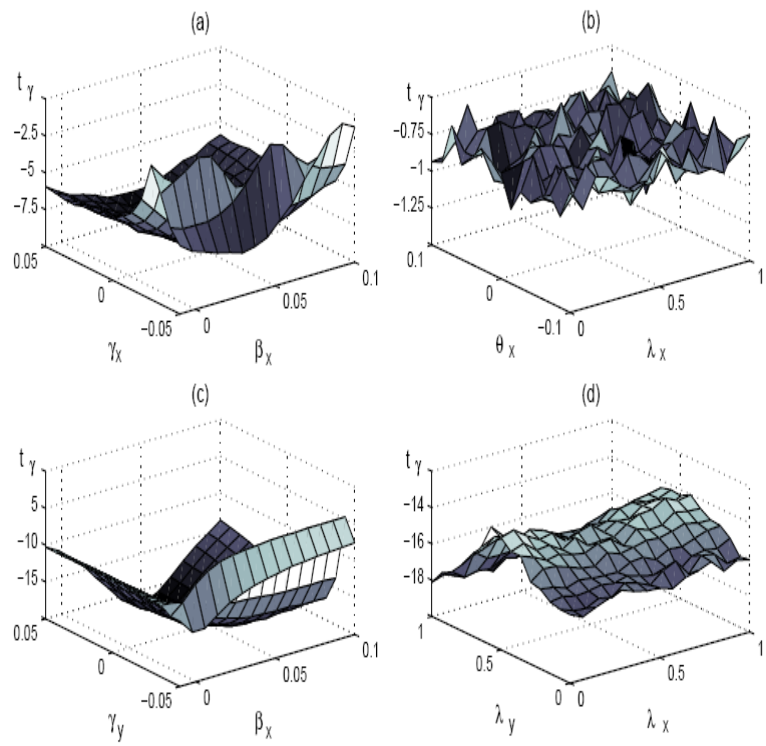


Figure 2

Graphs of $t_{\hat{\gamma}}$. Parameter values from table 2: graph (a)-panel a; graph (b)-panel b; graph (c)-panel c; graph (d)-panel d; T=500



From table 3, it is clear that the test indicates spurious cointegration in finite samples.

A second set of experiments was performed using larger values for the various parameters, see table 4.

Results are shown in figures 3 and 4, corresponding to samples of size 100 and 500. As can be seen, the t-statistic takes both positive and negative values, but in this case graphs (a), (c) and (d) indicate that as the sample size grows, the statistic tends to move towards positive values. For this second set of experiments, our results resemble those of Leybourne and Newbold (2003), in the sense that the t-statistic tends to be more negative the closer are the breaks to the beginning of the sample, according to graph (d) in figures 3 and 4.

Table 3

Rejection rates for $t_{\widehat{\gamma}}$

Combinations of DGPs in the assumption

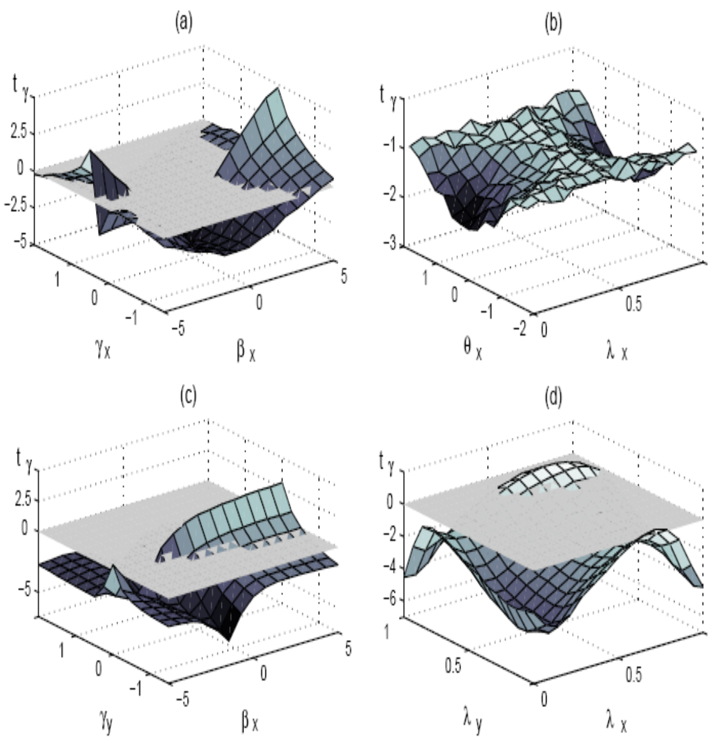
T	1-1	1-2	1-3	1-4	2-2	2-3	2-4	4-2
50	.99	.99	.99	.99	.99	.99	.99	.55
100	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
200	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
500	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table 4

DGPs	*Panel a* 3 $y{\sim}I(1)$ 2 $x{\sim}TS$+break	*Panel b* 3 $y{\sim}I(1)$ 1 $x{\sim}MS$+break	*Panel c* 2 $y{\sim}TS$+break 2 $x{\sim}TS$+break	*Panel d* 2 $y{\sim}TS$+break 4 $x{\sim}I(1)$+break
μ_y	2.7	2.7	2.7	2.7
μ_x	0.5	0.5	0.5	0.5
β_y	0.04	0.04	0.9	0.04
β_x	[**-5,5**]		[**-5,5**]	0.9
θ_y			-1.7	-1.7
θ_x	1.5	[**-1.8,1.8**]	1.5	
γ_y			[**-1.5,2**]	0.07
γ_x	[**-1.5,2**]		1.5	1.5
λ_y		0.3	0.3	[**0,1**]
λ_x	0.7	[**0,1**]	0.7	[**0,1**]
ρ_y	0	0.7	0	0
ρ_x	0.7	0.7	0	0

Figure 3

Graphs of $t_{\hat{\gamma}}$. Parameter values from table 4: graph (a)-panel a; graph (b)-panel b; graph (c)-panel c; graph (d)-panel d; T=100



Turning to the case of cointegrated variables, consider DGPs (3) and (4) with $\mu_x = 0.4$, $\mu_y = 0.7$, $\beta_y = 0.09$, $\theta_x \in [-0.1, 0.1]$, $\lambda_x \in (0, 1)$, and u_{yt}, u_{xt} are *iid* random variables. Figure 5 shows graphs of the EG test t-statistic, based on regression (5), for samples of size $T = 100$ (graph (a)), and $T = 500$ (graph (b)), using 1 000 replications. Graph (c) shows the behavior of the normalized statistic. Note from graphs (a) and (b) how the statistic becomes more negative as the sample size grows. Furthermore, it does not depend on the size nor the location of the break. Graph (c) shows that convergence of the simulated value to the asymptotic one is fast. These results allow us to argue that the large sample results seem to hold in finite samples: the test correctly rejects the unit root in the residuals of (5), thus indicating cointegration.

Figure 4

Graphs of $t_{\hat{\gamma}}$. Parameter values from table 4: graph (a)-panel a; graph (b)-panel b; graph (c)-panel c; graph (d)-panel d; T=500

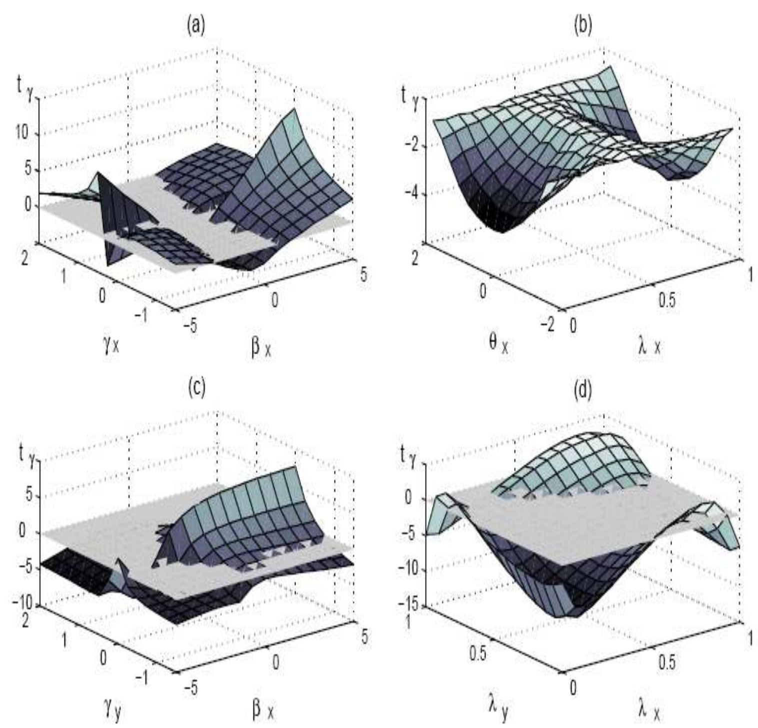


Figure 5
Graphs of $t_{\hat{\gamma}}$. DGPs (3) [I(1)] and (4) [I(1) + breaks]

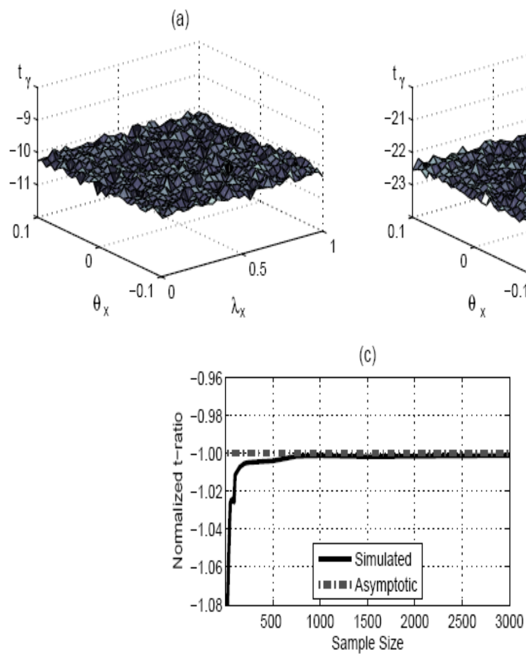
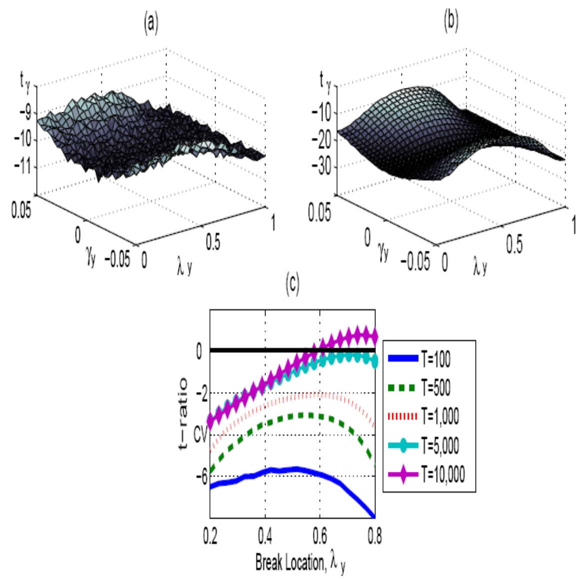


Figure 6 shows similar graphs but uses DGP (8) and (9) with $\gamma_y \in [-0.05, 0.05]$, and $\gamma_y \in (0, 1)$, i.e. two $CI(1{,}1)$ variables with a trend-break in y (the rest of the parameter values are those of the previous experiment). Graph (a) shows that, with $T = 100$, the statistic does not seem to be very sensitive to values of the break fraction. Graph (b) shows, however, that as sample size increases, the dependence of the statistic on the break fraction starts to show up, as the asymptotic results of theorem 3 indicate. Graph (c) shows the behavior of the t-statistic for different values of the break fraction and different sample sizes, together with the critical value at the 10% level, using $T = 100$.[11]

[11] We report a single critical value (corresponding to T=100) for simplicity,

Figure 6
Graphs of $t_{\hat{\gamma}}$. DGPs (8) [I(1)] and (9) [I(1) + breaks]



As can be seen, for samples of size $T = 100$, the test rejects the null hypothesis, thus correctly indicating the presence of cointegration; the break fraction does not seem to be of any relevance. As the sample size grows, however, the sign of the statistic becomes a function of the break fraction. For the case $T = 500$, for instance, the break fraction affects inference. For very large sample sizes ($T = 10\ 000$), the behavior of the statistic approaches the one predicted by the asymptotic theory presented in theorem 3. For sample sizes of practical relevance in macroeconomics ($100 \leq T \leq 500$), inference based on the t-statistic could be influenced by the break fraction, which would lead to incorrect inferences when the break is around the middle of the sample. For larger samples, the null is not rejected, incorrectly rejecting cointegration, for any value of the break fraction. As expected from results in theorem 3, the statistic takes

without affecting our conclusions. We use γ_y=0.14. Increasing the value of the size of the break makes the finite sample behavior closer to the asymptotic one.

positive values when the break occurs in the second half of the sample, and *vice versa*.

4. Empirical evidence

To illustrate the possibility of spurious cointegration using real data, we present an empirical exercise using data on *marriages* (to be precise, the proportion of Church of England marriages to all marriages) and *mortality* rates per 1 000 persons, studied by Yule (1926). The data are annual and span the period 1886 to 1911. Note that this section does not pretend to offer a complete time series analysis of these data. Its purpose is simply to present an illustration on the possibility of finding a cointegration relationship using real data, comprising variables which in principle have no relationship with each other.

We start by applying Augmented Dickey-Fuller tests under a variety of lag length selection criteria. In particular, we use the general-to-specific procedure advocated in Perron (1997), as well as several information criteria (AIC,BIC,HQ). We also apply Ng-Perron tests using the improving power correction suggested by Perron and Qu (2007). Individual results are available upon request.

Pretesting for unit roots (see table 5, column labeled O.I.) indicate that each variable follows a unit root process.[12] Since the theoretical results presented above apply when at least one of the variables has undergone a structural break, we followed Zivot and Andrews (1992, ZA henceforth), Perron (1997), and Kapetanios (2005), and test for a unit root allowing for breaks in the trend function. As can be seen from table 5, when allowing for a broken trend in *mortality*, methods in ZA, Perron (1997), and Kapetanios (2005), detect a break in/or around 1889, while the unit root is rejected. On the other hand, for *marriages*, the unit root cannot be rejected for either test.

When structural breaks are not taken into account, both variables would appear to be $I(1)$. The Engle-Granger test statistic is -5.824 when we run *marriages* on *mortality*, and -6.263 when we run *mortality* on *marriages*. As can be seen, this example, which has been used to illustrate the spurious regression phenomenon, can also be used to exemplify the presence of spurious cointegration, due to the presence of a structural break in one of the variables.

[12] We applied the procedures in Dickey and Pantula (1987) and Pantula (1989) to determine the order of integration of a time series.

Therefore, there appears to be a long-run equilibrium relationship between *marriages* and *mortality* rates in England and Wales, for the years 1866-1911, according to the Engle-Granger test. Of course, this is a spurious finding, since one does not expect to find cointegration between these two variables.

Figure 7
Mortality rate (thin line) and Anglican marriages (thick line): (a) time series, (b) scatter plot

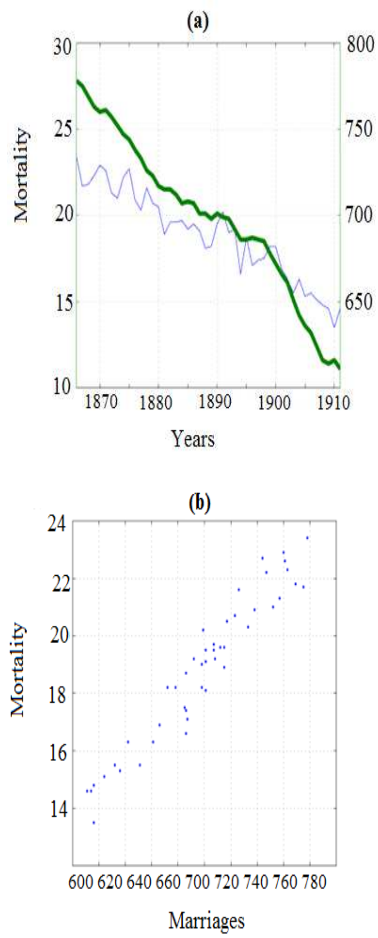


Table 5

Unit root tests and unit root tests allowing for structural breaks

Variable	T	*O.I.*[1,3]	*ZA*[2,3]	*Perron (1997)*[2,3]	*Kapetanios*[3]
Marriages	46	$I(1)^4$	$I(1)$	$I(1)$	$I(1)$
Mortality	46	$I(1)$	$I(0) + break$ 1889	$I(0) + break$ 1888	$I(0) + break$ 1889

Notes: [1]: O.I. stands for order of integration. The results of this column are based on ADF and Ng and Perron tests, as discussed in the text. [2]: The results indicating a Broken TS model, hold for any of the three models considered in ZA and Perron: with a changing mean, a changing trend, or a combination of them. For these two tests we used a trimming of 5 observations. [3]: The maximum length of the AR augmentation is set according to the formula. $k_{max}=integer[12(T/100)^{1/4}]$. See for instance Ng and Perron (2001) [4]: For this series, evidence based on the Ng and Perron tests is mixed: while the MZ_α and the MSB indicate that the process is $I(2)$, the MZ_t and the MP_T favour an $I(1)$ model.

5. Conclusions

Results in this paper represent an extension of Gonzalo and Lee's (1998) results and provide a theoretical explanation of Leybourne and Newbold's (2003) simulated evidence. We find that the EG test is sensitive to miss-specification of the trend behavior and can lead to spurious rejection of the null hypothesis of no cointegration for independent time series with breaks. It can also reject cointegration when the variables are indeed cointegrated. In particular, it has been shown that the Engle-Granger test for cointegration, based on the DF t-statistic, does not posses a limiting distribution, but diverges at rate root-T. Given the dependency of the asymptotic distribution on the various parameters in the DGP, the paper analyzed the behavior of the t-statistic through Monte Carlo simulations. Results show that the divergence of the EG test statistic can occur in either direction. For a particular set of (empirically relevant) parameters, the simulation experiments indicate that the statistic tends to minus infinity, and, therefore, can misleadingly indicate cointegration among independent variables. However, we also find that a different set of parameter values indicate divergence towards infinity, which would lead to correct inference. We also showed that a trend-break in the dependent variable of two $CI(1,1)$ variables affects the behavior of the test statistic, and could lead to incorrect inference, depending on the break fraction, the sample size, and the break size.

Through an empirical exercise, we showed that spurious cointegration can arise when there are breaks in the data generating process. We believe the results presented are relevant, given the difficulty in distinguishing among $I(1)$ variables from stationary variables around broken trends, that is, given the low power of unit root tests against broken trend stationary alternatives. All in all, the EG test should be used with caution, since the presence of neglected breaks could produce spurious rejections of the null hypothesis of no cointegration among independent time series, or over-rejections of a genuine cointegration relationship. Given the potential negative impact of neglected breaks on inference using the EG test, we adhere to Gonzalo and Lee's (1998: 149) recommendation in the sense that "...pre-testing for individual unit roots is not enough. We have to be sure that the variables do not have any other trending or long-memory behavior different from that of a unit root process".

References

Arai, Y. and E. Kurozumi. 2007. Testing for the null hypothesis of cointegration with a structural break, *Econometric Reviews*, 26(6): 705-739.

Campbell, J.Y. and P. Perron. 1991. Pitfalls and opportunities: What macroeconomists should know about unit roots, in O.J. Blanchard and S. Fischer (Eds.), *NBER Macroeconomics Annual*, MIT Press, Cambridge, pp. 141-201.

Campos, J., N.R. Ericsson and D.F. Hendry. 1996. Cointegration tests in the presence of structural breaks, *Journal of Econometrics*, 70(1): 187-220.

Dickey, D.A. and S.G. Pantula. 1987. Determining the order of differencing in autoregressive processes, *Journal of Business & Economic Statistics*, 5: 455-461.

Elliot, G. 1998. On the robustness of cointegration methods when regressors almost have unit roots, *Econometrica*, 66(1): 149-158.

Enders, W. 2004. *Applied Econometric Time Series*, Second Edition, Wiley.

Engle, R.F. and C.W.J. Granger. 1987. Cointegration and error correction: representation, estimation and testing, *Econometrica*, 55: 251-276.

Gonzalo, J. and T.H. Lee. 1998. Pitfalls in testing for long-run relationships, *Journal of Econometrics*, 86(1): 129-154.

Gregory, A. and B. Hansen. 1996a. Residual-based tests for cointegration in models with regime shifts, *Journal of Econometrics*, 70: 99-126.

———. 1996b. Tests for cointegration in models with regime and trend shifts, *Oxford Bulletin of Economics and Statistics*, 58(3): 555-560.

Gregory, A., J. Nason and D. Watt. 1996. Testing for structural breaks in cointegrated relationships, *Journal of Econometrics*, 71(1-2): 321-341.

Inoue, A. 1999. Tests of cointegrating rank with a trend-break, *Journal of Econometrics*, 90(2): 215-237.

Kapetanios, G. 2005. Unit-root testing against the alternative hypothesis of up to m structural breaks, *Journal of Time Series Analysis*, 26(1): 123-133.

Kellard, N. 2006. On the robustness of cointegration tests when assessing market efficiency, *Finance Research Letters*, 3: 57-64.

Leybourne, S.J. and P. Newbold. 2003. Spurious rejections by cointegration tests induced by structural breaks, *Applied Economics*, 35: 1117-1121.

MacKinnon, J.G. 1991. Critical values for co-integration tests, in R.F. Engle and C.W.J. Granger (eds.), *Long-Run Economic Relationships*, Oxford University Press, pp. 267-276.

Maddala, G. and I.M. Kim. 1998. *Unit Roots, Cointegration and Structural Change*, Cambridge University Press.

Montañés, A. and M. Reyes. 1998. Effect of a shift in the trend function on Dickey-Fuller unit root tests, *Econometric Theory*, 14(3): 355-363.

Ng, S. and P. Perron. 2001. Lag length selection and the construction of unit root tests with good size and power, *Econometrica*, 69: 1519-1554.

Noriega, A. and D. Ventosa-Santaulària. 2006. Spurious regression under broken-trend stationarity, *Journal of Time Series Analysis*, 27(5): 671-684.

———. 2007. Spurious regression and trending variables, *Oxford Bulletin of Economics and Statistics*, 69(3): 439-444.

Ogaki, M. and Park, J. 1997. A cointegration approach to estimating preference parameters, *Journal of Econometrics*, 82(1): 107-134.

Pantula, S.G. 1989. Testing for unit roots in time series data, *Econometric Theory*, 5(2): 256-271.

Perron, P. 1989. The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, 57: 1361-1401.

——. 1997. Further evidence on breaking trend functions in macroeconomic variables, *Journal of Econometrics*, 80: 355-385.

——. 2003. Statistical adequacy and the testing of trend versus difference stationarity: Some comments. (unpublished).

—— and X. Zhu. 2005. Structural breaks with deterministic and stochastic trends, *Journal of Econometrics*, 129: 65-119.

Perron, P. and Z. Qu. 2007. A simple modification to improve the finite sample properties of Ng and Perron's unit root tests, *Economics Letters*, 94(1): 12-19.

Phillips, P.C.B. 1986. Understanding spurious regressions in econometrics, *Journal of Econometrics*, 33: 311-340.

—— and S. Oularis. 1990. Asymptotic properties of residual based tests for cointegration, *Econometrica*, 58: 165-193.

Yule, G.U. 1926. Why do we sometimes get nonsense-correlations between time-series? A study in sampling and the nature of time-series, *Journal of the Royal Statistical Society*, 89(1): 1-63.

Zivot, E. and D.W.K. Andrews. 1992. Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis, *Journal of Business & Economic Statistics*, 10(3): 251-270.

Appendix

Proof of THEOREM 1

Here we present the proof on how to obtain the order in probability of two combinations of DGPs, namely, the combinations 2-2 and 4-4: $z_t = \mu_z + \theta_z DU_{zt} + \beta_z t + \gamma_z DT_{zt} + u_{zt}$, and $z_t = \mu_z + \beta_z t + \gamma_z DT_{zt} + S_{zt}$, respectively, for $z = y, x$. The proof for both cases can be presented jointly because the leading terms in the asymptotic expressions for DGPs 2 and 4 (the deterministic linear trend and the trend break dummies) are the same. We omit the proof for the rest of the combinations since they follow the same steps. All sums run from $t = 1$ to T. The OLS estimator of γ from (2) is:

$$\hat{\gamma} = \left(\sum \Delta \hat{u}_t \hat{u}_{t-1}\right) \left(\sum \hat{u}_{t-1}^2\right)^{-1}$$

Where $\sum \Delta \hat{u}_t \hat{u}_{t-1} = \sum \Delta y_t y_{t-1} - \hat{\alpha} \sum \Delta y_t - \hat{\delta} \sum \Delta y_t x_{t-1} - \hat{\delta} \sum \Delta x_t$ $y_{t-1} + \hat{\alpha}\hat{\delta} \sum \Delta x_t + \hat{\delta}^2 \sum \Delta x_t x_{t-1}$.

From direct calculation, and using the fact that, for (1), $\hat{\alpha} = O_p(T), \hat{\delta} = O_p(1), \sum \hat{u}_{(t-1)}^2 = O_p(T^3)$ (see Noriega and Ventosa-Santaulària, 2006), it is simple to show that each element of $\sum \hat{u}_t$ $\hat{u}_{(t-1)}$ is $O_p(T^2)$.

Therefore,

$$\hat{\gamma} = \frac{O_p(T^2)}{O_p(T^3)}$$

which implies that $T\hat{\gamma} = O_p(1)$.

Now define the residuals $\hat{\varepsilon}_t$ from (2) as:

$$\hat{\varepsilon}_t = \hat{u}_t - \hat{\gamma} \hat{u}_{(t-1)}$$

The estimated variance is:

$$\hat{\sigma}_\varepsilon^2 = T^{-1} \left[\sum (\Delta \hat{u}_t)^2 + \hat{\gamma}^2 \sum \hat{u}_{t-1}^2 - 2\hat{\gamma} \sum \Delta \hat{u}_t \hat{u}_{t-1}\right]$$

where, again, direct calculations indicate that

$$\sum (\Delta \hat{u}_t)^2 = \sum \left(\Delta y_t - \hat{\delta} \Delta x_t\right)^2 = O_p(T)$$

Hence

$$\hat{\sigma}_\varepsilon^2 = T^{-1}\left[O_p(T) + (T\hat{\gamma})^2 T^{-2} O_p\left(T^3\right) - 2(T\hat{\gamma}) T^{-1} O_p\left(T^2\right)\right]$$
$$= O_p(1).$$

Finally, the t-statistic $t_{\hat{\gamma}}$ can be written as:

$$T\hat{\gamma}\left[\hat{\sigma}_\varepsilon^2 T^3 \left(\sum \hat{u}_{t-1}^2\right)^{-1}\right]^{1/2} = T^{-1/2} t_{\hat{\gamma}} = O_p(1)$$

which proves the theorem for DGPs 2-2 and 4-4.

Proof of THEOREM *2*

We present the proof using residuals from equation (5). The steps needed to prove the Theorem for the case of regression model (6) are the same. In the first step we obtain the limiting distribution of the OLS estimates $\hat{\alpha}_1$ and $\hat{\beta}_1$ from (5). For this we need the sample moments of y and x (All sums run from $t = 1$ to T):

$$\begin{pmatrix} \hat{\alpha}_1 \\ \hat{\beta}_1 \end{pmatrix} = \begin{pmatrix} T & \sum x_t \\ \sum x_t & \sum x_t^2 \end{pmatrix}^{-1} \begin{pmatrix} \sum y_t \\ \sum x_t \; y_t \end{pmatrix}$$

where simple calculations give

$$\sum x_t = x_0 T + \mu_x \sum t + \theta \sum DT_{xt} + \left(T^{-3/2} S_{xt}\right) T^{3/2}$$
$$= \frac{1}{2}\left[\mu_x + \theta(1-\lambda)^2\right] T^2 + O_p\left(T^{3/2}\right)$$

$$\sum x_t^2 = \frac{1}{3}\left[\mu_x^2 + \theta^2(1-\lambda)^3 + \mu_x\theta(1-\lambda)^2(\lambda+2)\right]T^3 + O_p\left(T^{5/2}\right)$$

$$\sum y_t = \frac{1}{2}\beta_y\left[\mu_x + \theta(1-\lambda)^2\right]T^2 + O_p\left(T^{3/2}\right)$$

$$\sum x_t y_t = \tfrac{1}{3}\beta_y\left[\mu_x^2 + \theta^2(1-\lambda)^3 + \mu_x\theta(1-\lambda)^2(\lambda+2)\right]T^3 + O_p\left(T^{5/2}\right)$$

Calculations are carried out using a *Mathematica* code, available upon request, which produces expressions for $\hat{\alpha}_1$ and $\hat{\beta}_1$, as functions of (decreasing powers of) the sample size. This ordering in terms of powers of the sample size allows us to identify the required normalization, used to establish the following asymptotic results:

$$\hat{\alpha}_1 \xrightarrow{p} \mu_y$$

$$\hat{\beta}_1 \xrightarrow{p} \beta_y$$

The second step consists in obtaining the asymptotic behavior of the elements of the t-statistic, $t_{\hat{\gamma}_1} = \hat{\gamma}_1\left[\hat{\sigma}^2_{\varepsilon_1}\left(\sum \hat{u}^2_{1t-1}\right)^{-1}\right]^{-1/2}$, where $\hat{\gamma}_1$ is the OLS estimator of γ_1 in (7), and $\hat{\sigma}^2_{\varepsilon_1} = T^{-1}\sum \varepsilon^2_{1t}$. We begin with $\hat{\gamma}_1$. Note that $\hat{\gamma}_1 = \left(\sum \Delta\hat{u}_{1t}\hat{u}_{1t-1}\right)\left(\sum \hat{u}^2_{1t-1}\right)^{-1}$, where $\hat{u}_{1t} = y_t - \hat{\alpha}_1 - \hat{\beta}_1 x_t$. The numerator can be written as $\sum \Delta\hat{u}_{1t}\hat{u}_{1t-1} = \sum\left[\left(\beta_y - \hat{\beta}_1\right)\Delta x_t + \Delta u_{yt}\right]\left[\left(\beta_y - \hat{\beta}_1\right)x_{t-1} + (\mu_y - \hat{\alpha}_1) + u_{yt-1}\right] = \sum \Delta u_{yt}u_{yt-1} + O_p(1)$, since $(\beta_y - \hat{\beta}_1)$ and $(\mu_y - \alpha_1)$ vanish asymptotically. Since $\sum \Delta u_{yt}u_{yt-1} = \sum u_{yt}u_{yt-1} - \sum u^2_{yt-1} = o_p(T)$, then:

$$T^{-1}\sum \Delta\hat{u}_{1t}\hat{u}_{1t-1} \xrightarrow{p} -\sigma_y^2$$

The denominator can be written as

$$\sum \hat{u}^2_{1t-1} = \sum\left[\left(\beta_y - \hat{\beta}_1\right)x_{t-1} + (\mu_y - \hat{\alpha}_1) + u_{yt-1}\right]^2$$

$$= \sum u^2_{yt-1} + o_p(1)$$

Hence

$$T^{-1}\sum \hat{u}_{1t-1}^{2} \xrightarrow{p} \sigma_{y}^{2}$$

Therefore,

$$\hat{\gamma}_1 \xrightarrow{p} -1$$

as shown in part *a*) of theorem 2.

Now define residuals $\hat{\varepsilon}_{1t} = \Delta\hat{u}_{1t} - \hat{\gamma}_1\hat{u}_{1t-1}$. The error variance, $\hat{\sigma}_{\varepsilon_1}^2$, is defined as

$$\hat{\sigma}_{\varepsilon_1}^2 = T^{-1}\left[\sum \left(\Delta\hat{u}_{1t}\right)^2 + \hat{\gamma}_1^2\sum \hat{u}_{1t-1}^2 - 2\hat{\gamma}_1\sum \Delta\hat{u}_{1t}\hat{u}_{1t-1}\right].$$

From results above, and the fact that $T^{-1}\sum\left(\Delta\hat{u}_{1t}\right)^2 \xrightarrow{p} \sigma_y^2$, it is easy to show that

$$\hat{\sigma}_{\varepsilon_1}^2 \xrightarrow{p} \sigma_y^2.$$

Therefore,

$$\hat{\gamma}_1\left[\hat{\sigma}_{\varepsilon_1}^2 T\left(\sum \hat{u}_{1t-1}^2\right)^{-1}\right]^{-1/2} = T^{-1/2}t_{\hat{\gamma}_1} \xrightarrow{p} -1$$

as shown in part *b*) of theorem 2.

Proof of THEOREM 3

The proof is very similar to that for theorem 2. As above, we first obtain the limiting distribution of the OLS estimates $\hat{\alpha}_1$ and $\hat{\beta}_1$ from (5), and $\hat{\alpha}_2$ and $\hat{\beta}_2$ from (6). Calculations are carried out using a *Mathematica* code,[13] which produces expressions for $\hat{\alpha}_1$, $\hat{\beta}_1$, $\hat{\alpha}_2$ and $\hat{\beta}_2$. From the code, we get the following asymptotic results:

$$T^{-1}\hat{\alpha}_1 \xrightarrow{p} -\gamma_y(\lambda_y - 1)^2\lambda_y$$

[13] http://dl.dropbox.com/u/1307356/EstEc_SpCo/MathCodeEstEc.zip.

$$\hat{\beta}_1 \xrightarrow{p} \beta_y + \frac{\gamma_y}{\mu_y}(\lambda_y - 1)^2 (1 + 2\lambda_y)$$

and

$$T^{-1}\hat{\alpha}_2 \xrightarrow{p} \frac{\gamma_y(\lambda_y-1)^2\lambda_y\mu_x\left(\gamma_y\lambda_y^2-\gamma_y-\beta_y\mu_x\right)}{\gamma_y^2(\lambda_y-1)^3(1+3\lambda_y)-2\beta_y\gamma_y(\lambda_y-1)^2(1+2\lambda_y)\mu_x-\beta_y^2\mu_x^2}$$

$$\hat{\beta}_2 \xrightarrow{p} \frac{\mu_x\left(3\gamma_y\lambda_y^2-\gamma_y-2\gamma_y\lambda_y^3-\beta_y\mu_x\right)}{\gamma_y^2(\lambda_y-1)^3(1+3\lambda_y)-2\beta_y\gamma_y(\lambda_y-1)^2(1+2\lambda_y)\mu_x-\beta_y^2\mu_x^2}$$

Note that, if there is no break in $y(\lambda_y = 1)$, then $\hat{\alpha}_1 \xrightarrow{p} 0$, $\hat{\alpha}_2 \xrightarrow{p} 0$, $\hat{\beta}_1 \xrightarrow{p} \beta_y$, and $\hat{\beta}_2 \xrightarrow{p} \frac{1}{\beta_y}$.

As in the proof of theorem 2, the second step consists in obtaining the asymptotic behavior of the elements of the t-statistic, $t_{\hat{\gamma}_i} = \hat{\gamma}_i\left[\hat{\sigma}^2_{\varepsilon_i}\left(\sum \hat{u}^2_{it-1}\right)^{-1}\right]^{-1/2}$ for $i = 1, 2$, where $\hat{\gamma}_i$ is the OLS estimator of γ_i in (7), and $\hat{\sigma}^2_{\varepsilon_i} = T^{-1}\sum \hat{\varepsilon}^2_{it}$. Note that, as before, $\hat{\gamma}_i = \left(\sum \Delta\hat{u}_{it}\hat{u}_{it-1}\right)\left(\sum \hat{u}^2_{it-1}\right)^{-1}$, where $\hat{u}_{1t} = y_t - \hat{\alpha}_1 - \hat{\beta}_1 x_t$ and $\hat{u}_{2t} = x_t - \hat{\alpha}_2 - \hat{\beta}_2 y_t$. Direct calculations, similar as those performed in the proof of theorem 2 show that $\sum \Delta\hat{u}_{it}\hat{u}_{it-1} = O_p\left(T^2\right)$, and $\sum \hat{u}^2_{it-1} = O_p\left(T^3\right)$, wich imply that $\hat{\gamma}_i = O_p\left(T^{-1}\right),\ i = 1, 2$. This proves parts *1.a*) and *2.a*) of the theorem.

Residuals are defined as $\hat{\varepsilon}_{it} = \Delta\hat{u}_{it} - \hat{\gamma}_i\hat{u}_{it-1}$, and the error variance, $\hat{\sigma}^2_{\varepsilon_i}$, is defined as

$$\hat{\sigma}^2_{\varepsilon_i} = T^{-1}\left[\sum\left(\Delta\hat{u}_{it}\right)^2 + \hat{\gamma}_i^2\sum \hat{u}^2_{it-1} - 2\hat{\gamma}_i\sum \Delta\hat{u}_{it}\hat{u}_{it-1}\right].$$

From direct calculations it follows that $\sum\left(\Delta\hat{u}_{it}\right)^2 = O_p\left(T\right)$ Therefore, combining this result with the orders in probability for $\sum \Delta\hat{u}_{it}$ $\hat{u}_{it-1}$ and $\sum \hat{u}^2_{it-1}$ we obtain $\hat{\sigma}^2_{\varepsilon_i} = O\left(1\right)$. Finally, since $\sum \hat{u}^2_{it-1} = O_p\left(T^3\right)$, then

$$T\hat{\gamma}_i\left[\hat{\sigma}^2_{\varepsilon_i}T^3\left(\sum \hat{u}^2_{it-1}\right)^{-1}\right]^{-1/2} = T^{-1/2}t_{\hat{\gamma}_i}$$

converges, which implies that $t_{\hat{\gamma}_i} = O_p\left(T^{1/2}\right)$ thus proving parts *1.b*) and *2.b*) of the theorem.

The asymptotic value of $\hat{\gamma}_1$

To obtain the asymptotic value of the parameter estimate, $\hat{\gamma}_1$, we will need the asymptotic expressions for $\hat{\alpha}_1$, $\hat{\beta}_1$, and $\hat{\sigma}^2$ from (5). The first two have been presented in the proof of theorem 3, and are reproduced here for convenience:

$$T^{-1}\hat{\alpha}_1 \xrightarrow{p} -\gamma_y(\lambda_y - 1)^2\lambda_y \overset{\text{def}}{=} \alpha,$$

$$\hat{\beta}_1 \xrightarrow{p} \beta_y + \frac{\gamma_y}{\mu_y}(\lambda_y - 1)^2(1 + 2\lambda_y) \overset{\text{def}}{=} \beta,$$

$$T^{-2}\hat{\sigma}^2 \xrightarrow{p} \frac{1}{3}\gamma_y^2\left(\lambda_y - \lambda_y^2\right)^3,$$

where the limit expression for $\hat{\sigma}^2$ is obtained from the same *Mathematica* code used in the proof of theorem 3. Now, the OLS formula of $\hat{\gamma}_1\gamma_1$ from (7) is:

$$\hat{\gamma}_1 = \frac{\sum \Delta\hat{u}_{1t}\hat{u}_{1t-1}}{\sum \hat{u}_{1t-1}^2}$$

Note that the denominator is asymptotically identical to $T\hat{\sigma}^2$. This implies that $\sum \hat{u}_{1t-1}^2 = O_p\left(T^3\right)$, since $T^{-2}\hat{\sigma}^2 = O_p(1)$. Let us now turn to the numerator. The sum can be decomposed in two sums:

$$\sum \Delta\hat{u}_{1t}\hat{u}_{1t-1} = \sum \Delta y_t\hat{u}_{t-1} - \hat{\beta}_1 \sum \Delta x_t\hat{u}_{t-1}.$$

Let us focus on the first element which, by simple substitution can be expressed as:

$$\sum \Delta y_t\hat{u}_{t-1} = \sum\left(\beta_y\Delta x_t + \gamma_y DU_{yt} + \Delta u_{yt}\right)\left(y_{t-1} - \hat{\alpha}_1 - \hat{\beta}_1 x_{t-1}\right).$$

Replacing Δx_t with $\mu_x + u_{xt}$, x_{t-1} with $x_{-1} + \mu_x (t-1) + S_{xt-1}$, $\hat{\alpha}_1$ with $T\alpha$, $\hat{\beta}_1$ with β, and y_{t-1} with $\mu_y + \beta_y x_{t-1} + \gamma_y DT_{yt-1} + u_{yt-1}$ we obtain (note that in order to save space, we only present the leading terms of the relevant asymptotic expression, that is, the asymptotically relevant terms):

$$\sum \Delta y_t \hat{u}_{t-1} = \sum \big(\beta_y \mu_x + \gamma_y DU_{yt}\big) \cdot \big(\mu_x \beta_y t + \gamma_y DT_{yt-1} - \hat{\alpha}_1 - \mu_x \hat{\beta}_1 t\big).$$

Simple but tedious algebra shows that

$$T^{-2} \sum \Delta y_t \hat{u}_{t-1} \xrightarrow{p} \lambda_y^2 \gamma_y^2 \left(\lambda_y^3 - \frac{5}{2}\lambda_y^2 + 2\lambda_y - \frac{1}{2} \right)$$

Note that we also replaced $\hat{\alpha}_1$ and $\hat{\beta}_1$ by their asymptotic expressions.

Using exactly the same procedure as before, the second sum $\sum \Delta x_t \hat{u}_{t-1} = O_p\left(T^k\right)$ where $k < 2$, so it can be ignored from the asymptotic calculus. In sum, we have:

$$\sum \Delta \hat{u}_t \hat{u}_{1t-1} \overset{a}{\sim} \sum \Delta y_t \hat{u}_{t-1},$$

where $\overset{a}{\sim}$ denotes $\sum \Delta \hat{u}_t \hat{u}_{1t-1} / \sum \Delta y_t \hat{u}_{t-1} \xrightarrow{p} 1$. Remember now that the denominator of $\hat{\gamma}_1$ is $O_p\left(T^3\right)$. Since the numerator is $O_p\left(T^2\right)$ we easily obtain that $\hat{\gamma}_1 = O_p\left(T^{-1}\right)$. As for the asymptotic expression, the above results allow us to write:

$$T\hat{\gamma}_1 \xrightarrow{p} \frac{\lambda_y^2 \gamma_y^2 \left(\lambda_y^3 - \frac{5}{2}\lambda_y^2 + 2\lambda_y - \frac{1}{2}\right)}{(1/3)\gamma_y^2 \big(\lambda_y - \lambda_y^2\big)^3}$$

Finally, using simple algebra this expression simplifies to:

$$T\hat{\gamma}_1 \xrightarrow{p} 3 \frac{\frac{1}{2} - \lambda_y}{\lambda_y \left(\lambda_y - 1\right)}$$

which completes the proof.