Spurious regression under broken trend stationarity

Noriega, Antonio E. and Ventosa Santaulària, Daniel

Escuela de Economia, Universidad de Guanajuato. UCEA

2005

Spurious regression under broken trend stationarity

Antonio E. Noriega[1]
Daniel Ventosa-Santaulària

Escuela de Economia
Universidad de Guanajuato

JEL Classification: *C22*
Keywords: Stationarity, Structural breaks, Spurious regression

Abstract

We study the phenomenon of spurious regression between two random variables, when the generating mechanism of individual series is assumed to follow a stationary process around a trend with (possibly) multiple breaks in the level and slope of trend. We develop the relevant asymptotic theory and show that the phenomenon of spurious regression occurs independently of the structure assumed for the errors. In contrast to previous findings, the presence of a spurious relationship will be less severe when breaks are present in the generating mechanism of individual series. This is true whether the regression model includes a linear trend or not. Simulations confirm our asymptotic results, and reveal that in finite samples, the phenomenon of spurious regression is sensitive to the presence of a linear trend in the regression model, and to the relative location of breaks within the sample.

[1]Corresponding author. Escuela de Economia, Universidad de Guanajuato. UCEA, Campus Marfil, Fracc. I, El Establo, Guanajuato, Gto. 36250, México. Phone and Fax +52-(473)7352925, e-mail: noriegam@quijote.ugto.mx.

1 Introduction

The spurious regression phenomenon occurs when there is a statistically significant relationship between two independent random variables. Since the contribution of Granger and Newbold (1974) on the issue of spurious regressions in econometrics, several articles have investigated the phenomenon under a variety of structures for the data generating processes. Phillips (1986) assumes that the individual series in a spurious regression are driftless random walks, while Marmol (1995) extends this to the general $I(d)$ case, with d being an integer number. Marmol (1996) studies the spurious regression problem with different orders of integration of the dependent and independent variables. Entorf (1997) analyzes random walks with drifts, given the relevance of such models, as argued by Nelson and Plosser (1982). Granger et. al. (1998) extend the analysis to positively autocorrelated autoregressive series on long moving averages, and Marmol (1998) and Tsay and Chung (1999) to long memory fractional integration processes.

More recently, Kim, Lee and Newbold (2004) (KLN henceforth), show that the phenomenon of spurious regression is still present even when the nonstationarity in individual series is of a deterministic nature. In particular, under the assumption of stationarity around a linear trend for the individual series, they show that the ordinary least squares estimator $\widehat{\delta}_1$ converges to the ratio of the trend coefficients, and $t_{\widehat{\delta}_1}$ (the t-statistic for the null hypothesis $H_0 : \delta_1 = 0$) diverges at rate $T^{3/2}$ in the following (spurious) regression:

$$y_t = \widehat{\alpha}_1 + \widehat{\delta}_1 x_t + \widehat{u}_t \tag{1}$$

where the DGP for each series is:

$$\begin{aligned} y_t &= \mu_y + \beta_y t + u_{yt}; & x_t &= \mu_x + \beta_x t + u_{xt}; \\ u_{yt} &= \phi_y u_{yt-1} + \varepsilon_{yt},\ \left|\phi_y\right| < 1; & u_{xt} &= \phi_x u_{xt-1} + \varepsilon_{xt},\ \left|\phi_x\right| < 1 \end{aligned} \tag{2}$$

and where ε_{zt} are $iid(0, \sigma_z^2)$ for $z = y, x$, and independent of each other. In this setting, KLN show that $\widehat{\delta}_1$ is a consistent estimator of $\frac{\beta_y}{\beta x}$.

In this paper, we consider a more general setting by allowing for structural breaks in both level and slope of trend in the Data Generating Process (DGP). The relevance of this analysis stems from the ample evidence of infrequent breaks in economic time series and the development of methods for estimating models with (single and multiple) structural breaks.[2] In a recent paper, Hansen (2001) argues that "The econometrics of structural change, ..., has dramatically altered the face of applied time series econometrics. (p.118)." Furthermore, recent theoretical research on economic growth has introduced mechanisms that induce

[2] See Perron (1989, 1992, 1997), Duck (1992), Raj (1992), Zivot and Andrews (1992), Stock (1994), Zelhorst and Haan (1995), Ben-David and Papell (1995, 1998), Lumsdaine and Papell (1997), Bai (1997a, b), Bai and Perron (1998a, b), Clemente, et.al. (1998), Noriega and de Alba (2001), and Perron and Zhu (2002).

multiple growth states, the theoretical counterpart of the broken trend models presented below (see Durlauf (1993), Cooper (1994), Acemoglu and Scott (1997), Lau (1997), Startz (1998), and Kejak (2003)).

In section 2 we derive the asymptotic behavior of statistics in regressions where the DGP consists of two independent processes with (possibly) multiple breaks, and show that the phenomenon of spurious regression is present in this setting, regardless of the errors' structure. Section 3 concludes.

2 Asymptotics for spurious regressions

2.1 The case of a single break

When allowing for a break in the individual series, the DGP in (2) can be extended to:

$$\begin{aligned} y_t &= \mu_y + \theta_y DU_{yt} + \beta_y t + \gamma_y DT_{yt} + u_{yt} \\ x_t &= \mu_x + \theta_x DU_{xt} + \beta_x t + \gamma_x DT_{xt} + u_{xt} \\ u_{yt} &= \ \phi_y u_{yt-1} + \varepsilon_{yt},\ \left|\phi_y\right| < 1; u_{xt} = \ \phi_x u_{xt-1} + \varepsilon_{xt},\ \left|\phi_x\right| < 1 \end{aligned} \tag{3}$$

where DU_{zt} and DT_{zt}, $(z = y, x)$ are dummy variables allowing changes in the trend's level and slope respectively, that is, $DU_{zt} = \mathbf{1}(t > T_{b_z})$ and $DT_{zt} = (t - T_{b_z})\mathbf{1}(t > T_{b_z})$, where $\mathbf{1}(\cdot)$ is the indicator function, and T_{b_z} is the unknown date of the break in z. We maintain the same structure for the innovations u_{yt} and u_{xt} as in KLN, although it can also be assumed that innovations obey the (general-level) conditions stated in Phillips (1986, p. 313).

We start with a lemma which collects useful results for subsequent analysis (all proofs are provided in the appendix).

Lemma 1. *Suppose $\{y_t\}_1^\infty$ and $\{x_t\}_1^\infty$ are generated by (3) with $\lambda_z = (T_{b_z}/T) \in (0,1)$, $z = y, x$. Then, as $T \to \infty$,*

$$T^{-2}\sum z_t \xrightarrow{p} \tfrac{1}{2}\left[\beta_z + (1-\lambda_z)^2\gamma_z\right] \equiv d_z$$

$$T^{-3}\sum y_t x_t \xrightarrow{p} \tfrac{1}{3}\beta_x\beta_y + \beta_x\gamma_y\lambda_y^+ + \beta_y\gamma_x\lambda_x^+ + \gamma_x\gamma_y\lambda \equiv g$$

$$T^{-3}\sum z_t^2 \xrightarrow{p} \tfrac{1}{3}\beta_z^2 + \tfrac{1}{3}(1-\lambda_z)^3\gamma_z^2 + 2\beta_z\gamma_z\lambda_z^+ \equiv g_z$$

$$T^{-3}\sum_{t=1}^T t z_t \xrightarrow{p} \tfrac{1}{3}\beta_z + \gamma_z\lambda_z^+ \equiv \psi_z$$

where

$\lambda_z^+ = \frac{1}{6}\left(\lambda_z + 2\right)\left(\lambda_z - 1\right)^2$
$\lambda = \frac{1}{3}\left(1-\lambda_u\right)^3 + \frac{1}{2}\lambda_d\left(1-\lambda_u\right)^2$
$\lambda_u = \max(\lambda_y, \lambda_x)$
$\lambda_l = \min(\lambda_y, \lambda_x)$
$\lambda_d = \lambda_u - \lambda_l$

and $\xrightarrow{p}$ signifies convergence in probability.

The following theorem collects the asymptotic behavior of the estimated parameters and associated $t-$statistics in model (1). It shows that $\widehat{\delta}_1$ does not converge to its true value of zero; instead, it converges to a constant, while its $t-$statistic diverges to infinity.

Theorem 1. *Let y_t and x_t be generated according to equation (3) and denote by $\widehat{\alpha}_1$,* and *$\widehat{\delta}_1$ the OLS estimates of α_1* and *δ_1 in (1). Then, as $T \to \infty$,*

a) $\widehat{\delta}_1 \xrightarrow{p} \delta_1 \equiv (g - d_x d_y)\left(g_x - d_x^2\right)^{-1}$

b) $T^{-1}\widehat{\alpha}_1 \xrightarrow{p} \alpha_1 \equiv (g_x d_y - g d_x)\left(g_x - d_x^2\right)^{-1}$

c) $T^{-1/2} t_{\widehat{\delta}_1} \xrightarrow{p} \delta_1 \left[\sigma_u^2 \left(g_x - d_x^2\right)^{-1}\right]^{-1/2}$

d) $T^{-1/2} t_{\widehat{\alpha}_1} \xrightarrow{p} \alpha_1 \left[\sigma_u^2 g_x \left(g_x - d_x^2\right)^{-1}\right]^{-1/2}$

where, σ_u^2 is defined in the appendix.

In Entorf (1997) and KLN, the limit of $\widehat{\delta}_1$ is $\frac{\beta_y}{\beta_x}$. The corresponding expression in part a) simplifies to this when $\gamma_y = \gamma_x = 0$. This can be easily seen by rewriting the limit of $\widehat{\delta}_1$ in part a) of theorem 1 as:

$$\delta_1 = \left[\frac{\beta_y}{\beta_x}\left(1 + l_x \frac{\gamma_x}{\beta_x}\right) + \frac{\gamma_y}{\beta_x}\left(l_y + l\frac{\gamma_x}{\beta_x}\right)\right]\left[1 + \frac{\gamma_x}{\beta_x}\left(2l_x + l_1 \frac{\gamma_x}{\beta_x}\right)\right]^{-1}$$

where

$l_z = (2\lambda_z + 1)(\lambda_z - 1)^2$
$l = 12\lambda - 3(\lambda_x - 1)^2 (\lambda_y - 1)^2,$
$l_1 = (3\lambda_x + 1)(1 - \lambda_x)^3.$

Thus, the limit of $\widehat{\delta}_1$ is a function of several ratios: trend magnitudes ($\frac{\beta_y}{\beta_x}$), relative (own) size of the break ($\frac{\gamma_z}{\beta_z}$), and relative (cross) size of the break ($\frac{\gamma_y}{\beta_x}$). According to part b) the constant term does not converge to a constant, instead it diverges at rate T. Part c) shows that, as in Phillips (1987), Entorf (1997), and KLN, the $t-$statistic diverges. The rate of divergence is the same as the one found by Phillips (1987), and slower than that of Entorf (1997) (T) and KLN ($T^{3/2}$). Therefore, we should expect less evidence of spurious regression when the generating process allows for structural breaks. It also should be noted that, in contrast to KLN, the limit of the normalized t-statistic is asymptotically invariant to the specification of the errors in the DGP (3). This is because the trend components in σ_u^2 dominate asymptotically, while in KLN they cancel

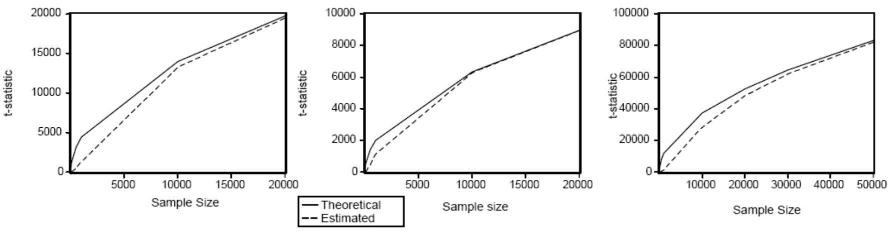

Figure 1: Graphs of t-statistics based on Theorem 1 and simulated data. Left panel: $\lambda_y = 0.55, \lambda_x = 0.45$; middle: $\lambda_y = 0.75, \lambda_x = 0.25$; right: $\lambda_y = 0.95, \lambda_x = 0.05$; all panels: $\beta_y = \beta_x = 0.2, \gamma_y = \gamma_x = 0.02$.

each other out. Thus, in large samples, the phenomenon of spurious regression will prevail when there are breaks in the individual series, and not necessarily as a consequence of the errors' specification, as in KLN.

Simulation experiments revealed that the finite sample rejection rate of the t-statistic in part c) of the above theorem is roughly the same whether variables have structural breaks or not. Rejection rates presented by KLN remain true in the presence of breaks. To illustrate our asymptotic results, we compare the theoretical t-statistic (computed from the above formula in Theorem 1) with the t-statistic computed from simulated data, based on the DGP in (3), for different sample sizes. Results are shown in Figure 1.

As can be seen, the proximity between the asymptotic and the simulated statistics varies significantly depending on the values of λ_y and λ_x. From these three particular examples, the difference between the simulated and the theoretical statistics is greater when the break dates are closer to the sample endpoints (right panel). Although in can not be inferred from the figure, our numerical simulations show that, for a sample size as small as 25 observations, the values of the t-statistic (both theoretical and simulated) are greater than 5.0, for all three analyzed examples, indicating a spurious relationship between y_t and x_t.

We now consider the case of a regression which allows for a linear trend. As discussed in KLN (2003), when the trend components in the individual series are sufficiently large to be detected, the applied researcher will run the following regression

$$y_t = \widehat{\alpha}_2 + \widehat{\beta}_2 t + \widehat{\delta}_2 x_t + \widehat{u}_t \tag{4}$$

where the DGP for each series is again (3). The next theorem presents the asymptotic behavior of the estimated parameters and associated $t-$statistics.

THEOREM 2. *Let y_t and x_t be generated according to equation (3) and denote by $\widehat{\alpha}_2$, $\widehat{\beta}_2$,* and *$\widehat{\delta}_2$ the OLS estimates of α_2, β_2* and *δ_2 in (4). Then, as $T \to \infty$,*

a) $\widehat{\delta}_2 \xrightarrow{p} \delta_2 \equiv \left[(\frac{1}{2}\psi_x - \frac{1}{3}d_x)d_y + (\frac{1}{2}d_x - \psi_x)\psi_y + \frac{1}{12}g\right] d^{-1}$

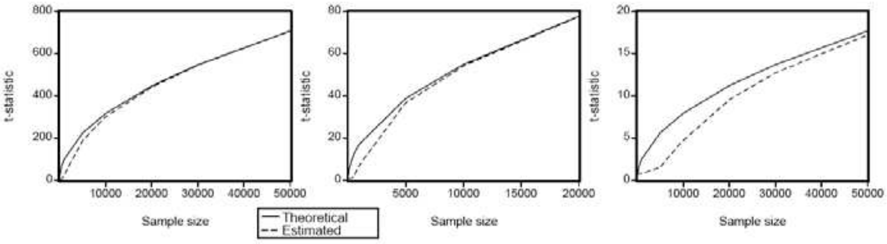

Figure 2: Graphs of t-statistics based on Theorem 2 and simulated data. Left panel: $\lambda_y = 0.55, \lambda_x = 0.45$; middle: $\lambda_y = 0.75, \lambda_x = 0.25$; right: $\lambda_y = 0.95, \lambda_x = 0.05$; all panels: $\beta_y = \beta_x = 0.2, \gamma_y = \gamma_x = 0.02$.

b) $T^{-1}\widehat{\alpha}_2 \xrightarrow{p} \alpha_2 \equiv \left[\left(\frac{1}{3}g_x - \psi_x^2\right) d_y + \left(d_x\psi_x - \frac{1}{2}g_x\right)\psi_y + \left(\frac{1}{2}\psi_x - \frac{1}{3}d_x\right) g\right] d^{-1}$

c) $\widehat{\beta}_2 \xrightarrow{p} \beta_2 \equiv \left[\left(d_x\psi_x - \frac{1}{2}g_x\right) d_y + \left(g_x - d_x^2\right)\psi_y + \left(\frac{1}{2}d_x - \psi_x\right) g\right] d^{-1}$

d) $T^{-1/2}t_{\widehat{\delta}_2} \xrightarrow{p} \frac{1}{12}\delta_2 \left[\eta_u^2 d^{-1}\right]^{-1/2}$

e) $T^{-1/2}t_{\widehat{\alpha}_2} \xrightarrow{p} \alpha_2 \left[\eta_u^2 \left(\frac{1}{3}g_x - \psi_x^2\right) d^{-1}\right]^{-1/2}$

f) $T^{-1/2}t_{\widehat{\beta}_2} \xrightarrow{p} \beta_2 \left[\eta_u^2 \left(g_x - d_x^2\right) d^{-1}\right]^{-1/2}$

where,
$d \equiv \psi_x(d_x - \psi_x) + \frac{1}{12}g_x - \frac{1}{3}d_x^2$, and η_u^2 *is defined in the appendix.*

As in Theorem 1, $\widehat{\delta}_2$ does not converge to its true value of zero, while the constant diverges to infinity at rate T. The trend parameter also converges to a constant. Part d) shows, as before, that the t-statistic diverges, thus indicating a spurious relationship between y and x.

Figure 2 presents a comparison between the asymptotic t-statistic, based on part d) of Theorem 2, and a simulated one based on the DGP (3).

As opposed to Figure 1, the difference between the simulated and the theoretical statistics in Figure 2 is greater when the break dates are closer to each other (left panel). Table 1 presents the behavior of the $t-$statistic for smaller sample sizes, and the three representative possibilities regarding the break fractions. It shows that when the breaks are far from each other, Theorem 2 implies that nearly $1,000$ observations are required in order to reject the null of no relationship at the 1% level. From the last two columns of the table, it is clear that the asymptotic value is a poor approximation to the simulated one for samples below $10,000$ (for which both statistics reject the null). On the other hand, the closer the break dates are to each other, the more likely the phenomenon of spurious regression is. For instance, if the breaks occur near the middle of the sample, the simulated (and theoretical) value(s) exceeds 2.0 for T just below 300. Therefore, the further away the breaks are from each other, the more observations are required to generate the phenomenon of spurious regression (more than 5000 in the right panel of the Table).

	$\lambda_{y=0.55}$	$\lambda_{x=0.45}$	$\lambda_{y=0.75}$	$\lambda_{x=0.25}$	$\lambda_{y=0.95}$	$\lambda_{x=0.05}$
T	Theoret.	Simul.	Theoret.	Simul.	Theoret.	Simul.
25	15.80	0.83	2.75	0.83	0.40	0.83
50	22.35	0.79	3.88	0.79	0.56	0.79
100	31.61	0.80	5.49	0.79	0.79	0.79
300	54.74	2.59	9.51	0.93	1.37	0.80
600	77.42	10.79	13.46	2.86	1.94	0.79
1,000	99.95	26.18	17.37	7.29	2.50	0.77
5,000	223.49	186.25	38.84	36.67	5.59	1.52
10,000	316.07	300.13	54.9	54.13	7.91	4.71

Table 1. t-statistics computed from: part d) of Theorem 2, and simulated data. Same value of parameters as in experiments for Figure 2.

From comparing the results of this simulation with the previous ones, it emerges that it is more likely in small samples to observe a spurious regression when not including a linear trend in the regression model.

2.2 The case of multiple breaks

In this section, we consider an even more general setting by allowing for multiple structural breaks in both level and slope of trend in the DGP. In particular, we assume that:

$$\begin{aligned} y_t &= \mu_y + \sum_{i=1}^{Ny} \theta_{iy} DU_{iyt} + \beta_y t + \sum_{i=1}^{My} \gamma_{iy} DT_{iyt} + u_{yt} \\ x_t &= \mu_x + \sum_{i=1}^{Nx} \theta_{ix} DU_{ixt} + \beta_x t + \sum_{i=1}^{Mx} \gamma_{ix} DT_{ixt} + u_{xt} \\ u_{yt} =& \ \ \phi_y u_{yt-1} + \varepsilon_{yt}, \ \left|\phi_y\right| < 1; u_{xt} = \ \ \phi_x u_{xt-1} + \varepsilon_{xt}, \ \left|\phi_x\right| < 1 \end{aligned} \tag{5}$$

where, N_z, M_z, the number of breaks in y_t and x_t, comprise the set of natural numbers, $for\ z = x, y$.

The following lemma collects useful results for subsequent analysis. All sums run from $t = 1$ to T, unless otherwise stated.

Lemma 2. *Suppose* $\{y_t\}_1^\infty$ *and* $\{x_t\}_1^\infty$ *are generated by (5) with* $\lambda_{iz} = (T_{bi_z}/T) \in (0,1), z = y, x, i = 1, 2..., M_z$. *Then, as* $T \to \infty$,

$T^{-2} \sum z_t \xrightarrow{p} \frac{1}{2} \left[\beta_z + \sum\limits_{i=1}^{Mz} (1 - \lambda_{iz})^2 \gamma_{iz} \right] \equiv \Delta_z$

$T^{-3}\sum y_t x_t \xrightarrow{p} \frac{1}{3}\beta_x\beta_y + \beta_x \sum_{i=1}^{My} \gamma_{iy}\lambda_{iy}^{+} + \beta_y \sum_{i=1}^{Mx} \gamma_{ix}\lambda_{ix}^{+} + \sum_{i=1}^{My}\sum_{j=1}^{Mx} \gamma_{iy}\gamma_{jx}\omega_{ij} \equiv \Gamma$

$T^{-3}\sum z_t^2 \xrightarrow{p} \frac{1}{3}\beta_z^2 + 2\beta_z \sum_{i=1}^{Mz} \gamma_{iz}\lambda_{iz}^{+} + \sum_{i=1}^{Mz}\sum_{j=1}^{Mz} \gamma_{iz}\gamma_{jz}\upsilon_{ij} \equiv \Gamma_z$

$T^{-3}\sum_{t=1}^{T} tz_t \xrightarrow{p} \frac{1}{3}\beta_z + \sum_{i=1}^{Mz} \gamma_{iz}\lambda_{iz}^{+} \equiv \Psi_z$

where

$\lambda_{iz}^{+} = \frac{1}{2}\lambda_{iz}(1-\lambda_{iz})^2 + \frac{1}{3}(1-\lambda_{iz})^3 = \frac{1}{6}\left(\lambda_{iz}+2\right)\left(\lambda_{iz}-1\right)^2$
$\omega_{ij} = \frac{1}{3}\left(1-\lambda_{u1,ij}\right)^3 + \frac{1}{2}\lambda_{d1,ij}\left(1-\lambda_{u1,ij}\right)^2$
$\lambda_{u1,ij} = \max(\lambda_{yi}, \lambda_{xj})$
$\lambda_{l1,ij} = \min(\lambda_{yi}, \lambda_{xj})$
$\lambda_{d1,ij} = \lambda_{u1,ij} - \lambda_{l1,ij}$

$\upsilon_{ij} = \frac{1}{3}\left(1-\lambda_{u2,ij}\right)^3 + \frac{1}{2}\lambda_{d2,ij}\left(1-\lambda_{u2,ij}\right)^2$
$\lambda_{u2,ij} = \max(\lambda_{xi}, \lambda_{xj})$
$\lambda_{l2,ij} = \min(\lambda_{xi}, \lambda_{xj})$
$\lambda_{d2,ij} = \lambda_{u2,ij} - \lambda_{l2,ij}$

The following theorem collects the asymptotic behavior of the estimated parameters and associated $t-$statistics in regression model (1).

THEOREM 3. *Let y_t and x_t be generated according to equation (5) and denote by* $\widehat{\alpha}_1$, and $\widehat{\delta}_1$ *the OLS estimates of* α_1 and δ_1 *in (1). Then, as* $T \to \infty$,

a) $\widehat{\delta}_1 \xrightarrow{p} \delta_1 \equiv \left(\Gamma - \Delta_x\Delta_y\right)\left(\Gamma_x - \Delta_x^2\right)^{-1}$

b) $T^{-1}\widehat{\alpha}_1 \xrightarrow{p} \alpha_1 \equiv \left(\Gamma_x\Delta_y - \Gamma\Delta_x\right)\left(\Gamma_x - \Delta_x^2\right)^{-1}$

c) $T^{-1/2}t_{\widehat{\delta}_1} \xrightarrow{p} \delta_1\left[\sigma_u^2\left(\Gamma_x - \Delta_x^2\right)^{-1}\right]^{-1/2}$

d) $T^{-1/2}t_{\widehat{\alpha}_1} \xrightarrow{p} \alpha_1\left[\sigma_u^2\Delta_x^2\left(\Gamma_x - \Delta_x^2\right)^{-1}\right]^{-1/2}$

where, σ_u^2 *is defined in the appendix.*

The theorem shows that the estimated parameters and t-statistics have similar asymptotic behavior as in the case of a single break. Hence, the phenomenon of spurious regression is present in large samples under multiple breaks.

We consider again the case of a regression which allows for a linear trend, as in the previous section. The next theorem presents the asymptotic behavior of the estimated parameters and associated $t-$statistics.

THEOREM 4. *Let y_t and x_t be generated according to equation (5) and denote by* $\widehat{\alpha}_2$, $\widehat{\beta}_2$, and $\widehat{\delta}_2$ *the OLS estimates of* α_2, β_2 and δ_2 *in (4). Then, as* $T \to \infty$,

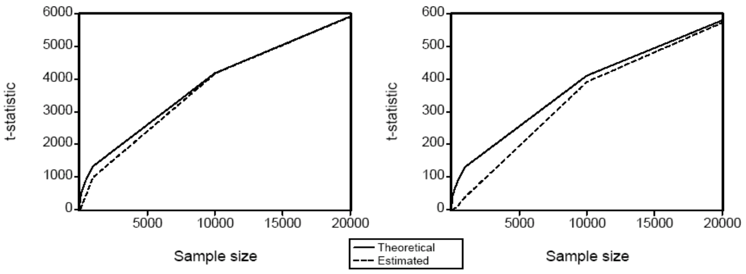

Figure 3: Graphs of t-statistics based on Theorems 3 and 4 and simulated data. Left panel based on regression model (1), while right panel based on model (4). Both panels: $M_y = 1, M_x = 3; \lambda_{y,1} = 0.5$, $\lambda_{1x} = 0.2, \lambda_{2x} = 0.4, \lambda_{3x} = 0.6; \beta_y = \beta_x = 0.2, \gamma_{iy} = \gamma_{ix} = 0.02$.

a) $\widehat{\delta}_2 \xrightarrow{p} \delta_2 \equiv \left[(\frac{1}{2}\Psi_x - \frac{1}{3}\Delta_x)\Delta_y + (\frac{1}{2}\Delta_x - \Psi_x)\Psi_y + \frac{1}{12}\Gamma\right] D^{-1}$

b) $T^{-1}\widehat{\alpha}_2 \xrightarrow{p} \alpha_2 \equiv \left[\left(\frac{1}{3}\Gamma_x - \Psi_x^2\right)\Delta_y + \left(\Delta_x\Psi_x - \frac{1}{2}\Gamma_x\right)\Psi_y + \left(\frac{1}{2}\Psi_x - \frac{1}{3}\Delta_x\right)\Gamma\right] D^{-1}$

c) $\widehat{\beta}_2 \xrightarrow{p} \beta_2 \equiv \left[\left(\Delta_x\Psi_x - \frac{1}{2}\Gamma_x\right)\Delta_y + \left(\Gamma_x - \Delta_x^2\right)\Psi_y + \left(\frac{1}{2}\Delta_x - \Psi_x\right)\Gamma\right] D^{-1}$

d) $T^{-1/2}t_{\widehat{\delta}_2} \xrightarrow{p} \frac{1}{12}\delta_2\left[\eta_u^2 D^{-1}\right]^{-1/2}$

e) $T^{-1/2}t_{\widehat{\alpha}_2} \xrightarrow{p} \alpha_2\left[\eta_u^2\left(\frac{1}{3}\Gamma_x - \psi_x^2\right)D^{-1}\right]^{-1/2}$

f) $T^{-1/2}t_{\widehat{\beta}_2} \xrightarrow{p} \beta_2\left[\eta_u^2\left(\Gamma_x - \Delta_x^2\right)D^{-1}\right]^{-1/2}$

where,
$D \equiv \Psi_x(\Delta_x - \Psi_x) + \frac{1}{12}\Gamma_x - \frac{1}{3}\Delta_x^2$, and η_u^2 *is defined in the appendix.*

Results are in line with those of Theorem 2. To asses the validity of our theoretical findings, the left panel in Figure 3 shows a comparison between the asymptotic t-statistic, based on part c) of Theorem 3, and a simulated one based on the DGP (5), while the right panel shows a comparison between the asymptotic t-statistic, based on part d) of Theorem 4, and the same DGP used for the left panel.

As can be seen, convergence between the theoretical and simulated statistics is achieved at around $10,000$ observations in both cases. However, the values of the t-statistics when a regression without trend is used are roughly 10 times larger than those for a spurious regression with a linear trend. According to our simulation experiments, this occurs for a variety of experimental designs (changing the number, location and size of breaks in both variables). Therefore, it is much more likely to find a spurious relationship when the regression equation does not contain a linear trend. This is also the case when analyzing smaller samples (not shown): while the spurious regression phenomenon is present with samples as low as 25 for a regression model without a linear trend, when a trend is included the rejections start to occur with 200 observations, similar to the case of a single break.

3 Conclusions

This paper has investigated the spurious regression phenomenon when the processes generating the individual series are stationary around a linear trend subject to (possibly) multiple structural breaks. Our results indicate that, whether there is a single break in each series or multiple breaks in both, the phenomenon of spurious regression will occur asymptotically, independently of the errors' structure in the data generating process, contrary to results of Kim, Lee and Newbold (2004). However, the rate of divergence of the t-statistic was found to be lower ($T^{1/2}$) than in the case of stationary processes around linear trends without breaks ($T^{3/2}$), implying that the presence of a spurious relationship will be less severe when breaks are present in the generating mechanism of individual series. This is true whether the regression model includes a linear trend or not.

Simulations confirm our asymptotic results: it takes several thousand observations for the theoretical t-statistic to converge to the simulated one, under a variety of locations of breaks. They also reveal that in finite samples, the phenomenon of spurious regression is sensitive to the presence of a linear trend in the regression model, and to the relative location of breaks in the generating mechanism. In particular, a spurious rejection necessitates hundreds of observations when the regression includes a linear trend, while it occurs with very small samples when the trend is not present. In other words, it is more likely to find a statistical significant relationship between two independent stationary series subject to breaks when the regression model does not include a linear trend in its deterministic specification.

4 Appendix

Proof of Lemma 1

From the DGP (3) we have
$\sum z_t = T\mu_z + \theta_z \sum DU_{zt} + \beta_z \sum t + \gamma_z \sum DT_{zt} + \sum u_{zt}$
where
$\sum DU_{zt} = (1-\lambda_z)T$, $\sum t = \frac{1}{2}(T^2+T)$, and $\sum DT_{zt} = \sum_1^{(1-\lambda_z)T} t = \frac{1}{2}(1-\lambda_z)\left[(1-\lambda_z)T^2+T\right]$.
Hence,
$\sum z_t = \frac{1}{2}\left[\beta_z + (1-\lambda_z)^2\gamma_z\right]T^2 + O_p(T)$,
and
$T^{-2}\sum z_t = \frac{1}{2}\left[\beta_z + (1-\lambda_z)^2\gamma_z\right] + O_p(T^{-1})$.
To prove the second part we write, from equations (3)

$$\sum y_t x_t = \beta_x\beta_y \sum t^2 + \beta_x\gamma_y \sum DT_{yt}t + \beta_y\gamma_x \sum DT_{xt}t + \gamma_x\gamma_y \sum DT_{yt}DT_{xt} + O_p(T^2),$$

where

$$\sum t^2 = \frac{1}{3}T^3 + O(T^2),$$

$$\sum DT_{zt}t = \sum_1^{(1-\lambda_z)T} t(\lambda_z T + t) = \lambda_z^+ T^3 + O(T^2),$$

$$\sum DT_{yt}DT_{xt} = \sum_1^{(1-\lambda_u)T} t[t+\lambda_d T)] = \left[\frac{1}{3}(1-\lambda_u)^3 + \frac{1}{2}(1-\lambda_u)^2\lambda_d\right]T^3 + O(T^2).$$

Therefore,

$$T^{-3}\sum y_t x_t$$
$$= \frac{1}{3}\beta_x\beta_y + \beta_x\gamma_y\lambda_y^+ + \beta_y\gamma_x\lambda_x^+ + \gamma_x\gamma_y\left[\frac{1}{3}(1-\lambda_u)^3 + \frac{1}{2}(1-\lambda_u)^2\lambda_d\right] + O_p(T^{-1}).$$

To prove the third part, we write

$$\sum z_t^2 = \beta_z^2 \sum t^2 + \gamma_z^2 \sum DT_{zt}^2 + 2\beta_z\gamma_z \sum DT_{zt}t + O_p(T^2),$$

where
$\sum DT_{zt}^2 = \sum_1^{(1-\lambda_z)T} t^2 = \frac{1}{3}(1-\lambda_z)^3T^3 + O(T^2)$.
Therefore,
$T^{-3}\sum z_t^2 = \frac{1}{3}\beta_z^2 + \frac{1}{3}(1-\lambda_z)^3\gamma_z^2 + 2\lambda_z^+\beta_z\gamma_z + O_p(T^{-1})$.
The last part can be proven as follows

$$\sum_{t=1}^T tz_t = \sum_{t=1}^T t(\mu_z + \theta_z DU_{zt} + \beta_z t + \gamma_z DT_{zt} + u_{zt})$$
$$= \left[\frac{1}{3}\beta_z + \frac{1}{3}(1-\lambda_z)^3\gamma_z + \frac{1}{2}\gamma_z\lambda_z(1-\lambda_z)^2\right]T^3 + O_p(T^2).$$

Therefore,
$T^{-3}\sum_{t=1}^T tz_t = \frac{1}{3}\beta_z + \gamma_z\lambda_z^+ + O_p(T^{-1})$.

Proof of Theorem 1

Write the regression model $y_t = \alpha_1 + \delta_1 x_t + u_t$ in matrix form:

$$y = X\beta + u$$

The vector of OLS estimators is defined as:

$$\widehat{\beta} = (X'X)^{-1}X'y = \left[T\sum x_t^2 - \left(\sum x_t\right)^2\right]^{-1} \left\{ \begin{array}{c} \sum x_t^2 \sum y_t - \sum x_t \sum x_t y_t \\ T\sum x_t y_t - \sum x_t \sum y_t \end{array} \right\}$$

Using results from lemma 1,

$$\left\{ \begin{array}{c} T^{-1}\widehat{\alpha}_1 \\ \widehat{\delta}_1 \end{array} \right\} \xrightarrow{p} \left[(g_x - d_x^2)\right]^{-1} \left\{ \begin{array}{c} (g_x d_y - g d_x) \\ (g - d_x d_y) \end{array} \right\} \equiv \left\{ \begin{array}{c} \alpha_1 \\ \delta_1 \end{array} \right\},$$

proving parts a) and b).
To prove c) and d) write the t-statistics as:

$$t_{\widehat{\alpha}} = \widehat{\alpha}_1 \left[\widehat{\sigma}_u^2 (X'X)_{11}^{-1}\right]^{-1/2},$$

$$t_{\widehat{\delta}} = \widehat{\delta}_1 \left[\widehat{\sigma}_u^2 (X'X)_{22}^{-1}\right]^{-1/2},$$

and $(X'X)_{ii}^{-1}$, the i^{th} diagonal element of $(X'X)^{-1}$, as

$$(X'X)_{11}^{-1} = \frac{T^{-3}\sum x_t^2 + o_p(1)}{T^{-4}\left(\sum x_t^2 - \left(\sum z_t\right)^2\right) + o_p(1)},$$

$$(X'X)_{22}^{-1} = \frac{T}{T^{-4}\left(\sum x_t^2 - \left(\sum z_t\right)^2\right) + o_p(1)}.$$

Note also that

$$\widehat{\sigma}_u^2 = T^{-1}\sum \widehat{u}_t^2 = T^{-1}\sum \left(y_t - \widehat{\alpha}_1 - \widehat{\delta}_1 x_t\right)^2$$

$$= T^{-1}\left(\sum y_t^2 + \widehat{\alpha}_1^2 T + \widehat{\delta}_1^2 \sum x_t^2 - 2\widehat{\alpha}_1 \sum y_t - 2\widehat{\delta}_1 \sum x_t y_t + 2\widehat{\alpha}_1 \widehat{\delta}_1 \sum x_t\right).$$

Using lemma 1,

$$T^{-2}\widehat{\sigma}_u^2 \xrightarrow{p} \left(g_y + \delta_1^2 g_x - 2\delta_1 g\right) + 2(\delta_1 d_x - d_y)\alpha_1 + \alpha_1^2 \equiv \sigma_u^2.$$

Thus,

$$T^{-1}\widehat{\alpha}_1 \left[T^{-2}\widehat{\sigma}_u^2 T (X'X)_{11}^{-1}\right]^{-1/2} = T^{-1/2}\widehat{\alpha}_1 \left[\widehat{\sigma}_u^2 (X'X)_{11}^{-1}\right]^{-1/2} = O_p(1),$$

and

$$\widehat{\delta}_1 \left[T^{-2}\widehat{\sigma}_u^2 T^3 (X'X)_{22}^{-1}\right]^{-1/2} = T^{-1/2}\widehat{\delta}_1 \left[\widehat{\sigma}_u^2 (X'X)_{22}^{-1}\right]^{-1/2} = O_p(1).$$

The limits of these two expressions yield the formulae in c) and d).

Proof of Theorem 2

Write the regression model $y_t = \alpha_2 + \beta_2 t + \delta_2 x_t + u_t$ in matrix form:

$$y = X\beta + u$$

The vector of OLS estimators is $\widehat{\beta} = (X'X)^{-1}X'y$, and we define

$$X'X = \left\{ \begin{array}{ccc} a & b & c \\ b & d & m \\ c & m & n \end{array} \right\},$$

where
$a = T$, $b = \sum_{t=1}^T t = \frac{1}{2}T^2 + O(T)$, $c = \sum_{t=1}^T x_t = d_x T^2 + O(T)$,
$d = \sum_{t=1}^T t^2 = \frac{1}{3}T^3 + O(T^2)$, $m = \sum_{t=1}^T t x_t = \psi_x T^3 + O(T^2)$, $n = \sum_{t=1}^T x_t^2 = g_x T^3 + O(T^2)$.
and

$$\widehat{\beta} = [\det(X'X)]^{-1} \begin{Bmatrix} dn - m^2 & cm - bn & bm - cd \\ cm - bn & an - c^2 & bc - am \\ bm - cd & bc - am & ad - b^2 \end{Bmatrix} \begin{Bmatrix} \sum_{t=1}^T y_t \\ \sum_{t=1}^T t y_t \\ \sum_{t=1}^T x_t y_t \end{Bmatrix}$$

with

$$\begin{aligned} \det(X'X) &= 2bcm + adn - c^2 d - am^2 - b^2 n \\ &= 2\sum_{t=1}^T t \sum_{t=1}^T x_t \sum_{t=1}^T t x_t + T \sum_{t=1}^T t^2 \sum_{t=1}^T x_t^2 - \left(\sum_{t=1}^T x_t\right)^2 \sum_{t=1}^T t^2 \\ &\quad - T\left(\sum_{t=1}^T t x_t\right)^2 - \left(\sum_{t=1}^T t\right)^2 \sum_{t=1}^T x_t^2 \equiv \det_x = O(T^7) \end{aligned}$$

Using results from lemma 1,

$T^{-7}\det(X'X) \xrightarrow{p} \psi_x(d_x - \psi_x) + \frac{1}{12} g_x - \frac{1}{3} d_x^2$

To prove parts a)-c) we derive expression for the numerators of $\widehat{\delta}_2, \widehat{\alpha}_2$, and $\widehat{\beta}_2$ as follows:

$$\begin{aligned} num\widehat{\delta}_2 &= (bm - cd)\sum_{t=1}^T y_t + (bc - am)\sum_{t=1}^T t y_t + (ad - b^2)\sum_{t=1}^T x_t y_t = O(T^7) \\ num\widehat{\alpha}_2 &= (dn - m^2)\sum_{t=1}^T y_t + (cm - bn)\sum_{t=1}^T t y_t + (bm - cd)\sum_{t=1}^T x_t y_t = O(T^8) \\ num\widehat{\beta}_2 &= (cm - bn)\sum_{t=1}^T y_t + (an - c^2)\sum_{t=1}^T t y_t + (bc - am)\sum_{t=1}^T x_t y_t = O(T^7) \end{aligned}$$

Using lemma 1 and the above definitions it is easy to show that

$$\begin{aligned} T^{-7} num\widehat{\delta}_2 &\xrightarrow{p} (\tfrac{1}{2}\psi_x - \tfrac{1}{3}d_x)d_y + (\tfrac{1}{2}d_x - \psi_x)\psi_y + \tfrac{1}{12}g \\ T^{-8} num\widehat{\alpha}_2 &\xrightarrow{p} \left(\tfrac{1}{3}g_x - \psi_x^2\right)d_y + \left(d_x\psi_x - \tfrac{1}{2}g_x\right)\psi_y + \left(\tfrac{1}{2}\psi_x - \tfrac{1}{3}d_x\right)g \\ T^{-7} num\widehat{\beta}_2 &\xrightarrow{p} \left(d_x\psi_x - \tfrac{1}{2}g_x\right)d_y + \left(g_x - d_x^2\right)\psi_y + \left(\tfrac{1}{2}d_x - \psi_x\right)g \end{aligned}$$

Combining these expression with that for $\det(X'X)$ gives the results in a)-c).

To prove d)-f) write the t-statistics as:

$$\begin{aligned} t_{\widehat{\delta}_2} &= \widehat{\delta}_2 \left[\widehat{\eta}_u^2 (X'X)_{33}^{-1}\right]^{-1/2}, \\ t_{\widehat{\alpha}_2} &= \widehat{\alpha}_2 \left[\widehat{\eta}_u^2 (X'X)_{11}^{-1}\right]^{-1/2}, \\ t_{\widehat{\beta}_2} &= \widehat{\beta}_2 \left[\widehat{\eta}_u^2 (X'X)_{22}^{-1}\right]^{-1/2}, \end{aligned}$$

and $(X'X)_{ii}^{-1}$, the i^{th} diagonal element of $(X'X)^{-1}$, as

$$\begin{aligned} (X'X)_{33}^{-1} &= \frac{T^{-4}\left(\frac{1}{12}T^4\right) + o_p(1)}{T^{-7}\det_x + o_p(1)}, \\ (X'X)_{11}^{-1} &= \frac{T^{-6}\left(\frac{1}{3}\sum x_t^2 - \left(\sum_{t=1}^T t x_t\right)^2\right) + o_p(1)}{T^{-7}\det_x + o_p(1)}, \\ (X'X)_{22}^{-1} &= \frac{T^{-4}\left(\sum x_t^2 - \left(\sum_{t=1}^T x_t\right)^2\right) + o_p(1)}{T^{-7}\det_x + o_p(1)}. \end{aligned}$$

Note also that

$$\widehat{\eta}_u^2 = T^{-1}\sum \widehat{u}_t^2 = T^{-1}\sum \left(y_t - \widehat{\alpha}_2 - \widehat{\beta}_2 t - \widehat{\delta}_2 x_t\right)^2$$

Using lemma 1,

$T^{-2}\widehat{\eta}_u^2 \xrightarrow{p} \left(g_y + \frac{1}{3}\beta_2^2 + \delta_2^2 g_x - 2\beta_2\psi_y - 2\delta_2 g + 2\beta_2\delta_2\psi_x\right) + (\beta_2 + 2\delta_2 d_x - 2d_y)\alpha_2 + \alpha_2^2 \equiv \eta_u^2.$

Thus,

$$\widehat{\delta}_2 \left[T^{-2}\widehat{\eta}_u^2 T^3 (X'X)_{33}^{-1}\right]^{-1/2} = T^{-1/2}\widehat{\delta}_2 \left[\widehat{\eta}_u^2 (X'X)_{33}^{-1}\right]^{-1/2} = O_p(1),$$

$$T^{-1}\widehat{\alpha}_2 \left[T^{-2}\widehat{\eta}_u^2 T (X'X)_{11}^{-1}\right]^{-1/2} = T^{-1/2}\widehat{\alpha}_2 \left[\widehat{\eta}_u^2 (X'X)_{11}^{-1}\right]^{-1/2} = O_p(1),$$

$$\widehat{\beta}_2 \left[T^{-2}\widehat{\eta}_u^2 T^3 (X'X)_{22}^{-1}\right]^{-1/2} = T^{-1/2}\widehat{\beta}_2 \left[\widehat{\eta}_u^2 (X'X)_{22}^{-1}\right]^{-1/2} = O_p(1).$$

The limits of these expressions yield the formulae in d)-f).

Proof of Lemma 2

From the DGP (5) we have

$\sum z_t = T\mu_z + \theta_{1,z}\sum DU_{1,zt} + ... + \theta_{N_z,z}\sum DU_{N_z,zt} + \beta_z \sum t + \gamma_{1,z}\sum DT_{1,zt} + ... + \gamma_{M_z,z}\sum DT_{M_z,zt} + \sum u_{zt}$

where

$\sum DU_{i,zt} = (1-\lambda_{i,z})T$, and $\sum DT_{i,zt} = \sum_1^{(1-\lambda_{i,z})T} t = \frac{1}{2}(1-\lambda_{i,z})\left[(1-\lambda_{i,z})T^2 + T\right]$.

Hence,

$\sum z_t = \frac{1}{2}\left[\beta_z + \sum_{i=1}^{Mz}(1-\lambda_{i,z})^2\gamma_{i,z}\right]T^2 + O_p(T),$

and

$T^{-2}\sum z_t = \frac{1}{2}\left[\beta_z + \sum_{i=1}^{Mz}(1-\lambda_{i,z})^2\gamma_{i,z}\right] + O_p(T^{-1}).$

To prove the second part we write, from the *DGP* (5)

$\sum y_t x_t = \beta_x\beta_y\sum t^2 + \beta_x\sum\left(\sum_{i=1}^{My}\gamma_{i,y}DT_{i,yt}t\right) + \beta_y\sum\left(\sum_{i=1}^{Mx}\gamma_{i,x}DT_{i,xt}t\right) + \sum\left(\sum_{i=1}^{My}\gamma_{i,y}DT_{i,yt}\sum_{j=1}^{Mx}\gamma_{j,x}DT_{j,xt}\right)$
$+O_p(T^2),$

where

$$\sum DT_{i,zt}t = \sum_1^{(1-\lambda_{i,z})T} t(\lambda_{i,z}T + t) = \lambda_{i,z}^+ T^3 + O(T^2),$$

$$\sum DT_{i,yt}DT_{i,xt} = \sum_1^{(1-\lambda_{u1},ij)T} t[t+\lambda_{d1,ij}T)] = \left[\frac{1}{3}(1-\lambda_{u1,ij})^3 + \frac{1}{2}(1-\lambda_{u1,ij})^2\lambda_{d1,ij}\right]T^3 + O(T^2).$$

Therefore,

$T^{-3}\sum y_t x_t$
$= \frac{1}{3}\beta_x\beta_y + \beta_x\sum_{i=1}^{My}\gamma_{i,y}\lambda_{i,y}^+ + \beta_y\sum_{i=1}^{Mx}\gamma_{i,x}\lambda_{i,x}^+ + \sum_{i=1}^{My}\gamma_{i,y}\sum_{j=1}^{Mx}\gamma_{j,x}\omega_{ij} + O_p(T^{-1}).$

In order to obtain $\sum x_t^2$, we can simply replace y_t by x_t in $\sum y_t x_t$:

$T^{-3}\sum x_t^2 = \frac{1}{3}\beta_x^2 + 2\beta_x\sum_{i=1}^{Mx}\gamma_{i,x}\lambda_{i,x}^+ + \sum_{i=1}^{Mx}\sum_{j=1}^{Mx}\gamma_{i,x}\gamma_{j,x}\upsilon_{ij} + O_p(T^{-1}).$

Proof of Theorem 3

Write the regression model $y_t = \alpha_1 + \delta_1 x_t + u_t$ in matrix form:

$$y = X\beta + u$$

The vector of OLS estimators is defined as:

$$\widehat{\beta} = (X'X)^{-1}X'y = \left[T\sum x_t^2 - \left(\sum x_t\right)^2\right]^{-1} \begin{Bmatrix} \sum x_t^2 \sum y_t - \sum x_t \sum x_t y_t \\ T\sum x_t y_t - \sum x_t \sum y_t \end{Bmatrix}$$

Using results from lemma 2,

$$\begin{Bmatrix} T^{-1}\widehat{\alpha}_1 \\ \widehat{\delta}_1 \end{Bmatrix} \xrightarrow{p} \left[\Delta_x^2 - (\Delta_x)^2\right]^{-1} \begin{Bmatrix} \Delta_x^2\Delta_y - \Delta\Delta_x \\ \Delta - \Delta_x\Delta_y \end{Bmatrix} \equiv \begin{Bmatrix} \alpha_1 \\ \delta_1 \end{Bmatrix},$$

proving parts a) and b).

To prove c) and d) write the t-statistics as:

$$t_{\widehat{\alpha}_1} = \widehat{\alpha}_1 \left[\widehat{\sigma}_u^2 (X'X)_{11}^{-1}\right]^{-1/2},$$

$$t_{\widehat{\delta}_1} = \widehat{\delta}_1 \left[\widehat{\sigma}_u^2 (X'X)_{22}^{-1}\right]^{-1/2},$$

and $(X'X)_{ii}^{-1}$, the i^{th} diagonal element of $(X'X)^{-1}$, as

$$(X'X)_{11}^{-1} = \frac{T^{-3}\sum x_t^2 + o_p(1)}{T^{-4}\left(\sum x_t^2 - \left(\sum x_t\right)^2\right) + o_p(1)},$$

$$(X'X)_{22}^{-1} = \frac{T}{T^{-4}\left(\sum x_t^2 - \left(\sum x_t\right)^2\right) + o_p(1)}.$$

Note also that

$$\widehat{\sigma}_u^2 = T^{-1}\sum \widehat{u}_t^2 = T^{-1}\sum \left(y_t - \widehat{\alpha}_1 - \widehat{\delta}_1 x_t\right)^2$$

$$= T^{-1}\left(\sum y_t^2 + \widehat{\alpha}_1^2 T + \widehat{\delta}_1^2 \sum x_t^2 - 2\widehat{\alpha}_1 \sum y_t - 2\widehat{\delta}_1 \sum x_t y_t + 2\widehat{\alpha}_1\widehat{\delta}_1 \sum x_t\right).$$

Using lemma 2,

$$T^{-2}\widehat{\sigma}_u^2 \xrightarrow{p} \left(\Delta_y^2 + \delta_1^2\Delta_x^2 - 2\delta_1\Delta\right) + 2(\delta_1\Delta_x - \Delta_y)\alpha_1 + \alpha_1^2 \equiv \sigma_u^2,$$

And thus,

$$T^{-1}\widehat{\alpha}_1 \left[T^{-2}\widehat{\sigma}_u^2 T (X'X)_{11}^{-1}\right]^{-1/2} = T^{-1/2}\widehat{\alpha}_1 \left[\widehat{\sigma}_u^2 (X'X)_{11}^{-1}\right]^{-1/2} = O_p(1),$$

$$\widehat{\delta}_1 \left[T^{-2}\widehat{\sigma}_u^2 T^3 (X'X)_{22}^{-1}\right]^{-1/2} = T^{-1/2}\widehat{\delta}_1 \left[\widehat{\sigma}_u^2 (X'X)_{22}^{-1}\right]^{-1/2} = O_p(1).$$

The limits of these expressions yield the formulae in c)-d).

Proof of Theorem 4

Write the regression model $y_t = \alpha_2 + \beta_2 t + \delta_2 x_t + u_t$ in matrix form:

$$y = X\beta + u$$

The vector of OLS estimators is $\widehat{\beta} = (X'X)^{-1}X'y$, and we define

$$X'X = \begin{Bmatrix} a & b & c \\ b & d & m \\ c & m & n \end{Bmatrix},$$

where

$a = T$, $b = \sum_{t=1}^{T} t = \frac{1}{2}T^2 + O(T)$, $c = \sum_{t=1}^{T} x_t = \Delta_x T^2 + O(T)$,
$d = \sum_{t=1}^{T} t^2 = \frac{1}{3}T^3 + O(T^2)$, $m = \sum_{t=1}^{T} tx_t = \Psi_x T^3 + O(T^2)$, $n = \sum_{t=1}^{T} x_t^2 = \Gamma_x T^3 + O(T^2)$.
and

$$\widehat{\beta} = [\det(X'X)]^{-1} \begin{Bmatrix} dn - m^2 & cm - bn & bm - cd \\ cm - bn & an - c^2 & bc - am \\ bm - cd & bc - am & ad - b^2 \end{Bmatrix} \begin{Bmatrix} \sum_{t=1}^{T} y_t \\ \sum_{t=1}^{T} ty_t \\ \sum_{t=1}^{T} x_t y_t \end{Bmatrix}$$

with

$$\begin{aligned}\det(X'X) &= 2bcm + adn - c^2 d - am^2 - b^2 n \\ &= 2\sum_{t=1}^{T} t \sum_{t=1}^{T} x_t \sum_{t=1}^{T} tx_t + T\sum_{t=1}^{T} t^2 \sum_{t=1}^{T} x_t^2 - \left(\sum_{t=1}^{T} x_t\right)^2 \sum_{t=1}^{T} t^2 \\ &\quad - T\left(\sum_{t=1}^{T} tx_t\right)^2 - \left(\sum_{t=1}^{T} t\right)^2 \sum_{t=1}^{T} x_t^2 \equiv \det{}_x = O(T^7)\end{aligned}$$

Using results from lemma 2,
$T^{-7}\det(X'X) \xrightarrow{p} \Psi_x(\Delta_x - \Psi_x) + \frac{1}{12}\Gamma_x - \frac{1}{3}\Delta_x^2$
To prove parts a)-c) we derive expression for the numerators of $\widehat{\delta}_2, \widehat{\alpha}_2$, and $\widehat{\beta}_2$ as follows:

$$num\widehat{\delta}_2 = (bm - cd)\sum_{t=1}^{T} y_t + (bc - am)\sum_{t=1}^{T} ty_t + (ad - b^2)\sum_{t=1}^{T} x_t y_t = O(T^7)$$
$$num\widehat{\alpha}_2 = (dn - m^2)\sum_{t=1}^{T} y_t + (cm - bn)\sum_{t=1}^{T} ty_t + (bm - cd)\sum_{t=1}^{T} x_t y_t = O(T^8)$$
$$num\widehat{\beta}_2 = (cm - bn)\sum_{t=1}^{T} y_t + (an - c^2)\sum_{t=1}^{T} ty_t + (bc - am)\sum_{t=1}^{T} x_t y_t = O(T^7)$$

Using lemma 2 and the above definitions it is easy to show that

$$T^{-7} num\widehat{\delta}_2 \xrightarrow{p} (\tfrac{1}{2}\Psi_x - \tfrac{1}{3}\Delta_x)\Delta_y + (\tfrac{1}{2}\Delta_x - \Psi_x)\Psi_y + \tfrac{1}{12}\Gamma$$
$$T^{-8} num\widehat{\alpha}_2 \xrightarrow{p} \left(\tfrac{1}{3}\Gamma_x - \Psi_x^2\right)\Delta_y + \left(d_x\Psi_x - \tfrac{1}{2}\Delta_x\right)\Psi_y + \left(\tfrac{1}{2}\Psi_x - \tfrac{1}{3}\Delta_x\right)\Gamma$$
$$T^{-7} num\widehat{\beta}_2 \xrightarrow{p} \left(\Delta_x\Psi_x - \tfrac{1}{2}\Gamma_x\right)\Delta_y + \left(\Gamma_x - \Delta_x^2\right)\Psi_y + \left(\tfrac{1}{2}\Delta_x - \Psi_x\right)\Gamma$$

Combining these expression with that for $\det(X'X)$ gives the results in a)-c).

To prove d)-f) write the t-statistics as:

$$t_{\widehat{\delta}_2} = \widehat{\delta}_2\left[\widehat{\eta}_u^2 (X'X)_{33}^{-1}\right]^{-1/2},$$
$$t_{\widehat{\alpha}_2} = \widehat{\alpha}_2\left[\widehat{\eta}_u^2 (X'X)_{11}^{-1}\right]^{-1/2},$$
$$t_{\widehat{\beta}_2} = \widehat{\beta}_2\left[\widehat{\eta}_u^2 (X'X)_{22}^{-1}\right]^{-1/2},$$

and $(X'X)_{ii}^{-1}$, the i^{th} diagonal element of $(X'X)^{-1}$, as

$$(X'X)_{33}^{-1} = \frac{T^{-4}\left(\frac{1}{12}T^4\right) + o_p(1)}{T^{-7}\det_x + o_p(1)},$$
$$(X'X)_{11}^{-1} = \frac{T^{-6}\left(\frac{1}{3}\sum x_t^2 - \left(\sum_{t=1}^{T} tx_t\right)^2\right) + o_p(1)}{T^{-7}\det_x + o_p(1)},$$
$$(X'X)_{22}^{-1} = \frac{T^{-4}\left(\sum x_t^2 - \left(\sum_{t=1}^{T} x_t\right)^2\right) + o_p(1)}{T^{-7}\det_x + o_p(1)}.$$

Note also that

$$\widehat{\eta}_u^2 = T^{-1}\sum \widehat{u}_t^2 = T^{-1}\sum\left(y_t - \widehat{\alpha}_2 - \widehat{\beta}_2 t - \widehat{\delta}_2 x_t\right)^2$$

Using lemma 2,

$T^{-2}\widehat{\eta}_u^2 \xrightarrow{p} \left(\Gamma_y + \frac{1}{3}\beta_2^2 + \delta_2^2\Gamma_x - 2\beta_2\Psi_y - 2\delta_2\Gamma + 2\beta_2\delta_2\Psi_x\right) + (\beta_2 + 2\delta_2\Delta_x - 2\Delta_y)\alpha_2 + \alpha_2^2 \equiv \eta_u^2.$

Thus,

$\widehat{\delta}_2 \left[T^{-2}\widehat{\eta}_u^2 T^3 (X'X)^{-1}_{33}\right]^{-1/2} = T^{-1/2}\widehat{\delta}_2 \left[\widehat{\eta}_u^2 (X'X)^{-1}_{33}\right]^{-1/2} = O_p(1),$

$T^{-1}\widehat{\alpha}_2 \left[T^{-2}\widehat{\eta}_u^2 T (X'X)^{-1}_{11}\right]^{-1/2} = T^{-1/2}\widehat{\alpha}_2 \left[\widehat{\eta}_u^2 (X'X)^{-1}_{11}\right]^{-1/2} = O_p(1),$

$\widehat{\beta}_2 \left[T^{-2}\widehat{\eta}_u^2 T^3 (X'X)^{-1}_{22}\right]^{-1/2} = T^{-1/2}\widehat{\beta}_2 \left[\widehat{\eta}_u^2 (X'X)^{-1}_{22}\right]^{-1/2} = O_p(1).$

The limits of these expressions yield the formulae in d)-f).

5 References

Acemoglu, D. and A. Scott (1997), "Asymmetric Business Cycles: Theory and Time Series Evidence", *Journal of Monetary Economics*, 40(3), 501-533.

Bai, J. (1997a), "Estimating Multiple Breaks One at a Time", *Econometric Theory,* 13(3), 315-352.

_ _ _ _ _(1997b), "Estimation of a Change Point in Multiple Regression Models", *The Review of Economics and Statistics,* 79(4), 551-563.

Bai, J, and P. Perron (1998a), "Estimating and Testing Linear Models with Multiple Structural Changes", *Econometrica*, 66(1), 47-78.

Bai, J, and P. Perron (1998b), "Computation and Analysis of Multiple Structural Change Models", Mimeo.

Ben-David, D. and D.H. Papell (1995), "The Great Wars, the Great Crash, and Steady State Growth: Some New Evidence About an Old Stylized Fact", *Journal of Monetary Economics*, 36, 453-475.

——— (1998), "Slowdowns and Meltdowns: Postwar Growth Evidence from 74 Countries", *Review of Economics and Statistics*, 80, 561-571

Clemente, J., A. Montañéz and M. Reyes (1998), "Testing for a Unit Root in Variables with a Double Change in the Mean", *Economics Letters*, 59, 175-182.

Cooper, R. (1994), "Equilibrium Selection in Imperfectly Competitive Economies with Multiple Equilibria", *Economic Journal*, 104 (426), 1106-1122.

Duck, N.W. (1992), "UK Evidence on Breaking Trend Functions", *Oxford Economic Papers*, 44, 426-439.

Durlauf, S. (1993), "Nonergodic Economic Growth", *Review of Economic Studies*, 60, 349-366.

Entorf, H. (1997), "Random Walks with Drifts: Nonsense Regression and Spurious Fixed-Effect Estimation", *Journal of Econometrics*, 80, 287-296.

Granger, C.W.J. and P. Newbold (1974), "Spurious Regression in Econometrics", *Journal of Econometrics*, 2, 111-120.

Granger, C.W.J., N. Hyung, and Y. Jeon (1998), "Spurious Regressions with Stationary Series", Mimeo.

Hansen, B. (2001), "The New Econometrics of Structural Change: Dating Breaks in U.S. Labour Productivity", *Journal of Economic Perspectives*, 15(4), 117-128.

Kejak, M. (2003), "Stages of Growth in Economic Development", *Journal of Economic Dynamics and Control*, 27 (5), 771-800 .

Kim, T.-H., Y.-S. Lee and P. Newbold (2003), "Spurious Regressions with Processes Around Linear Trends of Drifts", Discussion Paper in Economics, No. 03-07, University of Nottingham.

Kim, T.-H., Y.-S. Lee and P. Newbold (2004), "Spurious Regressions with Stationary Processes Around Linear Trends", *Economics Letters*, 83, 257-262.

Lau, S.-H. P. (1997), "Using Stochastic Growth Models to Understand Unit Roots and Breaking Trends", *Journal of Economic Dynamics and Control*, 21, 1645-1667.

Lumsdaine, R.L. and D.H. Papell (1997), " Multiple Trend Breaks and the Unit Root Hypothesis", *The Review of Economics and Statistics*, 79, 212-218.

Marmol (1995), "Spurious Regressions for $I(d)$ Processes", *Journal of Time Series Analysis*, 16, 313-321.

Marmol (1996), "Correlation Theory for Spuriously Related Higher Order Integrated Processes", *Economics Letters*, 50, 163-73.

Marmol (1998), "Spurious Regression Theory with Nonstationary Fractionally Integrated Processes", *Journal of Econometrics*, 84, 233-50.

Nelson, C. and C. Plosser (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications", *Journal of Monetary Economics*, September 1982, 139-162.

Noriega, A. E. and E. de Alba (2001), "Stationarity and Structural Breaks: Evidence from Classical and Bayesian Approaches", *Economic Modelling*, 18, 503-524.

Perron, P. (1989), "The Great Crash, The Oil Price Shock, and The Unit Root Hypothesis", *Econometrica*, 57, 1361-1401.

—— (1992), "Trend, Unit Root and Structural Change: A Multi-Country Study with Historical Data", In *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 144-149.

—— (1997), "Further Evidence on Breaking Trend Functions in Macroeconomic Variables", *Journal of Econometrics*, 80, 355-385.

Perron, P. and X. Zhu (2002), "Structural Breaks with Deterministic and Stochastic Trends", Mimeo.

Phillips, P.C.B. (1986), "Understanding Spurious Regressions in Econometrics", *Journal of Econometrics*, 33, 311-340.

Raj, B. (1992), "International Evidence on Persistence in Output in the Presence of an Episodic Change", *Journal of Applied Econometrics*, 7, 281-293.

Startz, R. (1998), "Growth States and Shocks", *Journal of Economic Growth*, 3, 203-215.

Stock, J.H. (1994), "Unit Roots, Structural Breaks and Trends", *Handbook of Econometrics*, IV, 2740-2841.

Tsay, W.-J. and C.-F. Chung (1999), "The Spurious Regression of Fractionally Integrated Processes", Mimeo.

Zelhorst, D. and Haan, Jakob de (1995), "Testing for a Break in Output: New International Evidence", *Oxford Economic Papers*, 47, 357-362.

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 and Economic Statistics*, 10, 251-270.