Spurious regression under deterministic and stochastic trends

Antonio E. Noriega[1]
Daniel Ventosa-Santaulària[2]

Escuela de Economia
Universidad de Guanajuato

JEL Classification: *C22*
Keywords: Unit roots, Trend stationarity, Structural breaks, Spurious regression

Abstract

This paper analyses the asymptotic and finite sample implications of a mixed nonstationary behavior among the dependent and explanatory variables in a linear spurious regression model. We study the cases when the nonstationarity in the dependent variable is deterministic (stochastic), while the nonstationarity in the explanatory variable is stochastic (deterministic). In particular, we derive the asymptotic distribution of statistics in a spurious regression equation when one variable follows a difference stationary process (a random walk with and without drift), while the other is characterized by deterministic nonstationarity (a linear trend model with and without structural breaks in the trend function). We find that the divergence rate is sensitive to the assumed mixture of nonstationarity in the data generating process, and the phenomenon of spurious regression itself, contrary to previous findings, depends on the presence of a linear trend in the regression equation. Simulation experiments and real data confirm our asymptotic results.

[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.

[2] Escuela de Economia, Universidad de Guanajuato., e-mail: daniel@ventosa-santaularia.com

1 Introduction

It is well documented by now that the phenomenon of spurious regression is present under different forms of nonstationarity in the data generating process (DGP). It has been shown that when the variables y_t and x_t are nonstationary, independent of each other, ordinary least squares applied to the regression model

$$y_t = \alpha + \delta x_t + u_t$$

have the following implications: *1)* the estimator of δ ($\widehat{\delta}$) does not converge to its true value of zero, and *2)* the t-statistic for testing the null hypothesis $H_0 : \delta = 0$ ($t_{\widehat{\delta}}$) diverges, thus indicating the presence of an asymptotic spurious relationship between y_t and x_t.

The rate at which $t_{\widehat{\delta}}$ diverges depends on the type of nonstationarity present in the process generating y_t and x_t. In Phillips (1986), where a driftless random walk is assumed for both variables, the t-statistic is $O_p(T^{1/2})$. For the case of a random walk with drift, Entorf (1997) shows that $t_{\widehat{\delta}}$ diverges at rate T. 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: they find that, under a linear trend stationary assumption for both variables, the t-statistic is $O_p(T^{3/2})$. Extending KLN's results, Noriega and Ventosa-Santaulària (2005) (NVS hereafter), show that adding breaks in the DGP still generates the phenomenon of spurious regression, but at a reduced divergence rate; i.e. $t_{\widehat{\delta}}$ is $O(T^{1/2})$ under either single or multiple breaks in each variable. In all these works, the implicit assumption is that both variables share the same type of nonstationarity, either stochastic (Phillips, Entorf), or deterministic (KLN, NVS).[3]

This paper analyses the implications of a mixed nonstationary behavior among the dependent and explanatory variables. We develop the asymptotic theory for the cases when the nonstationarity in y_t is deterministic (stochastic), while the nonstationarity in x_t is stochastic (deterministic). In particular, we derive the limit distribution of statistics in the above regression equation when the DGP has mixed forms of nonstationarity: one variable follows a difference stationary process (a random walk with and without drift), while the other is characterized by deterministic nonstationarity (a linear trend model with and without structural breaks in the trend function).

Indeed, recent empirical studies show that macroeconomic time series differ in the nature of their nonstationary behavior. For instance, in a multi-country study with historical data on real per capita GDP, Perron (1992) finds that some series are well characterized as (broken) trend-stationary ($I(0)$ around a linear trend, and $I(0)$ around a linear trend with one structural break), while others

[3] Some related papers share this same feature: Marmol (1995, 1996, 1998), Cappuccio and Lubian (1997), Granger et. al. (1998) and Tsay and Chung (1999).

follow a unit root process. For the U.S. economy, Perron (1997) finds that real, nominal and real per capita GNP, employment, industrial production, nominal wage, and the money stock, stochastic nonstationarity (a unit root) can be rejected in favor of deterministic nonstationarity (a model with one structural break in the trend function); on the other hand, the consumer price index, velocity, interest rate and GNP deflator follow an $I(1)$ process. Similar results are reported in Lumsdaine and Papell (1997) (US data) and Mehl (2000) (Japanese data) when testing the null of a unit root against a deterministic nonstationary alternative with two structural breaks in the trend function. Noriega and de Alba (2001) also find mixed forms of nonstationarity ($I(1)$ and $I(0)$ with one break) using macro data under both classical and Bayesian approaches.

Looking ahead to the results reported below, we find that the divergence rate is sensitive to the assumed mixture of nonstationarity in the DGP, and the phenomenon of spurious regression itself depends on the presence of a linear trend in the regression equation, and on the presence of structural breaks in the DGP. For instance, if the dependent variable is $I(1)$ while the explanatory is $I(0)$ without breaks, a spurious relationship will show up both asymptotically and in finite samples, unless a linear trend is included in the regression model. We show this using asymptotic theory, simulation experiments, and real (macro) data. The rest of the paper is organized as follows. Section 2 reports asymptotic results for spurious regression under mixed forms of nonstationarity. Section 3 deals with experimental and empirical results. Last section concludes.

2 Asymptotics for spurious regressions

In a simple regression equation, stochastic nonstationarity may be a feature of the dependent variable, while deterministic nonstationarity of the explanatory variable, or viceversa. Since this can not be known a priori, we consider both possibilities: the dependent variable follows a trend-stationary (TS) process, while the explanatory variable follows a difference stationary (DS) one, and viceversa. Under difference stationarity, we study the cases of a unit root process with and without a drift, while under trend-stationarity we study the cases of a linear trend with and without multiple structural breaks in both level and slope of trend. These four types of nonstationary behavior for economic time series have been found to be empirically relevant in the literature, as argued above.

The following assumption summarizes the DGPs considered below for both the dependent and the explanatory variables.

Assumption. The DGPs for $\{y_t, x_t\}_{t=1}^{\infty}$ are as follows:

DGP A1
$$y_t = \mu_y + \beta_y t + u_{yt}$$
$$x_t = x_{t-1} + u_{xt} = x_0 + S_{xt}$$

DGP A2
$$y_t = \mu_y + \beta_y t + u_{yt}$$
$$x_t = \mu_x + x_{t-1} + u_{xt} = x_0 + \mu_x t + S_{xt}$$

$DGP\ A3$
$$y_t = \mu_y + \sum_{i=1}^{Ny} \theta_{iy} DU_{iyt} + \beta_y t + \sum_{i=1}^{My} \gamma_{iy} DT_{iyt} + u_{yt}$$
$$x_t = \mu_x + x_{t-1} + u_{xt} = x_0 + \mu_x t + S_{xt}$$

$DGP\ B1$
$$x_t = \mu_x + \beta_x t + u_{xt}$$
$$y_t = y_{t-1} + u_{yt} = y_0 + S_{yt}$$

$DGP\ B2$
$$x_t = \mu_x + \beta_x t + u_{xt}$$
$$y_t = \mu_y + y_{t-1} + u_{yt} = y_0 + \mu_y t + S_{yt}$$

$DGP\ B3$
$$x_t = \mu_x + \sum_{i=1}^{Nx} \theta_{ix} DU_{ixt} + \beta_x t + \sum_{i=1}^{Mx} \gamma_{ix} DT_{ixt} + u_{xt}$$
$$y_t = \mu_y + y_{t-1} + u_{yt} = y_0 + \mu_y t + S_{yt}$$

where, for $z = x, y$, $S_{zt} = \sum_{i=1}^{t} u_{zi}$, z_0 is an initial condition, $u_{zt} = \phi_z u_{zt-1} + \varepsilon_{zt}$, $|\phi_z| < 1$, ε_{zt} are $iid(0, \sigma_z^2)$ independent of each other, and DU_{izt}, DT_{izt} are dummy variables allowing changes in the trend's level and slope respectively, that is, $DU_{izt} = \mathbf{1}(t > T_{b_{iz}})$ and $DT_{izt} = (t - T_{b_{iz}})\mathbf{1}(t > T_{b_{iz}})$, where $\mathbf{1}(\cdot)$ is the indicator function, and $T_{b_{iz}}$ is the i^{th} unknown date of the break in z. We denote the break fraction as $\lambda_z = (T_{b_z}/T) \in (0, 1)$.

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). As will be shown below: *1)* the relevant limiting expressions do not depend on initial conditions (z_0), and *2)* they depend on σ_z^2 only under DGPs $i2, i = A, B$.

We begin by considering the following spurious ordinary least squares regression model:

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

used as a vehicle for testing the null hypothesis $H_0 : \delta_1 = 0$.

The following theorems collect the asymptotic behavior of the estimated parameters and associated $t-$statistics in model (1) when the dependent and explanatory variables have mixed forms of nonstationarity, according to the DGPs in the Assumption. Proofs and definitions for some of the objects in the next theorems are collected in the Appendix.

Theorem 1A. *Let y_t and x_t be generated according to the DGP Ai, $i = 1, 2, 3$, and denote the corresponding OLS estimates of* α_1 and δ_1 in (1) *by* $\widehat{\alpha}_{1A(i)}$ and $\widehat{\delta}_{1A(i)}$. *Then, as $T \to \infty$:*

Case (1): DGP A1.

a) $T^{-1}\widehat{\alpha}_{1A(1)} \xrightarrow{d} \frac{1}{2}\beta_y N_1 D_1^{-1}$

b) $T^{-1/2}\widehat{\delta}_{1A(1)} \xrightarrow{d} \frac{1}{2}\beta_y N_{2x} (\sigma_x D_1)^{-1}$

c) $T^{-1/2} t_{\widehat{\alpha}_{1A(1)}} \xrightarrow{d} N_1 \left(\frac{1}{3} D_{2x} \int W_x^2\right)^{-1/2}$

d) $T^{-1/2} t_{\widehat{\delta}_{1A(1)}} \xrightarrow{d} \sqrt{3} N_{2x} D_{2x}^{-1/2}$

Case (2): DGP A2

a) $T^{-1/2} \widehat{\alpha}_{1A(2)} \xrightarrow{d} \sigma_x \frac{\beta_y}{\mu_x} N_{3x}$

b) $\widehat{\delta}_{1A(2)} \xrightarrow{p} \frac{\beta_y}{\mu_x}$

c) $T^{-1/2} t_{\widehat{\alpha}_{1A(2)}} \xrightarrow{d} \frac{1}{2} N_{3x} D_{2x}^{-1/2}$

d) $T^{-1} t_{\widehat{\delta}_{1A(2)}} \xrightarrow{d} \frac{\mu_x}{\sigma_x} (12 D_{2x})^{-1/2}$

Case (3): DGP A3

a) $T^{-1} \widehat{\alpha}_{1A(3)} \xrightarrow{p} 24 V_{4y}$

b) $\widehat{\delta}_{1A(3)} \xrightarrow{p} (\beta_y + 24 V_{5y}) \mu_x^{-1}$

c) $T^{-1/2} t_{\widehat{\alpha}_{1A(3)}} \xrightarrow{p} 12 V_{4y} G_{6y}^{-1/2}$

d) $T^{-1/2} t_{\widehat{\delta}_{1A(3)}} \xrightarrow{p} (\beta_y + 24 V_{5y})(12 G_{6y})^{-1/2}$

THEOREM 1B. *Let y_t and x_t be generated according to the DGP Bi, $i = 1, 2, 3$, and denote the corresponding OLS estimates of α_1* and δ_1 in (1) *by $\widehat{\alpha}_{1B(i)}$* and $\widehat{\delta}_{1B(i)}$. *Then, as $T \to \infty$:*

Case (1): DGP B1.

a) $T^{-1} \widehat{\alpha}_{1B(1)} \xrightarrow{d} -\sigma_y N_{3y}$

b) $T^{1/2} \widehat{\delta}_{1B(1)} \xrightarrow{d} \frac{\sigma_y}{\beta_x} 6 N_{2y}$

c) $T^{-1/2} t_{\widehat{\alpha}_{1B(1)}} \xrightarrow{d} -N_{3y} (4 D_{2y})^{-1/2}$

d) $T^{-1/2} t_{\widehat{\delta}_{1B(1)}} \xrightarrow{d} \sqrt{3} N_{2y} D_{2y}^{-1/2}$

Case (2): DGP B2.

a) $T^{-1/2} \widehat{\alpha}_{1B(2)} \xrightarrow{d} -\sigma_y N_{3y}$

b) $\widehat{\delta}_{1B(2)} \xrightarrow{p} \frac{\mu_y}{\beta_x}$

c) $T^{-1/2} t_{\widehat{\alpha}_{1B(2)}} \xrightarrow{d} -\frac{1}{2} N_{3y} D_{2y}^{-1/2}$

d) $T^{-1} t_{\widehat{\delta}_{1B(2)}} \xrightarrow{d} \frac{\mu_y}{\sigma_y} (12 D_{2y})^{-1/2}$

Case (3): DGP B3.

a) $T^{-1}\widehat{\alpha}_{1B(3)} \xrightarrow{p} \mu_y \left(V_{2x} - \beta_x V_{4x}\right) \left[2\beta_x \left(\frac{1}{48}\beta_x + V_{5x}\right) - 2V_{3x}\right]^{-1}$

b) $\widehat{\delta}_{1B(3)} \xrightarrow{p} \mu_y \left(\frac{1}{24}\beta_x + V_{5x}\right) \left[2\beta_x \left(\frac{1}{48}\beta_x + V_{5x}\right) - 2V_{3x}\right]^{-1}$

c) $T^{-1/2}t_{\widehat{\alpha}_{1B(3)}} \xrightarrow{p} \left(V_{2x} - \beta_x V_{4x}\right) \left[\left(\frac{1}{12}V_{1x} - V_{4x}^2\right) \left(\beta_x^2 + \beta_x G_{1x} + G_{3x} + 6G_{4x}\right)\right]^{-1/2}$

d) $T^{-1/2}t_{\widehat{\delta}_{1B(3)}} \xrightarrow{p} \left(\frac{1}{24}\beta_x + V_{5x}\right) \left[3\left(\frac{1}{12}V_{1x} - V_{4x}^2\right)\right]^{-1/2}$

Parts *b*) of the theorems show that, for $DGP\ i1, i = A, B$, the (normalized) estimated spurious parameter either goes to infinity ($DGP\ A1$), or collapses to zero ($DGP\ B1$). For the rest, it converges to well defined limits. Parts *d*) show that the spurious regression phenomenon is present under all DGPs, but occurs at a faster rate under $DGP\ i2, i = A, B$, which correspond to the models studied by Nelson and Plosser (1982). More over, the limiting distribution of the t-statistic for the hypothesis $H_0 : \delta = 0$ is symmetrical across DGPs Aj, and Bj, $j = 1, 2$.

We now consider the case of a (spurious) OLS 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

$$y_t = \widehat{\alpha}_2 + \widehat{\beta}_2 t + \widehat{\delta}_2 x_t + \widehat{u}_t \tag{2}$$

where y_t and x_t are generated from the DGPs in the Assumption. The next theorems present the relevant asymptotic results.

THEOREM 2A. *Let y_t and x_t be generated according to the DGP Ai, $i = 1, 2, 3$, and denote the corresponding OLS estimates of* α_2, β_2, and δ_2 *in (2) by* $\widehat{\alpha}_{2A(i)}$, $\widehat{\beta}_{2A(i)}$, and $\widehat{\delta}_{2A(i)}$. *Then, as $T \to \infty$:*

Case (1): DGP A1.

a) $\widehat{\alpha}_{2A(1)} \xrightarrow{p} \mu_y$

b) $\widehat{\beta}_{2A(1)} \xrightarrow{p} \beta_y$

c) $T\widehat{\delta}_{2A(1)} \xrightarrow{d} \frac{\sigma_y}{\sigma_x} N_{4x} D_{2x}^{-1}$

d) $T^{-1/2}t_{\widehat{\alpha}_{2A(1)}} \xrightarrow{d} \frac{\mu_y}{\sigma_y} \left(\frac{D_{2x}}{D_3}\right)^{1/2}$

e) $T^{-3/2}t_{\widehat{\beta}_{2A(1)}} \xrightarrow{d} \frac{\beta_y}{\sigma_y} \left(\frac{D_{2x}}{12D_1}\right)^{1/2}$

f) $t_{\widehat{\delta}_{2A(1)}} \xrightarrow{d} N_{4x} D_{2x}^{-1/2}$

Case (2): DGP A2.

a) $\widehat{\alpha}_{2A(2)} \xrightarrow{p} \mu_y$

b) $\widehat{\beta}_{2A(2)} \xrightarrow{p} \beta_y$

c) $T\widehat{\delta}_{2A(2)} \xrightarrow{d} \frac{\sigma_y}{\sigma_x} N_{4x} D_{2x}^{-1}$

d) $T^{-1/2} t_{\widehat{\alpha}_{2A(2)}} \xrightarrow{d} \frac{\mu_y}{\sigma_y} \left(\frac{D_{2x}}{D_3} \right)^{1/2}$

e) $T^{-1} t_{\widehat{\beta}_{2A(2)}} \xrightarrow{d} \frac{\beta_y}{\mu_x} \frac{\sigma_x}{\sigma_y} D_{2x}^{1/2}$

f) $t_{\widehat{\delta}_{2A(2)}} \xrightarrow{d} N_{4x} D_{2x}^{-1/2}$

Case (3): DGP A3.

a) $T^{-1}\widehat{\alpha}_{2A(3)} \xrightarrow{d} \left[\frac{1}{2} G_{2y} D_3 - G_{1y} N_1 + G_{5y} N_{3x}\right] D_{2x}^{-1}$

b) $T^{-1/2}\widehat{\beta}_{2A(3)} \xrightarrow{d} \mu_x \left[24 \left(V_{4y} \int W_y + V_{5y} \int r W_y\right) - G_{5y}\right] (\sigma_x D_{2x})^{-1}$

c) $T^{-1/2}\widehat{\delta}_{2A(3)} \xrightarrow{d} \left[G_{5y} - 24 \left(V_{4y} \int W_y + V_{5y} \int r W_y\right)\right] (\sigma_x D_{2x})^{-1}$

d) $T^{-1/2} t_{\widehat{\alpha}_{2A(3)}} \xrightarrow{d} \left(\frac{1}{2} G_{2y} D_3 - G_{1y} N_1 + G_{5y} N_{3x}\right) \left[D_3 \sigma_{u(A3)}\right]^{-1/2}$

e) $T^{-1/2} t_{\widehat{\beta}_{2A(3)}} \xrightarrow{d} \left[24 \left(V_{4y} \int W_y + V_{5y} \int r W_y\right) - G_{5y}\right] \left[\sigma_{u(A3)}\right]^{-1/2}$

f) $T^{-1/2} t_{\widehat{\delta}_{2A(3)}} \xrightarrow{d} \left[G_{5y} - 24 \left(V_{4y} \int W_y + V_{5y} \int r W_y\right)\right] \left[\sigma_{u(A3)}\right]^{-1/2}$

THEOREM 2B. *Let y_t and x_t be generated according to the DGP Bi, $i = 1, 2, 3$, and denote the corresponding OLS estimates of* α_2, β_2, and δ_2 *in (2) by* $\widehat{\alpha}_{2B(i)}$, $\widehat{\beta}_{2B(i)}$, and $\widehat{\delta}_{2B(i)}$, *Then, as $T \to \infty$:*

Case (1): DGP B1.

a) $T^{-1/2}\widehat{\alpha}_{2B(1)} \xrightarrow{d} -\sigma_y N_{3y}$

b) $\widehat{\beta}_{2B(1)} \xrightarrow{d} -\frac{\sigma_y}{\sigma_x} N_{4y} \beta_x$

c) $\widehat{\delta}_{2B(1)} \xrightarrow{d} \frac{\sigma_y}{\sigma_x} N_{4y}$

d) $T^{-1/2} t_{\widehat{\alpha}_{2B(1)}} \xrightarrow{d} -\sigma_x N_{3y} \left[\left(\mu_x^2 + 4\sigma_x^2\right) D_{2y}\right]^{-1/2}$

e) $t_{\widehat{\beta}_{2B(1)}} \xrightarrow{d} -N_{4y} D_{2y}^{-1/2}$

f) $t_{\widehat{\delta}_{2B(1)}} \xrightarrow{d} N_{4y} D_{2y}^{-1/2}$

Case (2): DGP B2.

a) $T^{-1/2}\widehat{\alpha}_{2B(2)} \xrightarrow{d} -\sigma_y N_{3y}$

b) $\widehat{\beta}_{2B(2)} \xrightarrow{d} \mu_y - \frac{\sigma_y}{\sigma_x} N_{4y}\beta_x$

c) $\widehat{\delta}_{2B(2)} \xrightarrow{d} \frac{\sigma_y}{\sigma_x} N_{4y}$

d) $T^{-1/2} t_{\widehat{\alpha}_{2B(2)}} \xrightarrow{d} -\sigma_x N_{3y} \left[D_{2y}\left(\mu_x^2 + 4\sigma_x^2\right)\right]^{-1/2}$

e) $t_{\widehat{\beta}_{2B(2)}} \xrightarrow{d} \left[\frac{\sigma_x}{\sigma_y}\frac{\mu_y}{\beta_x} - N_{4y}\right] D_{2y}^{-1/2}$

f) $t_{\widehat{\delta}_{2B(2)}} \xrightarrow{d} N_{4y} D_{2y}^{-1/2}$

Case (3): DGP B3.

a) $T^{-1/2}\widehat{\alpha}_{2B(3)} \xrightarrow{d} 24\sigma_y \left(V_{2x}\int rW_y - 2V_{1x}\int W_y + V_{4x}G_{5x}\right) G_{6x}^{-1}$

b) $\widehat{\beta}_{2B(3)} \xrightarrow{p} \mu_y$

c) $T^{1/2}\widehat{\delta}_{2B(3)} \xrightarrow{d} \sigma_y \left[24\left(V_{4x}\int W_y + V_{5x}\int rW_y\right) - G_{5x}\right] G_{6x}^{-1}$

d) $T^{-1/2} t_{\widehat{\alpha}_{2B(3)}} \xrightarrow{d} \left(V_{2x}\int rW_y - 2V_{1x}\int W_y + V_{4x}G_{5x}\right)\left(2\sqrt{\sigma_{u(B3)}V_{1x}}\right)^{-1}$

e) $T^{-1} t_{\widehat{\beta}_{2B(3)}} \xrightarrow{d} \mu_y G_{6x}\left\{24\sigma_y\left(\sigma_{u(B3)}\left[4\left(V_{3x} - \beta_x V_{5x}\right) - \frac{1}{12}\beta_x^2\right]\right)^{1/2}\right\}^{-1}$

f) $T^{-1/2} t_{\widehat{\delta}_{2B(3)}} \xrightarrow{d} \left[24\left(V_{4x}\int W_y + V_{5x}\int rW_y\right) - G_{5x}\right]\left(48\sigma_{u(B3)}\right)^{-1}$

Parts f) of the theorems show that the spurious regression parameter t-statistic only diverges to infinity under structural breaks; otherwise, it has a well defined limit. Hence, a spurious relationship will be present only under DGP $i3$, $i = A, B$. Note that, as opposed to the case of no trend in the regression, the spurious regression coefficient (and its t-statistic) converges to the same distribution across DGPs ij, $i = A, B$; $j = 1, 2$. Table A1 in the Appendix presents a summary of results concerning orders in probability of relevant statistics.

3 Experimental and empirical results

We computed rejection rates of the t-statistic for testing the null hypotheses $H_0 : \delta_j = 0, j = 1, 2$, in equations (1) and (2), respectively, using a 1.96 critical value (5% level) for a standard normal distribution.

In order to asses the usefulness of our limit theory in finite samples, rejection rates were based on both the asymptotic formulae and simulated data, for samples of size $T = 50, 100, 250, 500, 1000, 10000$, under each one of the DGPs in the Assumption. The value of the parameters in the DGPs are as follows[4]: $\sigma_z = 1$, $\phi_z = 0$ (but see below), $\mu_z = 0$, or 0.5, depending on the DGP containing a drift or not, $\beta_z = 0.03$ if the DGP contains a nonzero slope, and, whenever breaks are present, $M_z = 2$ and $\gamma_i = 0.03, i = 1, ..., M_z$, for $z = x, y$. Breaks occur at 10% and 30% of total data length. The number of replications is $10,000$.

The simulation experiments reveal a remarkable agreement in rejection rates between analytical results and simulated ones, as shown in Table 1. The only discrepancy occurs for the multibreaks model with trend (DGPs $i3$, $i = A, B$), for $T < 500$. Results also indicate that the rejection rates for the cases of the model without trend can be very high, even for values of the sample size as small as 50. Hence, in these cases, the phenomenon of spurious regression is likely to be present also in small (empirically relevant) samples.

Table 1

Rejection Rates for $t_{\widehat{\delta}}$; DGPs Aj, $j = 1, 2, 3$.

Sample size	DGP $A1$				DGP $A2$				DGP $A3$			
	no trend		trend		no trend		trend		no trend		trend	
	theo	sim	theo	sim	theo	sim	theo	sim	theo	sim	theo	sim
50	0.83	0.82	0.05	0.06	1.00	0.82	0.06	0.06	1.00	0.99	0.69	0.06
100	0.88	0.87	0.05	0.05	1.00	1.00	0.05	0.05	1.00	1.00	0.77	0.15
250	0.92	0.93	0.04	0.06	1.00	1.00	0.05	0.04	1.00	1.00	0.86	0.67
500	0.95	0.94	0.05	0.04	1.00	1.00	0.05	0.05	1.00	1.00	0.89	0.86
1,000	0.96	0.96	0.05	0.04	1.00	1.00	0.05	0.05	1.00	1.00	0.93	0.92
10,000	0.99	0.99	0.05	0.05	1.00	1.00	0.05	0.05	1.00	1.00	0.98	0.98

Rejection Rates for $t_{\widehat{\delta}}$; DGPs Bj, $j = 1, 2, 3$.

Sample size	DGP $B1$				DGP $B2$				DGP $B3$			
	no trend		trend		no trend		trend		no trend		trend	
	theo	sim	theo	sim	theo	sim	theo	sim	theo	sim	theo	sim
50	0.83	0.45	0.05	0.06	1.00	0.82	0.04	0.05	1.00	0.99	0.69	0.06
100	0.87	0.81	0.05	0.05	1.00	1.00	0.05	0.05	1.00	1.00	0.78	0.16
250	0.92	0.91	0.04	0.05	1.00	1.00	0.05	0.05	1.00	1.00	0.86	0.67
500	0.95	0.94	0.05	0.05	1.00	1.00	0.05	0.05	1.00	1.00	0.90	0.85
1,000	0.96	0.96	0.05	0.05	1.00	1.00	0.05	0.05	1.00	1.00	0.92	0.92
10,000	0.99	0.99	0.05	0.05	1.00	1.00	0.05	0.05	1.00	1.00	0.98	0.97

Note: theo and sim stand for theoretical (based on the corresponding theorem), and simulated.

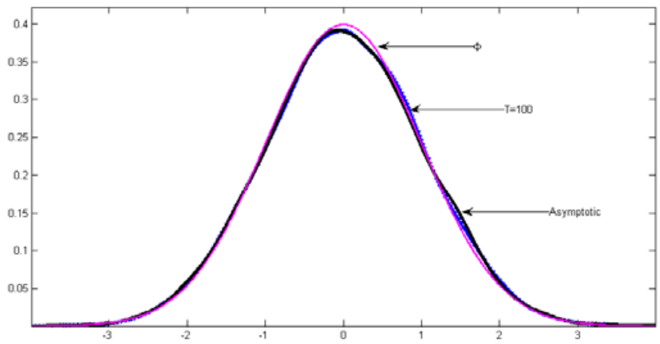
When including a linear trend in the regression model, however, rejection rates for DGPs ij, $i = A, B; j = 1, 2$ do fluctuate around the nominal level, for any sample size. Hence, assuming no breaks in the DGP, the inclusion of a linear

[4] We experimented with different values of the parameters, location and number of breaks, and obtained very similar results.

trend eliminates the spurious regression phenomenon, not only asymptotically, but also in finite samples.

As can be seen from the theorems, the limiting distribution of $t_{\widehat{\delta}}$ depends on σ_z^2 under DGPs $i2, i = A, B$. To asses the effect of autocorrelation, a simulation experiment was performed to compute rejection rates for $DGP\ i2, i = A, B$ using $\phi_z = 0.1, 0.5, 0.9$. Results revealed that rejection rates are greater that 60% even in the case of a sample as small as $T = 50$, and $\phi_z = 0.9$.

As the analytical results show, the presence of a linear trend in the regression equation ensures a well defined limit for the distribution of $t_{\widehat{\delta}_2}$, under DGPs ij, $i = A, B$; $j = 1, 2$. Figure 1 shows that this limit closely resembles a standard normal distribution. The Monte Carlo experiments used to generate the depicted densities were based on *1)* A non-parametric estimation using 10,000 simulated data samples of 100 observations (labeled $T = 100$), under $DGP\ B2$, and regression equation (2), with parameter values: $\sigma_z = 1$, $\phi_z = 0$, $\mu_z = 0.5$, $\beta_y = 0.03$, for $z = x, y$; and *2)* A non-parametric estimation using 30,000 replications and the asymptotic distributions of $t_{\widehat{\delta}_2}$ (labeled *Asymptotic*), for DGP $B2$, as stated in Theorem $2B$. Additionally, a standard normal distribution is included for comparison purposes.



Non-parametric estimation of $t_{\widehat{\delta}_2}$, for $DGP\ B2$.

To assess the empirical relevance of the above results, we utilize long, low frequency data of (log) real per capita GDP series from 1870 to 1986 for 10 countries: Australia, Canada, Denmark, France, Germany, Italy, Norway, Sweden, the United Kingdom and the United States[5]. According to Perron (1992), a unit root in the autoregressive representation can be rejected at the 5% level for Australia, Canada, Denmark, France, Germany and the U.K., when allowance is made for a single structural break in the trend function. For Italy, Norway and Sweden, the unit root can not be rejected; while for the U.S. there is neither a unit root nor a structural break. Hence, for the U.S., real per capita output is

[5] This data set was kindly provided by Pierre Perron, and is the same as used by Kormendi and Meguire (1990), Perron (1992) and Perron and Zhu (2002).

$I(0)$, for Italy, Norway and Sweden, $I(1)$, and for the remaining countries $I(0)$ with one break.

As predicted in theorems $1A$ and $1B$, when regression equation (1) is estimated, a spurious relationship will prevail asymptotically, and, as our simulation experiment shows, this relationship will hold also in small samples. On the other hand, when regression equation (2) is run under DGPs ij, $i = A, B; j = 1, 2$, no spurious regression will be present, as shown by theorems $2A$ and $2B$, and the simulations. Table 2 present the results of estimating (1) and (2) by OLS, to test the null hypothesis $H_0 : \delta = 0$, using pairs of variables according to DGP $B2$ (note that the same results are obtained under DGP $A2$).

Table 2

Significance of $t_{\widehat{\delta}}$ under DGP $B2$

y_t	x_t	Only Constant		Constant and Trend	
$I(1)$	$I(0)$	$\widehat{\delta}$	$t-stat$	$\widehat{\delta}$	$t-stat$
Italy	USA	0.948	22.70***	0.004	0.018
Norway	USA	1.126	34.04***	-0.035	-0.215
Sweden	USA	1.270	53.39***	0.023	0.306

Note: *** indicates rejection at the 1% level

As expected, inclusion of a linear trend does eliminate the spurious relationship between U.S output and the other three GDP variables. Table 3 gives empirical support of findings reported in theorem $2B$ under DGP $B3$.

Table 3

Significance of $t_{\widehat{\delta}}$ under DGPs $B3$

y_t	x_t	Only Constant		Constant and Trend	
$I(1)$	$I(0)+break$	$\widehat{\delta}$	$t-stat$	$\widehat{\delta}$	$t-stat$
Italy	Australia	1.585	36.19***	1.112	14.29***
	Canada	0.912	29.97***	0.859	6.76***
	Denmark	0.955	39.26***	1.843	18.03***
	France	1.137	59.19***	1.000	22.66***
	Germany	0.894	42.80***	0.849	13.93***
	UK	1.463	40.06***	1.750	12.96***
Norway	Australia	1.772	33.33***	0.821	16.65***
	Canada	1.056	41.58***	0.532	5.70***
	Denmark	1.099	64.39***	1.132	12.16***
	France	1.256	39.31***	0.632	14.86***
	Germany	1.013	48.23***	0.585	13.24***
	UK	1.681	63.24***	1.259	13.75***
Sweden	Australia	1.862	23.73***	0.328	10.68***
	Canada	1.168	47.57***	0.218	4.74***
	Denmark	1.215	89.66***	0.569	13.90***
	France	1.333	27.79***	0.239	9.05***
	Germany	1.093	36.47***	0.248	10.27***
	UK	1.828	48.74***	0.484	8.85***

In these regressions, a spurious relationship is present whether a linear trend is included or not, since one of the variables (the explanatory one in this case) underwent a structural break in the trend function.

4 Conclusions

This paper has presented an analysis of the spurious regression phenomenon when there is a mix of deterministic and stochastic nonstationarity among the dependent and the explanatory variables in a linear regression model. It has shown that: *1)* the asymptotic distribution of the t-statistic for testing a spurious relationship is sensitive to the assumed mixture of nonstationarity, and *2)* the phenomenon of spurious regression itself depends on the presence of a linear trend in the regression equation and on the presence of structural breaks in the DGP.

Thus, if it is believed that there might be a form of mixed nonstationarity (DS and TS with no breaks) among the dependent and explanatory variables in a regression equation, to avoid the phenomenon of a spurious relationship, a linear trend should be included in such regression model; otherwise, a spurious relationship will be present under any mix of DGPs. However, when structural breaks are a feature of the data (either in the dependent or the explanatory variable), the presence of a spurious relationship is unambiguous whether the regression model includes a linear trend or not.

5 Appendix

Proof of Theorems

The proofs were assisted by the software *Mathematica 4.1.* The corresponding codes are available from the authors upon request. Below, we describe the steps involved in the computerized calculations.

Write either regression model $y_t = \alpha_1 + \delta_1 x_t + u_t$ or $y_t = \alpha_2 + \beta_2 t + \delta_2 x_t + u_t$ in matrix form: $y = X\beta + u$. The vector of OLS estimators, $\widehat{\beta} = (X'X)^{-1}X'y$, is a function of the following objects:

For $DGP\ Ai, i = 1, 2$.

$\sum_{t=1}^{T} y_t = \frac{1}{2}\beta_y T^2 + (\mu_y + \frac{1}{2}\beta_y)T + \Sigma_{uy}T^{1/2}$

$\sum_{t=1}^{T} y_t^2 = \frac{1}{3}\beta_y^2 T^3 + (\frac{1}{2}\beta_y^2 + \mu_y\beta_y)T^2 + 2\beta_y\Sigma_{tuy}T^{3/2} + \left(\mu_y^2 + \frac{1}{6}\beta_y^2 + \Sigma_{u2y} + \mu_y\beta_y\right)T$
$+2\mu_y\Sigma_{uy}T^{1/2}$

$\sum_{t=1}^{T} ty_t = \frac{1}{3}\beta_y T^3 + \frac{1}{2}(\mu_y + \beta_y)T^2 + \Sigma_{tuy}T^{3/2} + \left(\frac{1}{2}\mu_y + \frac{1}{6}\beta_y\right)T$

For $DGP\ A1$.

$\sum_{t=1}^{T} x_t = \Sigma_{sx}T^{3/2} + x_0 T$

$\sum_{t=1}^{T} x_t^2 = \Sigma_{s2x}T^2 + 2x_0\Sigma_{sx}T^{3/2} + x_0^2 T$

$\sum_{t=1}^{T} tx_t = \Sigma_{sx}T^{5/2} + \frac{1}{2}x_0 T^2 + \frac{1}{2}x_0 T$

$\sum_{t=1}^{T} y_t x_t = \beta_y\Sigma_{tsx}T^{5/2} + \frac{1}{2}x_0\beta_y T^2 + \mu_y\Sigma_{sx}T^{3/2} + \left(x_0\mu_x + \frac{1}{2}x_0\beta_y + \Sigma_{sxuy}\right)T$
$+x_0\Sigma_{uy}T^{1/2}$

For $DGP\ Ai, i = 2, 3$.

$\sum_{t=1}^{T} x_t = \frac{1}{2}\mu_x T^2 + \Sigma_{sx}T^{3/2} + (x_0 + \frac{1}{2}\mu_x)T$

$\sum_{t=1}^{T} x_t^2 = \frac{1}{3}\mu_x^2 T^3 + 2\mu_x\Sigma_{tsx}T^{5/2} + (\frac{1}{2}\mu_x^2 + \Sigma_{s2x} + x_0\mu_x)T^2 + 2x_0\Sigma_{sx}T^{3/2} + (x_0^2$
$+\frac{1}{6}\mu_x^2 + x_0\mu_x)T$

$\sum_{t=1}^{T} tx_t = \Sigma_{tsx}T^{5/2} + \frac{1}{2}x_0 T^2 + \frac{1}{2}x_0 T$

For $DGP\ A2$.

$\sum_{t=1}^{T} y_t x_t = \frac{1}{3}\beta_y\mu_x T^3 + \beta_y\Sigma_{tsx}T^{5/2} + \left(\frac{1}{2}x_0\beta_y + \frac{1}{2}\mu_x\mu_y + \frac{1}{2}\mu_x\beta_y\right)T^2 +$
$\left(\mu_x\Sigma_{tuy} + \mu_y\Sigma_{sx}\right)T^{3/2} + \left(x_0\mu_y + \frac{1}{2}x_0\beta_y + \frac{1}{2}\mu_x\mu_y + \frac{1}{6}\mu_x\beta_y + \Sigma_{sxuy}\right)T + x_0\Sigma_{uy}T^{1/2}$

For $DGP\ A3$.

$\sum_{t=1}^{T} y_t = \frac{1}{2}\left(\beta_y + G_{2y}\right)T^2 + \left[\frac{1}{2}\beta_y + \sum_{i=1}^{M}\theta_i(1-\lambda_i) + \frac{1}{2}\sum_{i=1}^{M}\gamma_i(1-\lambda_i) + \mu_y\right]T$
$+\Sigma_{uy}T^{1/2}$

$\sum_{t=1}^{T} y_t^2 = \left(\frac{1}{3}\beta_y^2 + \frac{1}{3}G_{3y} + G_{4y}\right)T^3 + O(T^2)$

$\sum_{t=1}^{T} ty_t = \left(\frac{1}{3}\beta_y + \frac{1}{6}G_{1y}\right)T^3 + O(T^2)$

$\sum_{t=1}^{T} y_t x_t = \mu_x\left(\frac{1}{3}\beta_y + \frac{1}{6}G_{1y}\right)T^3 + \left(\beta_y \Sigma_{tsx} + \sum_{i=1}^{M_y}\gamma_i \Sigma_{ts1sx}\right)T^{5/2} + O(T^2)$

with

$\Sigma_{uy} = T^{-1/2}\sum_{t=1}^{T} u_{yt}$
$\Sigma_{tuy} = T^{-3/2}\sum_{t=1}^{T} tu_{yt}$
$\Sigma_{u2y} = T^{-1}\sum_{t=1}^{T} u_{yt}^2$
$\Sigma_{sx} = T^{-3/2}\sum_{t=1}^{T} S_{xt}$
$\Sigma_{s2x} = T^{-2}\sum_{t=1}^{T} S_{xt}^2$
$\Sigma_{tsx} = T^{-5/2}\sum_{t=1}^{T} tS_{xt}$
$\Sigma_{sxuy} = T^{-1}\sum_{t=1}^{T} S_{xt}u_{yt}$
$\Sigma_{ts1sx} = T^{-5/2}\left(\sum_{t=T_b+1}^{T} tS_{xt} - \lambda_i \sum_{t=T_b+1}^{T} S_{xt}\right)$

(For $DGP\ Bi, i = 1, 2, 3$, simply invert the roles of x and y). Using these expressions, *Mathematica* computes the limiting distribution of the parameter vector by factoring out $(X'X)^{-1}X'y$ in powers of the sample size. In this way, the orders in probability can be determined, and the limiting expression obtained, by retaining only the asymptotically relevant terms, upon a suitable normalization. The expressions presented in the theorems result from the factorization of these limits. The proof of theorems 2A and 2B follows the same steps.

Definitions. We make notational economies by writing the various stochastic processes without the argument. Integrals are understood to be taken over the interval $[0, 1]$, and with respect to Lebesgue measure, unless otherwise indicated. Thus, we use, for instance, W_z, $\int W_z$, and $\int rW_z$ in place of $W_z(r)$, $\int_0^1 W_z(r)dr$, and $\int_0^1 rW_z(r)dr$, where $W_z(r)$ is the standard Wiener process on $r \in [0, 1]$.

For $z = x, y$,

$N_1 = \int W_x^2 - 2\int rW_x \int W_x$
$N_{2z} = 2\int rW_z - \int W_z$
$N_{3z} = 6\int rW_z - 4\int W_z$
$N_{4x} = \int W_x dW_y + N_{3x}W_y(1) - 6N_{2x}\left(W_y(1) - \int W_y\right)$
$N_{4y} = \int W_y dW_x + N_{3y}W_x(1) - 6N_{2y}\left(W_x(1) - \int W_x\right)$
$D_1 = \int W_x^2 - \left(\int W_x\right)^2$
$D_{2z} = \int W_z^2 - 12\int rW_z\left(\int rW_z - \int W_z\right) - 4\left(\int W_z\right)^2$
$D_3 = 4\int W_x^2 - 12\left(\int rW_x\right)^2$
$G_{1z} = \sum_{i=1}^{M_z}\gamma_{iz}(1-\lambda_{iz})^2(\lambda_{iz}+2)$
$G_{2z} = \sum_{i=1}^{M_z}\gamma_{iz}(1-\lambda_{iz})^2$
$G_{3z} = \sum_{i=1}^{M_z}\gamma_{iz}^2(1-\lambda_{iz})^3$

$G_{4z} = \sum_{i=1}^{M_z-1} \sum_{j=i+1}^{M_z} \gamma_{iz}\gamma_{jz} \left[\frac{2}{3}(1-\lambda_{u(i,j)})^3 + \lambda_{d(i,j)}(1-\lambda_{u(i,j)})^2\right]$
$G_{5z} = \sum_{i=1}^{M_z} \gamma_{iz} \int_{\lambda_{iz}}^{1} (r-\lambda_{iz})W_z$
$G_{6z} = \frac{1}{3}G_{3z} - G_{1z}\left(\frac{1}{3}G_{1z} - G_{2z}\right) + 2G_{4z} - G_{2z}^2$
$V_{1x} = \frac{1}{12}\left(\frac{1}{3}G_{3x} - \frac{1}{12}G_{1x}^2 + 2G_{4x}\right)$
$V_{2x} = \left(-\frac{1}{24}G_{2x}G_{1x} + \frac{1}{12}G_{3x} + \frac{1}{2}G_{4x}\right)$
$V_{3x} = \frac{1}{4}\left(\frac{1}{4}G_{2x}^2 - \frac{1}{3}G_{3x} - 2G_{4x}\right)$
$V_{4z} = \frac{1}{12}\left(G_{2z} - \frac{1}{2}G_{1z}\right)$
$V_{5z} = \frac{1}{24}\left(2G_{1z} - 3G_{2z}\right)$
$\sigma_{u(A3)} = 2G_{1y}G_{5y}N_{2x} - G_{5y}^2 + D_{2x}\left(2G_{4y} + \frac{1}{3}G_{3y}\right) - \frac{1}{4}G_{2y}^2D_3 - \frac{1}{3}G_{1y}^2D_1$
$\quad -G_{2y}G_{5y}N_{3x} + G_{2y}G_{1y}N_1$
$\sigma_{u(B3)} = \int W_y\left(V_{2x}\int rW_y - V_{1x}\int W_y\right) - \frac{1}{48}\left(G_{5x}^2 + G_{6x}\int W_y^2\right) + V_{3x}\left(\int rW_y\right)^2$
$\quad +G_{5x}\left(V_{4x}\int W_y - V_{5x}\int rW_y\right)$
$\lambda_{u(i,j)} = \max(\lambda_{z,i}, \lambda_{z,j}),\ i,j = 1,2,...,M_z$
$\lambda_{l(i,j)} = \min(\lambda_{z,i}, \lambda_{z,j})$
$\lambda_{d(i,j)} = \lambda_{u(i,j)} - \lambda_{l(i,j)}$

Table A1

DGPs, regression models and orders in probability of statistics

Process for y Process for x	Regression	$\widehat{\delta}$	$t_{\widehat{\delta}}$
$I(1)$ $I(0)+trend$	$y_t = \alpha + \delta x_t + u_t$	$O_p(T^{-1/2})$	$O_p(T^{1/2})$
	$x_t = \alpha + \delta y_t + u_t$	$O_p(T^{1/2})$	$O_p(T^{1/2})$
	$y_t = \alpha + \beta t + \delta x_t + u_t$	$O_p(1)$	$O_p(1)$
	$x_t = \alpha + \beta t + \delta y_t + u_t$	$O_p(T)$	$O_p(1)$
$I(1)+drift$ $I(0)+trend$	$y_t = \alpha + \delta x_t + u_t$	$O_p(1)$	$O_p(T)$
	$x_t = \alpha + \delta y_t + u_t$	$O_p(1)$	$O_p(T)$
	$y_t = \alpha + \beta t + \delta x_t + u_t$	$O_p(1)$	$O_p(1)$
	$x_t = \alpha + \beta t + \delta y_t + u_t$	$O_p(T)$	$O_p(1)$
$I(1)+drift$ $I(0)+trend$ $+breaks$	$y_t = \alpha + \delta x_t + u_t$	$O_p(1)$	$O_p(T^{1/2})$
	$x_t = \alpha + \delta y_t + u_t$	$O_p(1)$	$O_p(T^{1/2})$
	$y_t = \alpha + \beta t + \delta x_t + u_t$	$O_p(T^{-1/2})$	$O_p(T^{1/2})$
	$x_t = \alpha + \beta t + \delta y_t + u_t$	$O_p(T^{1/2})$	$O_p(T^{1/2})$

6 References

Cappuccio, N. and D. Lubian (1997), "Spurious Regressions Between I(1) Processes with Long Memory Errors", *Journal of Time Series Analysis*, 18, 341-354.

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

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

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.

Kormendi, R. C. and P. Meguire (1990), "A Multicountry Characterization of the Nonstationarity of Aggregate Output", *Journal of Money, Credit and Banking*, 22, 77-93.

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.

Mehl, A. (2000), "Unit Root Tests with Double Trend Breaks and the 1990s Recession in Japan", *Japan and the World Economy*, 12, 363-379.

Nelson, C.R. and C. Plosser (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications", *Journal of Monetary Economics,* 10, 139-162.

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

Noriega, A. and D. Ventosa-Santaulària (2005), "Spurious Regression Under Broken Trend Stationarity", Working Paper EM-0501, School of Economics, University of Guanajuato.

Perron, P. (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.

Perron, P. (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.

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