Author(s): Ba Chu and Roman Kozhan

Article Title: Spurious Regressions of Stationary AR(p) Processes with Structural Breaks

Year of publication: 2010

Link to published article:
http://dx.doi.org/10.2202/1558-3708.1781

Publisher statement: None

Studies in Nonlinear Dynamics & Econometrics

Volume 15, *Issue* 1 2011 *Article* 1

Spurious Regressions of Stationary *AR(p)* Processes with Structural Breaks

Ba Chu* Roman Kozhan†

*Carleton University, ba_chu@carleton.ca
†University of Warwick, roman.kozhan@wbs.ac.uk

Spurious Regressions of Stationary *AR(p)* Processes with Structural Breaks*

Ba Chu and Roman Kozhan

Abstract

When a pair of independent series is highly persistent, there is a spurious regression bias in a regression between these series, closely related to the classic studies of Granger and Newbold (1974). Although this is well known to occur with independent *I(1)* processes, this paper provides theoretical and numerical evidence that the phenomenon of spurious regression also arises in regressions between stationary *AR(p)* processes with structural breaks, which occur at different points in time, in the means and the trends. The intuition behind this is that structural breaks can increase the persistence levels in the processes (e.g., Granger and Hyung (2004)), which then leads to spurious regressions. These phenomena occur for general distributions and serial dependence of the innovation terms.

---

*Ba Chu, Department of Economics, Carleton University, B-857 Loeb Building, 1125 Colonel By Drive, Ottawa, ON K1S 5B6, Canada. Phone: +1 (613) 520 2600 (ext. 1546). Fax: +1 (613) 520 3906. E-mail: ba chu@carleton.ca. Web Page: http-server.carleton.ca/∼bchu. Roman Kozhan, Finance Group,Warwick Business School, University ofWarwick, Coventry CV4 7AL, UK. Phone: +44 (0)24 7652 4118. Fax: +44 (0)24 7652 3779. E-mail: roman.kozhan@wbs.ac.uk. Web Page: www.warwick.ac.uk/staff/R.Kozhan. Our sincere thanks are due to the editor – Prof. Bruce Mizrach and two anonymous referees, whose many insightful comments have helped us to substantially revise the original manuscript. The usual disclaimer applies.

1 Introduction

Simulation studies of Granger and Newbold [1974] warned that spurious relations may be found between the levels of trending time series that are actually *independent*. Later, Phillips [1986, 1998] provide an elegant asymptotic framework that vindicates the simulation results. Applied economists become increasingly aware of this problem. For instance, Ferson, Sarkissian, and Simin [2003] find that many predictive stock return regressions in the literature, based on individual predicting variables, may be spurious.

It has been known for some time that the spurious regressions do not hold only for independent random walks, but also for other persistent processes, such as high-order integrated processes (see, e.g., Marmol [1995], among others), fractionally integrated processes (see, e.g., Tsay and Chung [1999], among others), *I(1)* processes with infinite variance errors (see Tsay [1999]), and positively autocorrelated processes on long moving averages (see Granger, Hyung, and Jeon [2001]).

The goal of the current paper is to investigate the possible existence of spurious relations in a pair of stationary, invertible *AR(p)* processes with *weakly dependent* innovations and structural breaks, which occur at different points in time, in the means and the trends. This problem has an aesthetic appeal and also practical implications. We have shown that the strength of these types of spurious relationship is rather severe for the type of processes under our study; and that the rates of convergence for the OLS statistics to the corresponding limiting values do not depend on the starting values and the break locations of the underlying processes.

Our analytical framework, although simplified in a number of respects (such as only one break point is considered per se), proves tractable in addressing the main issues of spurious regressions. Nevertheless, we also provide a sketch of theoretical results for the case of many break points. Hitherto, the plan of this paper is as follows: Section **2** deals with structural breaks in means; Section **3** deals with structural breaks in trends; Section **4** provides some simulation evidences; and Section **5** concludes this paper. Last but not least, results of technical flavor but essential for the paper are collected in the appendices at the end of the paper.

2 Spurious Regression: Structural Breaks in Mean

The data generating processes (DGPs) for two *independent* stationary[1] time series are defined as follows:

[1] We shall note at this point that the terminology ‘stationary’ is used to merely mean that the roots of lag polynomials lie outside the unit circle.

$$\begin{aligned} A(L)X_t &= c^{(x)}\mathbf{1}_{\{t>[T\tau^{(x)}]\}} + u_t, \\ B(L)Y_t &= c^{(y)}\mathbf{1}_{\{t>[T\tau^{(y)}]\}} + v_t, \end{aligned} \tag{2.1}$$

where intercepts, $c^{(x)}$ and $c^{(y)}$, are the break levels of X_t and Y_t, respectively; $A(L)$ and $B(L)$ are lag polynomials with their roots outside the unit circle; $\tau^{(x)}$ and $\tau^{(y)}$ are the break points of X_t and Y_t, respectively; and innovation terms, u_t and v_t, are contemporaneously *independent* and fulfill Assumption 1 (below).

Let $\mathscr{F}_t$ and $\mathscr{F}^t$ denote the σ-fields generated, respectively, by (u_s, v_s), $-\infty < s \leq t$, and (u_s, v_s), $t \leq s < \infty$. Given a positive integer, k, we set

$$\alpha(k) = \sup\left\{ \left| P(A \bigcap B) - P(A)P(B) \right| : A \in \mathscr{F}_t \text{ and } B \in \mathscr{F}^{t+k} \right\}.$$

This is Rosenblatt's [1956] mixing coefficient. The stationary process is said to be α-mixing or strongly mixing if $\alpha(k) \longrightarrow 0$ as $k \longrightarrow \infty$.

Assumption 1. *The innovation terms, (u_t, v_t), are strictly stationary with zero means such that $E[|u_0|^{2(r+\delta)}] < \infty$ and $E[|v_0|^{2(r+\delta)}] < \infty$ for some integer, $r > 1$, and a positive generic constant, δ. The mixing coefficient satisfies $\alpha(k) = \mathscr{O}\left(k^{-\frac{r}{r-1}-\varepsilon}\right)$, where ε is a positive generic constant.*

Assumption 1 includes a wide variety of possible data-generating mechanisms. For example, both the ARMA process and the MA(∞) process can become strongly mixing under some regularity conditions (see, e.g., Gorodetskii [1977] and Withers [1981]). The mixing condition in this assumption essentially controls the extent of permissible temporal dependence in the process in the relation to the probability of outlier occurrences.

Eq. (2.1) can be rewritten as

$$\begin{aligned} X_t &= A^{-1}(L)c^{(x)}\mathbf{1}_{\{t>[T\tau^{(x)}]\}} + A^{-1}(L)u_t, \\ Y_t &= B^{-1}(L)c^{(y)}\mathbf{1}_{\{t>[T\tau^{(y)}]\}} + B^{-1}(L)v_t. \end{aligned}$$

Let us define the following inverse lag operators: $A^*(L) = A^{-1}(L)$ and $B^*(L) = B^{-1}(L)$. An application of the Beveridge-Nelson (BN) decomposition, $A^*(L) = A^*(1) + A^{\diamond}(L)(1-L)$ and $B^*(L) = B^*(1) + B^{\diamond}(L)(1-L)$, yields

$$\begin{aligned} A^*(L)u_t &= A^*(1)u_t + \Delta u_t^* \\ B^*(L)v_t &= B^*(1)v_t + \Delta v_t^*, \end{aligned}$$

where $u_t^* = A^{\diamond}(L)u_t$ and $v_t^* = B^{\diamond}(L)v_t$ with $A^{\diamond}$ and $B^{\diamond}$ having their roots outside the unit circle.

Now, let us consider the following regression equation:

$$Y_t = \hat{\gamma} + \hat{\beta} X_t + \hat{w}_t,$$

where $\hat{\gamma}$, $\hat{\beta}$ and $\hat{w}_t$ are the OLS estimators, defined as

$$\begin{aligned} \hat{\beta} &= \frac{\sum_1^T Y_t (X_t - \overline{X})}{\sum_1^T (X_t - \overline{X})^2}, \\ \hat{\gamma} &= \overline{Y} - \hat{\beta}\overline{X}, \\ \hat{w}_t &= Y_t - \hat{\gamma} - \hat{\beta} X_t. \end{aligned}$$

Lemma 1. *Suppose that (X_t, Y_t) are generated by Eq. (2.1). The innovations, u_t and v_t, are independent and satisfy Assumption 1. Then, as $T \longrightarrow \infty$,*

$$T^{-1}\sum_{t=1}^{T} X_t Y_t \overset{a.s.}{\Longrightarrow} \underbrace{A^*(1)B^*(1)c^{(x)}c^{(y)}\left(1 - \max(\tau^{(x)}, \tau^{(y)})\right)}_{\mathscr{L}_{XY}} \tag{2.2}$$

$$T^{-1}\sum_{t=1}^{T} X_t \overset{a.s.}{\Longrightarrow} \underbrace{(1 - \tau^{(x)})A^*(1)c^{(x)}}_{\mathscr{L}_X} \tag{2.3}$$

$$T^{-1}\sum_{t=1}^{T} Y_t \overset{a.s.}{\Longrightarrow} \underbrace{(1 - \tau^{(y)})B^*(1)c^{(y)}}_{\mathscr{L}_Y} \tag{2.4}$$

$$\frac{\sum_{t=1}^{T} X_t^2}{T} \overset{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} A^{*2}(1)\left\{\sigma_u^2 + \sigma_{\Delta u^*}^2 + (1 - \tau^{(x)}){c^{(x)}}^2\right\} \\ +2A^*(1)\sigma_{u\Delta u^*} \end{array}}_{\mathscr{L}_{X^2}} \tag{2.5}$$

$$\frac{\sum_{t=1}^{T} Y_t^2}{T} \overset{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} B^{*2}(1)\left\{\sigma_v^2 + \sigma_{\Delta v^*}^2 + (1 - \tau^{(y)}){c^{(y)}}^2\right\} \\ +2B^*(1)\sigma_{v\Delta v^*} \end{array}}_{\mathscr{L}_{Y^2}} \tag{2.6}$$

$$\frac{\sum_{t=s+1}^{T} (Y_t - \overline{Y})(Y_{t-s} - \overline{Y})}{T} \overset{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} B^{*2}(1)\left\{{c^{(y)}}^2 \tau^{(y)}\left(1 - \tau^{(y)}\right) + \sigma_{v_0 v_s} \right. \\ \left. +\sigma_{\Delta v_0^* \Delta v_s^*}\right\} + B^*(1)\left(\sigma_{v_0 \Delta v_s^*} + \sigma_{v_s \Delta v_0^*}\right) \end{array}}_{\mathscr{L}_{Y_0 Y_s}} \tag{2.7}$$

$$\frac{\sum_{t=s+1}^{T} (X_t - \overline{X})(X_{t-s} - \overline{X})}{T} \overset{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} A^{*2}(1)\left\{{c^{(x)}}^2 \tau^{(x)}\left(1 - \tau^{(x)}\right) + \sigma_{u_0 u_s} \right. \\ \left. +\sigma_{\Delta u_0^* \Delta u_s^*}\right\} + A^*(1)\left(\sigma_{u_0 \Delta u_s^*} + \sigma_{u_s \Delta u_0^*}\right) \end{array}}_{\mathscr{L}_{X_0 X_s}} \tag{2.8}$$

$$\frac{\sum_{t=s+1}^{T}(Y_t-\overline{Y})(X_{t-s}-\overline{X})}{T} \overset{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} A^*(1)B^*(1)c^{(x)}c^{(y)}\Big\{1-\max(\tau^{(x)},\tau^{(y)}) \\ -(1-\tau^{(x)})(1-\tau^{(y)})\Big\} \end{array}}_{\mathscr{L}_{Y_0X_s}} \tag{2.9}$$

$$\frac{\sum_{t=s+1}^{T}(Y_{t-s}-\overline{Y})(X_t-\overline{X})}{T} \overset{a.s.}{\Longrightarrow} \mathscr{L}_{Y_0X_s}, \tag{2.10}$$

where

$$\begin{array}{ll}
\sigma_u^2 = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=1}^{T}u_t^2; & \sigma_{\Delta u^*}^2 = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=1}^{T}(\Delta u_t^*)^2; \\
\sigma_{u\Delta u^*} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=1}^{T}u_t\Delta u_t^*; & \sigma_v^2 = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=1}^{T}v_t^2; \\
\sigma_{\Delta v^*}^2 = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=1}^{T}(\Delta v_t^*)^2; & \sigma_{v\Delta v^*} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=1}^{T}v_t\Delta v_t^*; \\
\sigma_{v_0v_s} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=s+1}^{T}v_tv_{t-s}; & \sigma_{\Delta v_0^*\Delta v_s^*} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=s+1}^{T}\Delta v_t^*\Delta v_{t-s}^*; \\
\sigma_{v_0\Delta v_s^*} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=s+1}^{T}v_t\Delta v_{t-s}^*; & \sigma_{v_s\Delta v_0^*} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=s+1}^{T}v_{t-s}\Delta v_t^*; \\
\sigma_{u_0u_s} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=s+1}^{T}u_tu_{t-s}; & \sigma_{\Delta u_0^*\Delta u_s^*} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=s+1}^{T}\Delta u_t^*\Delta u_{t-s}^*; \\
\sigma_{u_0\Delta u_s^*} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=s+1}^{T}u_t\Delta u_{t-s}^*; & \sigma_{u_s\Delta u_0^*} = \lim\limits_{T\longrightarrow\infty} T^{-1}\sum_{t=s+1}^{T}u_{t-s}\Delta u_t^*.
\end{array}$$

Moreover, Eqs. (2.2)-(2.10) *hold irrespective of the initial conditions assigned to* X_0 *and* Y_0.

Proof. The proof is presented in Appendix. □

The limiting behaviors of regression statistics are stated in Theorem 1 below.

Theorem 1. *Suppose that the conditions of Lemma 1 are satisfied. Then, as* $T \to \infty$,

$$\hat{\beta} \overset{a.s.}{\Longrightarrow} \frac{\mathscr{L}_{XY}-\mathscr{L}_X\mathscr{L}_Y}{\mathscr{L}_{X^2}-\mathscr{L}_X^2} = \mathscr{L}_{\widehat{\beta}};$$

$$\hat{\gamma} \overset{a.s.}{\Longrightarrow} \mathscr{L}_Y - \mathscr{L}_{\widehat{\beta}}\mathscr{L}_X = \mathscr{L}_{\widehat{\gamma}};$$

$$\hat{s}^2 = T^{-1}\sum_{1}^{T}\Big\{(Y_t-\overline{Y})-\hat{\beta}(X_t-\overline{X})\Big\}^2 \overset{a.s.}{\Longrightarrow} \begin{array}{l} \{\mathscr{L}_{Y^2}-\mathscr{L}_Y^2\}+\mathscr{L}_{\widehat{\beta}}^2\{\mathscr{L}_{X^2}-\mathscr{L}_X^2\} \\ -2\mathscr{L}_{\widehat{\beta}}\{\mathscr{L}_{XY}-\mathscr{L}_X\mathscr{L}_Y\} = \mathscr{L}_{\widehat{s}^2}; \end{array}$$

$$T^{-1/2}t_{\hat{\beta}} = \frac{\hat{\beta}}{\hat{s}\Big(T^{-1}\sum_{t=1}^{T}(X_t-\overline{X})^2\Big)^{-1/2}} \overset{a.s.}{\Longrightarrow} \frac{\mathscr{L}_{\widehat{\beta}}}{\mathscr{L}_{\widehat{s}^2}^{1/2}\{\mathscr{L}_{X^2}-\mathscr{L}_X^2\}^{-1/2}};$$

$$T^{-1/2}t_{\hat{\gamma}} = \frac{\hat{\gamma}}{\hat{s}} \frac{\left(T^{-1}\sum_{t=1}^{T}(X_t - \overline{X})^2\right)^{1/2}}{\left(T^{-1}\sum_{t=1}^{T}X_t^2\right)^{1/2}} \overset{a.s.}{\Longrightarrow} \frac{\mathscr{L}_{\widehat{\gamma}}}{\mathscr{L}_{\hat{s}^2}^{1/2}} \frac{\{\mathscr{L}_{X^2} - \mathscr{L}_X^2\}^{1/2}}{\mathscr{L}_{X^2}^{1/2}};$$

$$\hat{r}_s = \frac{\sum_{s+1}^{T}\hat{w}_t\hat{w}_{t-s}}{\sum_1^T \hat{w}_t^2} \overset{a.s.}{\Longrightarrow} \frac{\mathscr{L}_{Y_0Y_s} + \mathscr{L}_{\widehat{\beta}}^2\mathscr{L}_{X_0X_s} - 2\mathscr{L}_{\widehat{\beta}}\mathscr{L}_{Y_0X_s}}{\mathscr{L}_{Y^2} + \mathscr{L}_{\widehat{\gamma}}^2 + \mathscr{L}_{\widehat{\beta}}^2\mathscr{L}_{X^2} - 2\mathscr{L}_{\widehat{\gamma}}\mathscr{L}_Y - 2\mathscr{L}_{\widehat{\beta}}\mathscr{L}_{XY} + 2\mathscr{L}_{\widehat{\gamma}}\mathscr{L}_{\widehat{\beta}}\mathscr{L}_X}.$$

Proof. The proof is presented in Appendix. □

Remark 1. *Theorem 1 is a close-to-trivial but quite enlightening case of spurious regression. The OLS coefficient, $\widehat{\beta}$, does not tend to zero, as it is expected to when two* independent *d.g.p. are regressed on each other. Moreover, the t statistics diverge at the rate of $\mathscr{O}_{a.s.}(T^{1/2})$, implicating that the null hypothesis of zero β always fails to be rejected. The serial correlation coefficients of the regression residuals converge almost surely to a non-zero constant, which is not the case for regression of two independent processes. In general, all of these results differ from the conventional theory of regression with stationary processes.*

It is worth noting that the main results hold only for the case where structural breaks occur at different points in time. In other words, if the break point is at the same time the spurious regression effect disappears.

Corollary 1. *Suppose, there are no structural breaks in means. Then, as $T \to \infty$,*

$$T^{-1}\sum_{t=1}^{T}X_tY_t \overset{a.s.}{\Longrightarrow} c^{(x)}c^{(y)}A^*(1)B^*(1)$$
$$T^{-1}\sum_{t=1}^{T}X_t \overset{a.s.}{\Longrightarrow} c^{(x)}A^*(1)$$
$$T^{-1}\sum_{t=1}^{T}Y_t \overset{a.s.}{\Longrightarrow} c^{(y)}B^*(1).$$

Moreover, suppose that X_t has a structural break but Y_t does not. Then, as $T \to \infty$,

$$T^{-1}\sum_{t=1}^{T}X_tY_t \overset{a.s.}{\Longrightarrow} c^{(x)}c^{(y)}(1-\tau^{(x)})A^*(1)B^*(1)$$
$$T^{-1}\sum_{t=1}^{T}X_t \overset{a.s.}{\Longrightarrow} (1-\tau^{(x)})c^{(x)}A^*(1)$$
$$T^{-1}\sum_{t=1}^{T}Y_t \overset{a.s.}{\Longrightarrow} B^*(1)c^{(y)}.$$

The above limiting behaviors imply $\widehat{\beta} \overset{a.s.}{\Longrightarrow} 0$ (i.e., no spurious regression).

3 Spurious Regression: Structural Breaks in Trend

We shall consider the following d.g.p.:

$$\begin{aligned} A(L)X_t &= c^{(x)} + \mu^{(x)} t + \mu_1^{(x)} \left(t - [T\tau^{(x)}]\right) \mathbf{1}_{\{t>[T\tau^{(x)}]\}} + u_t, \\ B(L)Y_t &= c^{(y)} + \mu^{(y)} t + \mu_1^{(y)} \left(t - [T\tau^{(y)}]\right) \mathbf{1}_{\{t>[T\tau^{(y)}]\}} + v_t, \end{aligned} \tag{3.1}$$

where lag polynomials, $A(L)$ and $B(L)$, have their roots lying outside the unit circle; $c^{(x)}$ and $c^{(y)}$ are the intercepts of X_t and Y_t, respectively; $\mu^{(x)}$ and $\mu_1^{(x)}$ are, respectively, the permanent trend and the transitory trend, resulting from a break, of the process X_t; $\mu^{(y)}$ and $\mu_1^{(y)}$ are, respectively, the permanent trend and the transitory trend, resulting from a break, of the process Y_t; and innovations, u_t and v_t), are independent and satisfies Assumption 1.

A BN decomposition of the inverse lag operators $A^*(L)$ and $B^*(L)$, as defined in Section 2, yields

$$\begin{aligned} X_t &= A^*(1)c^{(x)} + \mu^{(x)} \left(A^*(1)t + A^\diamond(1)\right) \\ &+ \mu_1^{(x)} \left\{A^*(1)(t - [T\tau^{(x)}]) + A^\diamond(1)\right\} \mathbf{1}_{\{t>[T\tau^{(x)}]\}} + A^*(L)u_t \text{ and} \\ Y_t &= B^*(1)c^{(y)} + \mu^{(y)} \left(B^*(1)t + B^\diamond(1)\right) \\ &+ \mu_1^{(y)} \left\{B^*(1)(t - [T\tau^{(y)}]) + B^\diamond(1)\right\} \mathbf{1}_{\{t>[T\tau^{(y)}]\}} + B^*(L)v_t, \end{aligned}$$

where $u_t = A^*(1)u_t + \Delta u_t^*$ and $v_t = B^*(1)v_t + \Delta v_t^*$.

Next, we formulate the following OLS regression:

$$Y_t = \hat{\gamma}_1 + \hat{\gamma}_2 t + \hat{\beta} X_t + \hat{w}_t,$$

where

$$\begin{bmatrix} \hat{\gamma}_1 \\ \hat{\beta} \\ \hat{\gamma}_2 \end{bmatrix} = \begin{bmatrix} T & \sum_1^T X_t & \sum_1^T t \\ \sum_1^T X_t & \sum_1^T X_t^2 & \sum_1^T tX_t \\ \sum_1^T t & \sum_1^T tX_t & \sum_1^T t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_1^T Y_t \\ \sum_1^T X_tY_t \\ \sum_1^T tY_t \end{bmatrix}. \tag{3.2}$$

To facilitate the asymptotic argument for the OLS statistics, we shall first state Lemma 2 (below). (Also to avoid any unnecessary confusion, we shall note here that notations, $\mathscr{L}_\bullet$, are specific to Section 3 and independent of $\mathscr{L}_\bullet$ in other sections.)

Lemma 2. *Suppose that (X_t, Y_t) is generated by Eq. (3.1). The innovations, u_t and v_t, are independent and satisfy Assumption 1. Then, as $T \longrightarrow \infty$,*

$$T^{-2}\sum_{1}^{T} X_t \stackrel{a.s.}{\Longrightarrow} \underbrace{A^*(1)\left\{\mu^{(x)}\int_0^1 sds+\mu_1^{(x)}\int_{\tau^{(x)}}^1(s-\tau^{(x)})ds\right\}}_{\mathscr{L}_X} \tag{3.3}$$

$$T^{-3}\sum_{1}^{T} X_t^2 \stackrel{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} A^{*2}(1)\left\{\mu^{(x)2}\int_0^1 s^2ds+\mu_1^{(x)2}\int_{\tau^{(x)}}^1\left(s-\tau^{(x)}\right)^2ds\right. \\ \left. +\,2\mu^{(x)}\mu_1^{(x)}\int_{\tau^{(x)}}^1 s\left(s-\tau^{(x)}\right)ds\right\} \end{array}}_{\mathscr{L}_{X^2}} \tag{3.4}$$

$$T^{-2}\sum_{1}^{T} t \Longrightarrow \underbrace{\int_0^1 sds}_{\mathscr{L}_t} \tag{3.5}$$

$$T^{-3}\sum_{1}^{T} tX_t \stackrel{a.s.}{\Longrightarrow} \underbrace{A^*(1)\left\{\mu^{(x)}\int_0^1 s^2ds+\mu_1^{(x)}\int_{\tau^{(x)}}^1 s\left(s-\tau^{(x)}\right)ds\right\}}_{\mathscr{L}_{tX}} \tag{3.6}$$

$$T^{-3}\sum_{1}^{T} t^2 \Longrightarrow \underbrace{\int_0^1 s^2ds}_{\mathscr{L}_{t^2}} \tag{3.7}$$

$$T^{-2}\sum_{1}^{T} Y_t \stackrel{a.s.}{\Longrightarrow} \underbrace{B^*(1)\left\{\mu^{(y)}\int_0^1 sds+\mu_1^{(y)}\int_{\tau^{(y)}}^1(s-\tau^{(y)})ds\right\}}_{\mathscr{L}_Y} \tag{3.8}$$

$$T^{-3}\sum_{1}^{T} X_tY_t \stackrel{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} A^*(1)B^*(1)\left\{\mu^{(x)}\mu^{(y)}\int_0^1 s^2ds\right. \\ +\mu_1^{(x)}\mu_1^{(y)}\int_{\max(\tau^{(x)},\tau^{(y)})}^1\left(s-\tau^{(x)}\right)\left(s-\tau^{(y)}\right)ds \\ +\mu^{(x)}\mu_1^{(y)}\int_{\tau^{(y)}}^1 s\left(s-\tau^{(y)}\right)ds \\ \left. +\mu^{(y)}\mu_1^{(x)}\int_{\tau^{(x)}}^1 s\left(s-\tau^{(x)}\right)ds\right\} \end{array}}_{\mathscr{L}_{XY}} \tag{3.9}$$

$$T^{-2}\sum_{1}^{T} tY_t \stackrel{a.s.}{\Longrightarrow} \underbrace{B^*(1)\left\{\mu^{(y)}\int_0^1 s^2ds+\mu_1^{(y)}\int_{\tau^{(y)}}^1 s\left(s-\tau^{(y)}\right)ds\right\}}_{\mathscr{L}_{tY}} \tag{3.10}$$

$$T^{-2}\sum_{1}^{T} Y_t^2 \stackrel{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} B^{*2}(1)\left\{\mu^{(y)2}\int_0^1 s^2ds+\mu_1^{(y)2}\int_{\tau^{(y)}}^1\left(s-\tau^{(y)}\right)^2ds\right. \\ \left. +\,2\mu^{(y)}\mu_1^{(y)}\int_{\tau^{(y)}}^1 s\left(s-\tau^{(y)}\right)ds\right\}. \end{array}}_{\mathscr{L}_{Y^2}} \tag{3.11}$$

Moreover, Eqs. (3.3)-(3.11) *hold irrespective of the initial conditions assigned to* X_0 *and* Y_0.

Proof. The proof is presented in Appendix. □

Let's define

$$\mathscr{A} = \begin{bmatrix} 1 & \mathscr{L}_X & \mathscr{L}_t \\ \mathscr{L}_X & \mathscr{L}_{X^2} & \mathscr{L}_{tX} \\ \mathscr{L}_t & \mathscr{L}_{tX} & \mathscr{L}_{t^2} \end{bmatrix} \text{ and } \mathscr{B} = \begin{bmatrix} \mathscr{L}_Y \\ \mathscr{L}_{XY} \\ \mathscr{L}_{tY} \end{bmatrix},$$

where the elements of the matrices are defined in Lemma 2.

The limiting behavior of regression statistics are stated in Theorem 2.

Theorem 2. *Suppose that the conditions of Lemma 2 are satisfied. Then, as* $T \longrightarrow \infty$,

$$\begin{bmatrix} T^{-1}\widehat{\gamma}_1 \\ \widehat{\beta} \\ \widehat{\gamma}_2 \end{bmatrix} \overset{a.s.}{\Longrightarrow} \mathscr{A}^{-1}\mathscr{B} = \begin{bmatrix} \mathscr{L}_{\widehat{\gamma}_1} \\ \mathscr{L}_{\widehat{\beta}} \\ \mathscr{L}_{\widehat{\gamma}_2} \end{bmatrix} \tag{3.12}$$

$$T^{-2}\hat{s}^2 \overset{a.s.}{\Longrightarrow} \underbrace{\begin{array}{l} \mathscr{L}_{Y^2} + \mathscr{L}_{\widehat{\gamma}_2}{}^2\mathscr{L}_{t^2} + \mathscr{L}_{\widehat{\beta}}{}^2\mathscr{L}_{X^2} - 2\mathscr{L}_{\widehat{\gamma}_1}\mathscr{L}_Y - 2\mathscr{L}_{\widehat{\gamma}_2}\mathscr{L}_{tY} - 2\mathscr{L}_{\widehat{\beta}}\mathscr{L}_{XY} \\ +2\mathscr{L}_{\widehat{\gamma}_1}\mathscr{L}_{\widehat{\gamma}_2}\mathscr{L}_t + 2\mathscr{L}_{\widehat{\beta}}\mathscr{L}_{\widehat{\gamma}_1}\mathscr{L}_X + 2\mathscr{L}_{\widehat{\beta}}\mathscr{L}_{\widehat{\gamma}_2}\mathscr{L}_{tX}; \end{array}}_{\mathscr{L}_{\hat{s}^2}} \tag{3.13}$$

$$T^{-1/2}t_{\hat{\gamma}_1} \overset{a.s.}{\Longrightarrow} \frac{\mathscr{L}_{\widehat{\gamma}_1}}{\mathscr{L}_{\sigma_{\widehat{\gamma}_1}}} \tag{3.14}$$

$$T^{-1/2}t_{\hat{\beta}} \overset{a.s.}{\Longrightarrow} \frac{\mathscr{L}_{\widehat{\beta}}}{\mathscr{L}_{\sigma_{\widehat{\beta}}}} \tag{3.15}$$

$$T^{-1/2}t_{\hat{\gamma}_2} \overset{a.s.}{\Longrightarrow} \frac{\mathscr{L}_{\widehat{\gamma}_2}}{\mathscr{L}_{\sigma_{\widehat{\gamma}_2}}} \tag{3.16}$$

where

$$\begin{aligned} \mathscr{L}_{\sigma_{\widehat{\gamma}_1}} &= \mathscr{L}_{\hat{s}^2}^{1/2}\mathbb{I}_{(1)}\mathscr{A}^{-1}\mathbb{I}'_{(1)} \\ \mathscr{L}_{\sigma_{\widehat{\beta}}} &= \mathscr{L}_{\hat{s}^2}^{1/2}\mathbb{I}_{(2)}\mathscr{A}^{-1}\mathbb{I}'_{(2)} \\ \mathscr{L}_{\sigma_{\widehat{\gamma}_2}} &= \mathscr{L}_{\hat{s}^2}^{1/2}\mathbb{I}_{(3)}\mathscr{A}^{-1}\mathbb{I}'_{(3)} \end{aligned}$$

($\mathbb{I}_{(i)}$ *denotes a vector in which the i-th element is one and the other elements are zeros.)*

$$R^2 = \hat{\beta}^2 \frac{\sum_1^T (X_t - \overline{X})^2}{\sum_1^T (Y_t - \overline{Y})^2} \overset{a.s.}{\Longrightarrow} \mathscr{L}_{\widehat{\beta}}^2 \left\{ \frac{\mathscr{L}_X}{\mathscr{L}_Y} \right\}^2 \tag{3.17}$$

$$\hat{r}_s = \frac{1}{\mathscr{L}_{\hat{s}^2}} \left(\begin{array}{l} (1-2\mathscr{L}_{\widehat{\beta}})\mathscr{L}_{XY} - 2\mathscr{L}_{\widehat{\gamma}_1}\mathscr{L}_Y + 2\mathscr{L}_{\widehat{\gamma}_1}\mathscr{L}_{\widehat{\gamma}_2}\mathscr{L}_t + 2\mathscr{L}_{\widehat{\gamma}_1}\mathscr{L}_{\widehat{\beta}}\mathscr{L}_X \\ +\mathscr{L}^2_{\widehat{\beta}}\mathscr{L}_{X^2} - 2\mathscr{L}_{\widehat{\gamma}_2}\mathscr{L}_{tY} + \mathscr{L}^2_{\widehat{\gamma}_2}\mathscr{L}_{t^2} + 2\mathscr{L}_{\widehat{\beta}}\mathscr{L}_{\widehat{\gamma}_2}\mathscr{L}_{tX} \end{array} \right). \quad (3.18)$$

Proof. The proof is presented in Appendix. □

Remark 2. *The heuristics for the results in Theorem 2 is rather succinct. Since the limiting behaviors of the OLS estimates are dominated by the trend components, the OLS estimates do not converge to zero, as one expects when a process is regressed on another* unrelated *process, unless all the trends are canceled out. Also due to the presence of breaks in the trends, the OLS statistics – regression sum of squared errors, $\widehat{s}$, and t statistics – diverge at the rate of $\mathscr{O}_{a.s.}(T^{1/2})$, suggesting that the null hypotheses of zero regression coefficients always fail to be rejected in this case. In addition, the serial correlation coefficients of the regression residuals and the R^2 coefficient converge almost surely to non-zero constants, which is not the case in the conventional theory of regression for stationary processes.*

Corollary 2. *Suppose that X_t has a structural break but Y_t does not. Then,*

$$\begin{aligned} \mathscr{L}_Y &= \mu^{(y)}B^*(1)\mathscr{L}_t \\ \mathscr{L}_{XY} &= A^*(1)B^*(1)\left(\mu^{(x)}\mu^{(y)}\mathscr{L}_{t^2} + \mu_1^{(x)}\mu^{(y)}\left(\int_{\tau^{(x)}}^1 s^2 ds - \tau^{(x)}\int_{\tau^{(x)}}^1 sds\right)\right) \\ \mathscr{L}_{tY} &= \mu^{(y)}B^*(1)\mathscr{L}_{t^2} \\ \mathscr{L}_{Y^2} &= {\mu^{(y)}}^2 B^{*2}(1)\mathscr{L}_{t^2}. \end{aligned}$$

Theorem 2 still holds.

4 Simulations

The validity of our theorems for approximating the distributions of the OLS statistics in small samples can be legitimately questioned. To give some idea of the significance of our theoretical results, we shall provide some simulation studies. In the sequel, we run the following two regressions:

$$Y_t = \gamma + \beta X_t + w_t \quad (4.1)$$

$$Y_t = \gamma_1 + \gamma_2 t + \beta X_t + w_t, \quad (4.2)$$

where X_t and Y_t are generated by *AR(1)* processes with structural breaks in the mean and in the trend, as defined in Eq. (2.1) and Eq. (3.1) respectively.

Structural Break in Mean

First, we generate artificial data from

$$\begin{aligned} X_t &= \phi_x X_{t-1} + c^{(x)} \mathbf{1}_{\{t>[T\tau^{(x)}]\}} + u_t, \\ Y_t &= \phi_y Y_{t-1} + c^{(y)} \mathbf{1}_{\{t>[T\tau^{(y)}]\}} + v_t, \end{aligned} \tag{4.3}$$

where we set $\tau^{(x)} = 0.4$, $\tau^{(y)} = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, and $\phi_y = 0.75$. In this simulation, for simplicity we shall assume that the processes u_t and v_t are *independent white noises*. This is, unfortunately, because simulations for general weakly dependent innovations are quite complicated to implement. The processes in Eq. (4.3) are initialized by $X_0 = 10$ and $Y_0 = 10$. In order to evaluate the convergence rates of the OLS estimates, $\hat{\gamma}$ and $\hat{\beta}$, their t-statistics, t_γ and t_β, and the serial correlation coefficient, $\widehat{r}_s$, to the corresponding limits, we generate samples $\{X_t, Y_t\}_{t=1}^{T}$ of size T from $1,000$ to $5,000,000$.

Table 1: Simulations: Structural Breaks in Means

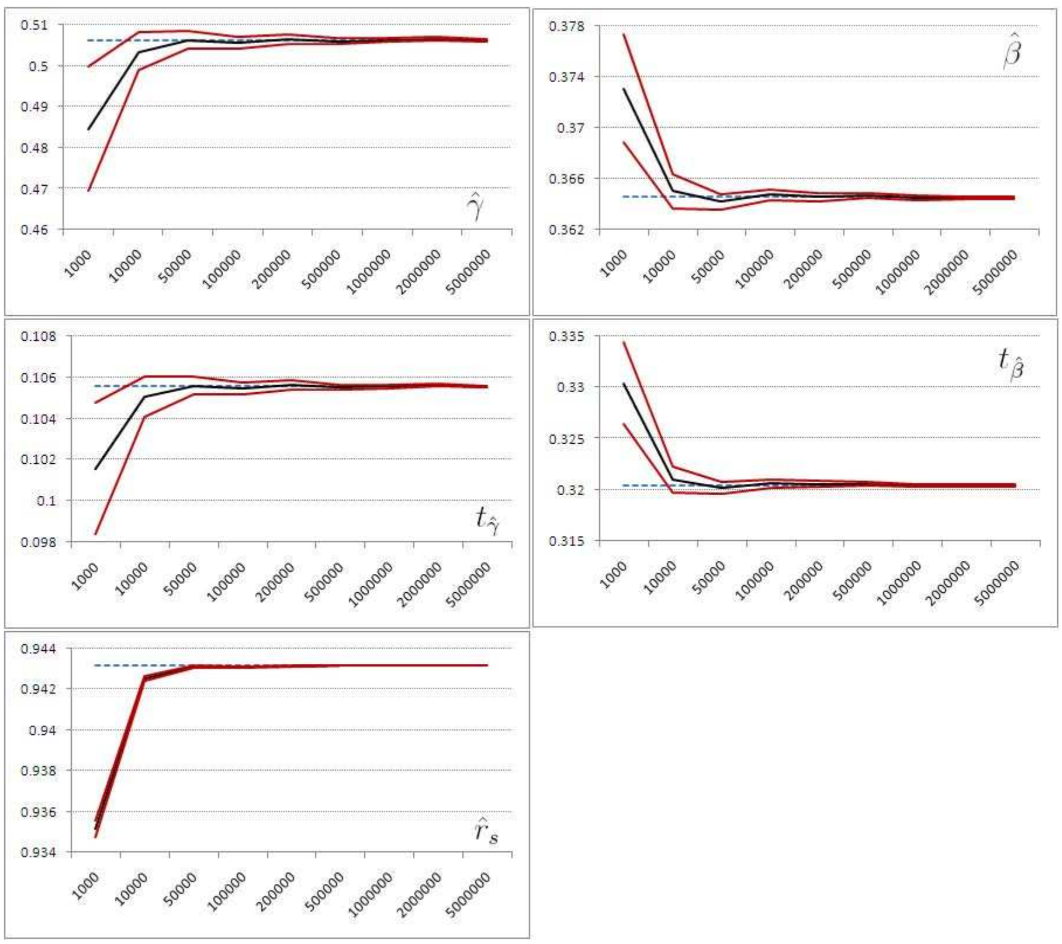
This table presents the simulation results for the OLS statistics convergence under a structural break in the mean. First, two *independent AR(1)* processes, defined in Eq. (4.3), are simulated, given $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $x_0 = 10$, and $y_0 = 10$. Then the OLS coefficients, their t-statistics, and the serial correlations of regression residuals are estimated from these simulated data. Limiting values are computed from Theorem 1. This procedure is repeated 1,000 times to generate standard errors of the mean of estimates and their 95% confidence bounds. Standard errors are given in parentheses.

T	$\hat{\gamma}$	$\hat{\beta}$	$\hat{t}_\gamma$	$\hat{t}_\beta$	$\hat{r}_s, s=1$
1,000	0.484633 ($7.763 \cdot 10^{-3}$)	0.373060 ($2.158 \cdot 10^{-3}$)	0.101586 ($1.632 \cdot 10^{-3}$)	0.330414 ($2.039 \cdot 10^{-3}$)	0.935153 ($1.997 \cdot 10^{-4}$)
10,000	0.503534 ($2.391 \cdot 10^{-3}$)	0.364988 ($6.643 \cdot 10^{-4}$)	0.105044 ($5.035 \cdot 10^{-4}$)	0.321041 ($6.266 \cdot 10^{-4}$)	0.942552 ($6.001 \cdot 10^{-5}$)
50,000	0.506336 ($1.074 \cdot 10^{-3}$)	0.364193 ($2.957 \cdot 10^{-4}$)	0.105572 ($2.259 \cdot 10^{-4}$)	0.320185 ($2.773 \cdot 10^{-4}$)	0.943115 ($2.602 \cdot 10^{-5}$)
100,000	0.505706 ($7.328 \cdot 10^{-4}$)	0.364738 ($2.077 \cdot 10^{-4}$)	0.105434 ($1.549 \cdot 10^{-4}$)	0.320596 ($1.926 \cdot 10^{-4}$)	0.943121 ($1.825 \cdot 10^{-5}$)
200,000	0.506516 ($5.592 \cdot 10^{-4}$)	0.364592 ($1.528 \cdot 10^{-4}$)	0.105626 ($1.176 \cdot 10^{-4}$)	0.320566 ($1.412 \cdot 10^{-4}$)	0.943150 ($1.272 \cdot 10^{-6}$)
500,000	0.506062 ($3.344 \cdot 10^{-4}$)	0.364675 ($9.348 \cdot 10^{-5}$)	0.105514 ($7.074 \cdot 10^{-5}$)	0.320572 ($8.691 \cdot 10^{-5}$)	0.943178 ($8.240 \cdot 10^{-6}$)
1,000,000	0.506276 ($2.453 \cdot 10^{-4}$)	0.364507 ($6.725 \cdot 10^{-5}$)	0.105554 ($5.147 \cdot 10^{-5}$)	0.320415 ($6.308 \cdot 10^{-5}$)	0.943182 ($5.743 \cdot 10^{-6}$)
2,000,000	0.506574 ($1.797 \cdot 10^{-4}$)	0.364501 ($4.960 \cdot 10^{-5}$)	0.105625 ($3.810 \cdot 10^{-5}$)	0.320435 ($4.550 \cdot 10^{-5}$)	0.943186 ($4.148 \cdot 10^{-6}$)
5,000,000	0.506203 ($1.074 \cdot 10^{-4}$)	0.364515 ($2.931 \cdot 10^{-5}$)	0.105540 ($2.255 \cdot 10^{-5}$)	0.320436 ($2.734 \cdot 10^{-5}$)	0.943183 ($2.542 \cdot 10^{-6}$)
Limit Value	**0.506329**	**0.364557**	**0.105564**	**0.320468**	**0.943194**

The simulated means of the OLS statistics, the corresponding limiting values, and the 95% confidence bounds are presented in Table 1 and Figure 1. It is noteworthy that the simulated means of the OLS statistics merely approach the corresponding limits for T roughly equal to 1,000 observations.

Figure 1: Convergence of Estimators: Structural Breaks in Means

This figure shows the plots of the mean OLS statistics (black line) and their 95% confidence bounds (red lines) of the regression $y_t = \gamma + \beta x_t + w_t$. The processes y_t and x_t are independent $AR(1)$ processes, defined in Eq. (4.3), with $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $x_0 = 10$, and $y_0 = 10$. As the number of observations, T, increases, the OLS statistics converge to their limiting values (dashed line), based on Theorem 1.



Structural Break in Trend

To illustrate the case with structural breaks in the trends, we generate artificial data from

$$\begin{aligned} X_t &= \phi_x X_{t-1} + c^{(x)} + \mu^{(x)} t + \mu_1^{(x)} (t - [T\tau^{(x)}]) \mathbf{1}_{\{t > [T\tau^{(x)}]\}} + u_t, \\ Y_t &= \phi_y Y_{t-1} + \alpha_y + \mu_y t + \mu_1^{(y)} (t - [T\tau_y]) \mathbf{1}_{\{t > [T\tau_y]\}} + v_t, \end{aligned} \tag{4.4}$$

where we set $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $\mu_x = 0.2$, $\mu_1^{(x)} = 0.1$, $\mu_y = 0.3$, $\mu_1^{(y)} = 0.05$, with u_t and v_t are, as earlier, *independent white noises*. Using initial values, $X_0 = 10$ and $Y_0 = 10$, we run the regression in Eq. (4.2) using the simulated data and estimate the OLS statistics. We repeat the simulation 1,000 times to compute the means and the 95% confidence bounds of the OLS coefficients $\hat{\gamma}_1$, $\hat{\gamma}_2$ and $\hat{\beta}$, their t-statistics, the serial correlation coefficient of the residuals, and the determination coefficient R^2. The simulation results are presented in Table 2 and Figure 2. As one can see, the simulated means of the OLS coefficients, except for $\widehat{\gamma}_1$, become rather stable as T reaches 1,000 observations. However, it is worth mentioning at this point that the discrepancies between the true slope coefficients and their OLS estimates, as seen in Figures 1 and 2, are rather wide due to a large graphic scaling being used, not because these discrepancies are really high.

Finally, we also performed some sensitivity analyses to check the robustness of the previous results with respect to different parameters values (cf. Table 3). We implemented these analyses by choosing different starting values for processes X_t and Y_t. The results are in line with those presented earlier – coefficient estimates and the values of their t-statistics approach to their corresponding limit values.

Table 4 presents simulation results for a variety of break points, τ_x and τ_y. Precisely, we used some big values for τ_x and τ_y. Although the limits of the OLS estimates depend on break locations, the convergence rates are essentially unaffected.

5 Discussion and Conclusion

Although the specification we adopt in this paper may omit some potential features of the data, such as ARCH effects, we find that using this fairly standard *AR(p)* framework allows us to successfully address the questions whether spurious regressions occur in the presence of structural breaks. In summary, the thrust of the present paper has been to show evidences of spurious regressions in the presence of structural breaks in the means and the trends of *AR(p)* processes by analyzing the limiting properties of the standard OLS statistics.

Table 2: Simulations: Structural Breaks in Trends

This table presents the simulation results for the OLS statistics convergence under structural breaks in the trend. First, two *independent AR*(1) processes, defined in Eq. (4.4), are simulated, given $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $\mu_x = 0.2$, $\mu_{x,b} = 0.1$, $\mu_y = 0.3$, $\mu_{y,b} = 0.05$, $x_0 = 10$, and $y_0 = 10$. Then the regression coefficients, their t-statistics, and the serial correlations of regression residuals are estimated from these simulated data. Limiting values are computed from Theorem 2. This procedure is repeated 1,000 times to generate the standard errors of the mean of OLS statistics and their 95% confidence bounds. Standard errors are presented in parentheses.

T	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}$	$\hat{t}_{\gamma_1}$	$\hat{t}_{\gamma_2}$	$\hat{t}_{\beta}$	$\hat{R}^2$	$\hat{r}_s, s=1$
1,000	0.00696617 *(1.14·10^{-5})*	1.056422 *(1.59·10^{-4})*	0.123978 *(1.210·10^{-4})*	0.394673 *(7.029·10^{-4})*	4.442476 *(2.938·10^{-3})*	0.691611 *(8.215·10^{-4})*	0.0181748 *(3.55·10^{-5})*	0.986649 *(3.069·10^{-5})*
10,000	0.00309036 *(3.83·10^{-7})*	1.054361 *(5.22·10^{-6})*	0.125696 *(3.921·10^{-6})*	0.176181 *(2.238·10^{-5})*	4.473360 *(9.925·10^{-5})*	0.708730 *(2.736·10^{-5})*	0.0187166 *(1.16·10^{-6})*	0.999308 *(2.488·10^{-7})*
50,000	0.00275121 *(3.40·10^{-8})*	1.054134 *(4.57·10^{-7})*	0.125878 *(3.438·10^{-7})*	0.156599 *(1.972·10^{-6})*	4.466557 *(9.090·10^{-6})*	0.708934 *(2.459·10^{-6})*	0.0187752 *(1.02·10^{-7})*	0.999881 *(9.720·10^{-9})*
100,000	0.00270894 *(1.23·10^{-8})*	1.054104 *(1.69·10^{-7})*	0.125902 *(1.271·10^{-7})*	0.154157 *(7.088·10^{-7})*	4.465557 *(3.083·10^{-6})*	0.708941 *(8.560·10^{-7})*	0.0187828 *(3.79·10^{-8})*	0.999942 *(2.537·10^{-9})*
200,000	0.00268780 *(4.24·10^{-9})*	1.054089 *(5.62·10^{-8})*	0.125914 *(4.226·10^{-8})*	0.152937 *(2.451·10^{-7})*	4.465054 *(1.119·10^{-6})*	0.708946 *(2.905·10^{-7})*	0.0187866 *(1.26·10^{-8})*	0.999971 *(2.467·10^{-10})*
500,000	0.00267512 *(1.11·10^{-9})*	1.054080 *(1.44·10^{-8})*	0.125921 *(1.082·10^{-8})*	0.152205 *(6.364·10^{-8})*	4.464749 *(2.778·10^{-7})*	0.708948 *(7.303·10^{-8})*	0.0187889 *(3.23·10^{-9})*	0.999989 *(9.920·10^{-11})*
1,000,000	0.00267089 *(3.90·10^{-10})*	1.054077 *(5.81·10^{-9})*	0.125924 *(4.037·10^{-9})*	0.151961 *(2.241·10^{-8})*	4.464648 *(9.873·10^{-8})*	0.708949 *(2.696·10^{-8})*	0.0187897 *(1.20·10^{-9})*	0.999994 *(2.523·10^{-11})*
2,000,000	0.00266878 *(1.34·10^{-10})*	1.054076 *(1.94·10^{-9})*	0.125925 *(1.402·10^{-9})*	0.151839 *(7.727·10^{-9})*	4.464597 *(3.611·10^{-8})*	0.708949 *(9.356·10^{-9})*	0.0187901 *(4.15·10^{-10})*	0.999997 *(6.040·10^{-12})*
Limit Value	**0.00266667**	**1.054070**	**0.125926**	**0.151717**	**4.464550**	**0.708949**	**0.0187905**	**1.00000**

Figure 2: Convergence of Estimators: Structural Breaks in Trends

This figure shows the plots of the mean OLS statistics (black line) and their 95% confidence bounds (red line) of the regression $y_t = \gamma + \beta x_t + w_t$. The processes y_t and x_t are independent $AR(1)$ processes, as defined in Eq. (4.4), with $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $\mu_x = 0.2$, $\mu_{x,b} = 0.1$, $\mu_y = 0.3$, $\mu_{y,b} = 0.05$, $x_0 = 10$, and $y_0 = 10$. As the number of observations, T, increases, the OLS statistics converge to their limiting values (dashed line), computed from Theorem 2.

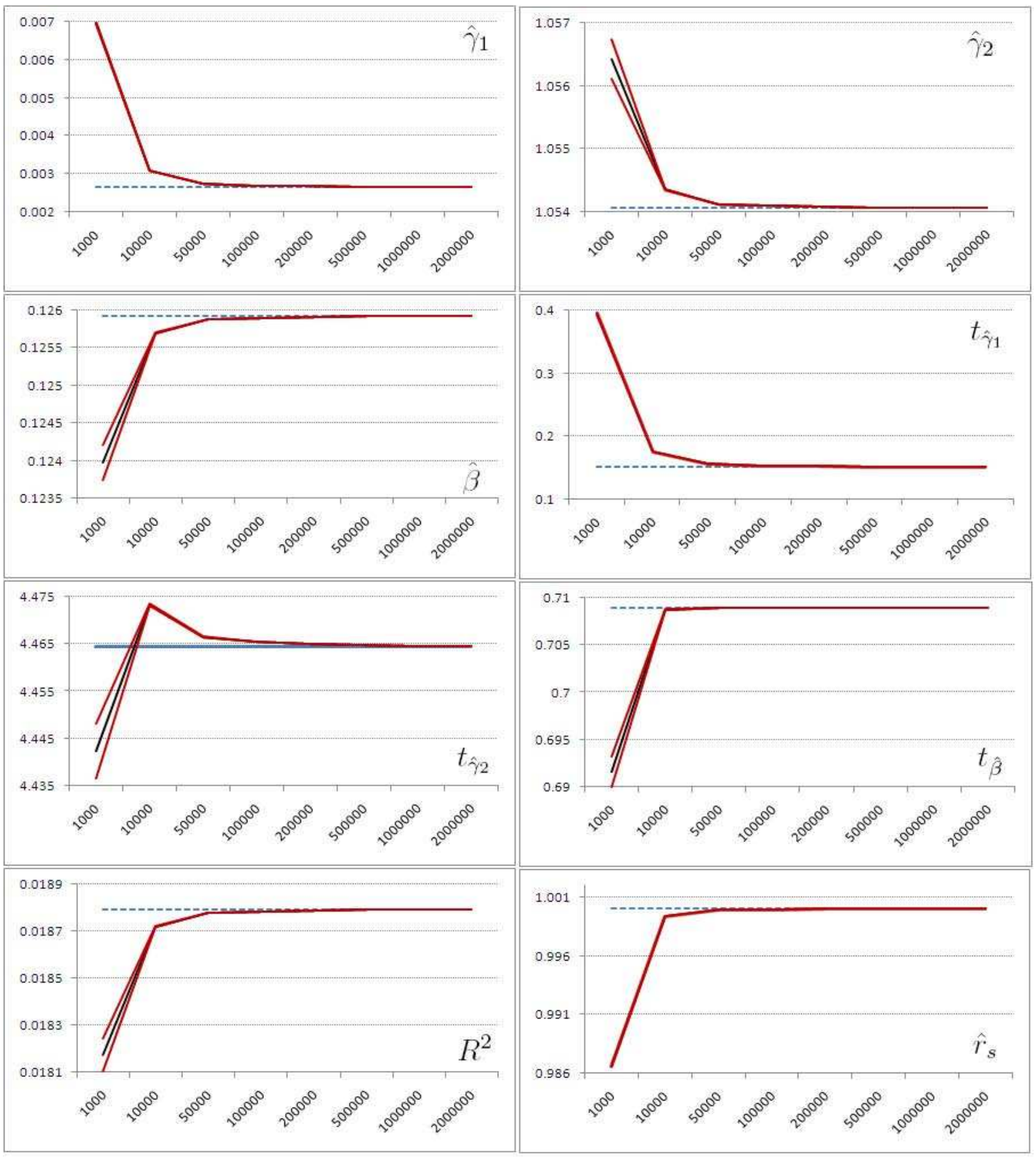


Table 3: Sensitivity Analysis: Starting Values

This table presents the simulation results for the OLS statistics convergence under a structural break in mean with different starting values. Panel A contains simulation results for the regression with structural breaks in the mean coefficients of the *AR(1)* processes. Two *independent AR(1)* processes, defined in Eq. (4.3), are simulated, given $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$. Number of observations in each simulated time series is 5,000,000. Panel B contains simulation results for the regression with structural breaks in trends. Two *independent* $AR(1)$ processes, defined in Eq. (4.4), are simulated, given $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $\mu_x = 0.2$, $\mu_{x,b} = 0.1$, $\mu_y = 0.3$, $\mu_{y,b} = 0.05$. Number of observations in each simulated time series is 2,000,000. In both cases, the OLS coefficients, their t-statistics, and the serial correlations of regression residuals are estimated from these simulated data. These procedures are repeated 1,000 times to generate standard errors of the mean of estimates and their 95% confidence bounds. Standard errors are given in parentheses.

	Panel A: Structural Break in Mean					*Panel B: Structural Break in Trend*							
	$\hat{\beta}$	$\hat{\gamma}$	$\hat{t}_\beta$	$\hat{t}_\gamma$	$\hat{r}_s, s=1$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}$	$\hat{t}_{\gamma_1}$	$\hat{t}_{\gamma_2}$	$\hat{t}_\beta$	$\hat{R}^2$	$\hat{r}_s, s=1$
$x_0 = 1$ $y_0 = 1$	0.506452 ($1.12\cdot10^{-4}$)	0.364529 ($3.02\cdot10^{-5}$)	0.105597 ($2.36\cdot10^{-5}$)	0.320451 ($2.83\cdot10^{-5}$)	0.943189 ($2.61\cdot10^{-6}$)	0.00266751 ($1.33\cdot10^{-10}$)	1.054074 ($1.43\cdot10^{-9}$)	0.125925 ($1.56\cdot10^{-9}$)	0.151765 ($6.87\cdot10^{-9}$)	4.464566 ($3.06\cdot10^{-8}$)	0.708948 ($8.14\cdot10^{-9}$)	0.0187903 ($3.99\cdot10^{-10}$)	0.999998 ($5.09\cdot10^{-12}$)
$x_0 = 1$ $y_0 = 100$	0.506833 ($1.12\cdot10^{-4}$)	0.364390 ($3.04\cdot10^{-5}$)	0.105681 ($2.34\cdot10^{-5}$)	0.320373 ($2.85\cdot10^{-5}$)	0.943176 ($2.63\cdot10^{-6}$)	0.00266877 ($1.39\cdot10^{-10}$)	1.054075 ($1.38\cdot10^{-9}$)	0.125924 ($1.41\cdot10^{-9}$)	0.151838 ($8.14\cdot10^{-9}$)	4.464596 ($3.46\cdot10^{-8}$)	0.708948 ($9.70\cdot10^{-9}$)	0.0187900 ($4.23\cdot10^{-10}$)	0.999997 ($6.14\cdot10^{-12}$)
$x_0 = 100$ $y_0 = 1$	0.506237 ($1.07\cdot10^{-4}$)	0.364523 ($2.97\cdot10^{-5}$)	0.105533 ($2.23\cdot10^{-5}$)	0.320399 ($2.82\cdot10^{-5}$)	0.943070 ($2.69\cdot10^{-6}$)	0.00266878 ($1.34\cdot10^{-10}$)	1.054075 ($1.90\cdot10^{-9}$)	0.125924 ($1.37\cdot10^{-9}$)	0.151838 ($7.69\cdot10^{-9}$)	4.464596 ($3.72\cdot10^{-6}$)	0.708948 ($9.21\cdot10^{-9}$)	0.0187900 ($4.06\cdot10^{-10}$)	0.999997 ($6.12\cdot10^{-12}$)
$x_0 = 100$ $y_0 = 100$	0.505982 ($1.11\cdot10^{-4}$)	0.364694 ($3.02\cdot10^{-5}$)	0.105502 ($2.34\cdot10^{-5}$)	0.320634 ($1.28\cdot10^{-5}$)	0.943140 ($2.59\cdot10^{-6}$)	0.00266878 ($1.31\cdot10^{-10}$)	1.054075 ($1.15\cdot10^{-9}$)	0.125924 ($1.36\cdot10^{-9}$)	0.151838 ($7.51\cdot10^{-9}$)	4.464596 ($3.39\cdot10^{-6}$)	0.708948 ($9.23\cdot10^{-9}$)	0.0187900 ($4.04\cdot10^{-10}$)	0.999997 ($5.98\cdot10^{-12}$)
Limit Value	**0.506329**	**0.364557**	**0.105564**	**0.320468**	**0.943194**	**0.00266667**	**1.054070**	**0.125926**	**0.151717**	**4.464550**	**0.708949**	**0.0187905**	**1.00000**

Table 4: Sensitivity Analysis: Structural Break Points

This table presents the simulation results for the OLS statistics convergence under a structural break in mean with different breaking point. Panel A contains simulation results for the regression with structural breaks in the mean coefficients of the *AR(1)* processes. Two *independent AR(1)* processes, defined in Eq. (4.3), are simulated, given $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $x_0 = 10$ and $y_0 = 10$. Number of observations in each simulated time series is 5,000,000. Panel B contains simulation results for the regression with structural breaks in trends. Two *independent* $AR(1)$ processes, defined in Eq. (4.4), are simulated, given $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $\mu_x = 0.2$, $\mu_{x,b} = 0.1$, $\mu_y = 0.3$, $\mu_{y,b} = 0.05$, $x_0 = 10$ and $y_0 = 10$. Number of observations in each simulated time series is 2,000,000. In both cases, the OLS coefficients, their t-statistics, and the serial correlations of regression residuals are estimated from these simulated data. These procedures are repeated 1,000 times to generate standard errors of the mean of estimates and their 95% confidence bounds. Standard errors are given in parentheses.

	Panel A: Structural Break in Mean					*Panel B: Structural Break in Trend*							
	$\hat{\gamma}$	$\hat{\beta}$	$\hat{t}_\gamma$	$\hat{t}_\beta$	$\hat{r}_s, s=1$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\beta}$	$\hat{t}_{\gamma_1}$	$\hat{t}_{\gamma_2}$	$\hat{t}_\beta$	$\hat{R}^2$	$\hat{r}_s, s=1$
$\tau_x = 0.05$, $\tau_y = 0.1$	5.043801 ($2.01 \cdot 10^{-4}$)	0.453884 ($4.15 \cdot 10^{-5}$)	0.725273 ($3.52 \cdot 10^{-5}$)	0.336147 ($3.46 \cdot 10^{-5}$)	0.905105 ($5.62 \cdot 10^{-6}$)	0.00426491 ($5.78 \cdot 10^{-10}$)	0.037291 ($3.70 \cdot 10^{-8}$)	0.906930 ($2.48 \cdot 10^{-8}$)	0.354447 ($5.00 \cdot 10^{-8}$)	0.048341 ($4.82 \cdot 10^{-8}$)	1.759291 ($9.32 \cdot 10^{-8}$)	0.947216 ($5.17 \cdot 10^{-8}$)	0.999999 ($4.04 \cdot 10^{-12}$)
Lim.Value	**5.043780**	**0.453940**	**0.725296**	**0.336208**	**0.905112**	**0.00426316**	**0.037256**	**0.906954**	**0.354288**	**0.0482933**	**1.75924**	**0.947266**	**1.00000**
$\tau_x = 0.05$, $\tau_y = 0.9$	0.560310 ($2.45 \cdot 10^{-4}$)	0.050442 ($4.97 \cdot 10^{-4}$)	0.076428 ($3.33 \cdot 10^{-5}$)	0.035436 ($3.49 \cdot 10^{-5}$)	0.928703 ($2.86 \cdot 10^{-6}$)	0.00089686 ($5.90 \cdot 10^{-10}$)	1.026891 ($3.88 \cdot 10^{-9}$)	0.119427 ($2.59 \cdot 10^{-9}$)	0.037078 ($2.44 \cdot 10^{-8}$)	0.662192 ($2.96 \cdot 10^{-8}$)	0.115243 ($2.52 \cdot 10^{-9}$)	0.021972 ($9.56 \cdot 10^{-9}$)	0.999993 ($1.92 \cdot 10^{-11}$)
Lim.Value	**0.560420**	**0.050437**	**0.076434**	**0.035431**	**0.928712**	**0.00089473**	**1.026880**	**0.119434**	**0.036988**	**0.662142**	**0.115242**	**0.021974**	**1.00000**
$\tau_x = 0.95$, $\tau_y = 0.1$	7.187114 ($6.14 \cdot 10^{-5}$)	0.050352 ($4.78 \cdot 10^{-5}$)	2.515653 ($5.19 \cdot 10^{-5}$)	0.035371 ($3.36 \cdot 10^{-5}$)	0.928695 ($2.85 \cdot 10^{-6}$)	-0.0160562 ($9.07 \cdot 10^{-11}$)	1.274526 ($2.45 \cdot 10^{-8}$)	0.119440 ($2.45 \cdot 10^{-9}$)	-2.542369 ($4.19 \cdot 10^{-8}$)	1.225199 ($3.76 \cdot 10^{-8}$)	0.115240 ($2.36 \cdot 10^{-8}$)	0.007390 ($3.03 \cdot 10^{-9}$)	0.999999 ($0.78 \cdot 10^{-12}$)
Lim.Value	**7.187390**	**0.050437**	**2.515740**	**0.035431**	**0.928712**	**-0.0160582**	**1.274530**	**0.119434**	**-2.542680**	**1.225290**	**0.115242**	**0.007390**	**1.00000**
$\tau_x = 0.95$, $\tau_y = 0.9$	0.686579 ($6.57 \cdot 10^{-5}$)	0.453898 ($4.20 \cdot 10^{-5}$)	0.253382 ($2.49 \cdot 10^{-5}$)	0.336183 ($3.50 \cdot 10^{-5}$)	0.905103 ($5.61 \cdot 10^{-6}$)	-0.0007212 ($1.38 \cdot 10^{-10}$)	0.295334 ($2.47 \cdot 10^{-8}$)	0.906977 ($2.46 \cdot 10^{-8}$)	-0.229584 ($4.55 \cdot 10^{-8}$)	0.570736 ($5.55 \cdot 10^{-8}$)	1.759187 ($9.75 \cdot 10^{-9}$)	0.570086 ($3.10 \cdot 10^{-8}$)	0.999997 ($2.88 \cdot 10^{-11}$)
Lim.Value	**0.686515**	**0.453940**	**0.253353**	**0.336208**	**0.905112**	**-0.0007229**	**0.295358**	**0.906954**	**-0.230137**	**0.570813**	**1.759240**	**0.570058**	**1.00000**

Appendices: Proofs

Proofs of Theorems 1 and 2 are based on the following lemma:

Lemma 3. *Suppose that the stationary process* $\boldsymbol{\xi}_t = (u_t, v_t)$ *is strongly mixing such that* $\alpha(k) = \mathcal{O}\left(k^{-\frac{r}{r-1}-\varepsilon}\right)$ *for some* $r > 1$ *and* $\varepsilon > 0$. *Let* $\eta_t = g(\boldsymbol{\xi}_t, \boldsymbol{\xi}_{t-1}, \ldots, \boldsymbol{\xi}_{t-\tau})$ *be a measurable function, for some finite* τ. *If* $E\,|\eta_t|^{r+\delta} < \infty$ *for some generic constant,* $\delta > 0$, *then*

$$T^{-1} \sum_{t=\tau+1}^{T} \{\eta_t - E[\eta_t]\} \overset{a.s.}{\Longrightarrow} 0.$$

Proof. Theorem 14.1 in Davidson [2002] asserts that the process, η_t, is also α-mixing of size $\frac{r}{r-1}$. The lemma immediately follows from an application of McLeish's [1975] SLLN for mixingales. □

Proof of Lemma 1

First, we shall verify Eq. (2.2). We shall note at the outset that roman numbers, used to indicate mathematical expressions, are specific to each equation of Lemma 1.

$$\begin{aligned}
\frac{\sum_1^T X_t Y_t}{T} &= T^{-1}\left(\sum_{1}^{[T\tau^{(x)}]} + \sum_{[T\tau^{(x)}]+1}^{[T\tau^{(y)}]} + \sum_{[T\tau^{(y)}]+1}^{T}\right) X_t Y_t \\
&= T^{-1} \sum_{1}^{[T\tau^{(x)}]} (A^*(1)u_t + \Delta u_t^*)(B^*(1)v_t + \Delta v_t^*) \\
&+ T^{-1} \sum_{[T\tau^{(x)}]+1}^{[T\tau^{(y)}]} (A^*(1)(c^{(x)} + u_t) + \Delta u_t^*)(B^*(1)v_t + \Delta v_t^*) \\
&+ T^{-1} \sum_{[T\tau^{(y)}]+1}^{T} \left(A^*(1)(c^{(x)} + u_t) + \Delta u_t^*\right)\left(B^*(1)(c^{(y)} + v_t) + \Delta v_t^*\right) \\
&= I + II + III.
\end{aligned}$$

An application of Hölder's inequality yields, for every t and s, $E[|u_t v_s|^{r+\delta}] \le E[|u_0|^{2(r+\delta)}]E[|v_0|^{2(r+\delta)}] < \infty$, where the last inequality is due to Assumption 1. Since $(A^*(1)u_t + \Delta u_t^*)(B^*(1)v_t + \Delta v_t^*)$ is some linear function of $\boldsymbol{\xi}_t$ and $\boldsymbol{\xi}_{t-1}$, Lemma 3 yields $I \overset{a.s.}{\Longrightarrow} 0$. Similarly, we can show that $II \overset{a.s.}{\Longrightarrow} 0$. Next, the last term

$$\begin{aligned} III &= T^{-1}\sum_{[T\tau^{(y)}]+1}^{T} A^*(1)B^*(1)c^{(x)}c^{(y)} + T^{-1}\sum_{[T\tau^{(y)}]+1}^{T} A^*(1)c^{(x)}\left(B^*(1)v_t+\Delta v_t^*\right) \\ &+ T^{-1}\sum_{[T\tau^{(y)}]+1}^{T} B^*(1)c^{(y)}\left(A^*(1)u_t+\Delta u_t^*\right) \end{aligned}$$

converges to $A^*(1)B^*(1)c^{(x)}c^{(y)}(1-\tau^{(y)})$ because the last two terms converge to zero by applying Lemma 3.

Eqs. (2.3) and (2.4) can be easily proved by using Lemma 3. Now, we shall prove Eq. (2.5). Some preliminary algebra yields

$$\begin{aligned} T^{-1}\sum_{t=1}^{T} X_t^2 &= T^{-1}\sum_{t=1}^{T}\Big\{A^{*2}(1){c^{(x)}}^2\mathbf{1}\{t>[T\tau^{(x)}]\} + A^{*2}(1)u_t^2 + \Delta^2 u_t^* \\ &+ 2A^{*2}(1)c^{(x)}\mathbf{1}\{t>[T\tau^{(x)}]\}u_t + 2A^*(1)\Delta u_t^*\mathbf{1}\{t>[T\tau^{(x)}]\} + 2A^*(1)u_t\Delta u_t^*\Big\} \\ &= I+II+III+IV+V+VI. \end{aligned}$$

One can verify that $I \Longrightarrow A^{*2}(1){c^{(x)}}^2 \lim_{T\longrightarrow\infty} T^{-1}\sum_{t=1}^{T}\mathbf{1}\{t>[T\tau^{(x)}]\} = A^{*2}(1){c^{(x)}}^2(1-\tau^{(x)})$. An application of Lemma 3 yields $II \overset{a.s.}{\Longrightarrow} A^{*2}(1)\sigma_u^2$; $III \overset{a.s.}{\Longrightarrow} \sigma^2_{\Delta u^*}$; $IV \overset{a.s.}{\Longrightarrow} 0$; $V \overset{a.s.}{\Longrightarrow} 0$; and $VI \overset{a.s.}{\Longrightarrow} 2A^*(1)\sigma_{u\Delta u^*}$. Hence, Eq. (2.5) follows. Eq. (2.6) can be similarly proved. To prove Eq. (2.7), let us rewrite

$$T^{-1}\sum_{t=s+1}^{T}\left(Y_t-\overline{Y}\right)\left(Y_{t-s}-\overline{Y}\right) = T^{-1}\sum_{s+1}^{T} Y_tY_{t-s} - \overline{Y}T^{-1}\left(\sum_{s+1}^{T} Y_t + \sum_{s+1}^{T} Y_{t-s}\right) + \overline{Y}^2.$$

Some algebra yields

$$\begin{aligned} T^{-1}\sum_{t=s+1}^{T} Y_tY_{t-s} &= B^{*2}(1){c^{(y)}}^2 T^{-1}\sum_{s+1}^{T}\mathbf{1}\{t>[T\tau^{(x)}]\}\mathbf{1}\{t-s>[T\tau^{(x)}]\} \\ &+ B^{*2}(1)T^{-1}\sum_{s+1}^{T} v_tv_{t-s} + T^{-1}\sum_{s+1}^{T}\Delta v_t^*\Delta v_{t-s}^* \\ &+ B^*(1)\left\{T^{-1}\sum_{s+1}^{T} v_t\Delta v_{t-s}^* + T^{-1}\sum_{s+1}^{T} v_{t-s}\Delta v_t^*\right\} \\ &+ B^*(1)c^{(y)}\left\{T^{-1}\sum_{s+1}^{T} B^*(L)v_t + T^{-1}\sum_{s+1}^{T} B^*(L)v_{t-s}\right\} \\ &= I+II+III+IV+V. \end{aligned}$$

One can verify that $I \Longrightarrow B^{*2}(1){c^{(y)}}^2(1-\tau^{(y)})$; in view of Lemma 3, under Assumption 1, $II \overset{a.s.}{\Longrightarrow} B^{*2}(1)\sigma_{v_0v_s}$; $III \overset{a.s.}{\Longrightarrow} \sigma_{\Delta v_0^*\Delta v_s^*}$; $IV \overset{a.s.}{\Longrightarrow} B^*(1)[\sigma_{v_0\Delta v_s^*} + \sigma_{v_s\Delta v_0^*}]$; and $V \overset{a.s.}{\Longrightarrow} 0$. Moreover, in view of Eq. (2.4), we obtain

$$\overline{Y}\left(T^{-1}\sum_{s+1}^{T}Y_t+T^{-1}\sum_{s+1}^{T}Y_{t-s}\right)\overset{a.s.}{\Longrightarrow}2(1-\tau^{(y)})^2B^{*2}(1)c^{(y)2}.$$

Hence Eq. (2.7) follows. Eq. (2.8) can be similarly proved. To prove Eq. (2.9), note that

$$T^{-1}\sum_{t=s+1}^{T}\left(Y_t-\overline{Y}\right)\left(X_{t-s}-\overline{X}\right)=T^{-1}\sum_{t=s+1}^{T}X_{t-s}Y_t-\overline{X}\times\overline{Y}=I-II.$$

One can show that

$$\begin{aligned} I &= A^*(1)B^*(1)c^{(x)}c^{(y)}T^{-1}\sum_{t=s+1}^{T}\mathbf{1}(t-s>[T\tau^{(x)}])\mathbf{1}(t>[T\tau^{(y)}]) \\ &+ B^*(1)c^{(y)}T^{-1}\sum_{t=s+1}^{T}\mathbf{1}(t>[T\tau^{(y)}])A^*(L)u_{t-s} \\ &+ A^*(1)c^{(x)}T^{-1}\sum_{t=s+1}^{T}\mathbf{1}(t-s>[T\tau^{(x)}])B^*(L)v_t=I.1+I.2+I.3. \end{aligned}$$

One can verify that $I.1\Longrightarrow A^*(1)B^*(1)c^{(x)}c^{(y)}(1-\tau^{(y)})$; in view of Lemma 3, under Assumption 1, $I.2\overset{a.s.}{\Longrightarrow}0$ and $I.3\overset{a.s.}{\Longrightarrow}0$. Moreover, Eqs. (2.3) and (2.4) yield

$$II\overset{a.s.}{\Longrightarrow}(1-\tau^{(x)})(1-\tau^{(y)})c^{(x)}c^{(y)}A^*(1)B^*(1).$$

Hence, Eq. (2.9) follows. Eq. (2.10) can be similarly proved.

Proof of Theorem 1

The OLS statistics are given by

$$\widehat{\beta} = \frac{T^{-1}\sum_1^T X_tY_t-\overline{X}\,\overline{Y}}{T^{-1}\sum_1^T X_t^2-\overline{X}^2}; \tag{5.1}$$

$$\widehat{\gamma} = \overline{Y}-\hat{\beta}\overline{X}; \tag{5.2}$$

$$\widehat{s}^2 = T^{-1}\sum_1^T(Y_t-\overline{Y})^2-\hat{\beta}^2T^{-1}\sum_1^T(X_t-\overline{X})^2; \tag{5.3}$$

$$T^{-1/2}t_{\widehat{\beta}} = \frac{\widehat{\beta}}{\widehat{s}\left(T^{-1}\sum_1^T(X_t-\overline{X})^2\right)^{-1/2}}; \tag{5.4}$$

$$T^{-1/2}t_{\widehat{\gamma}} = \frac{\widehat{\gamma}}{\widehat{s}}\frac{\left(T^{-1}\sum_1^T\left(X_t-\overline{X}\right)^2\right)^{1/2}}{\left(T^{-1}\sum_1^T X_t^2\right)^{1/2}}; \tag{5.5}$$

$$\widehat{r}_s = \frac{T^{-1}\sum_{t=1+s}^T\widehat{w}_t\widehat{w}_{t-s}}{T^{-1}\sum_1^T\widehat{w}_t^2}; \tag{5.6}$$

where

$$\begin{aligned} T^{-1}\sum_{t=s+1}^{T}\widehat{w}_t\widehat{w}_{t-s} &= T^{-1}\sum_{t=s+1}^{T}Y_tY_{t-s}-\widehat{\gamma}T^{-1}\sum_{t=s+1}^{T}(Y_t+Y_{t-s}) \\ &- \widehat{\beta}T^{-1}\sum_{t=s+1}^{T}(Y_tX_{t-s}+Y_{t-s}X_t) \\ &+ \widehat{\beta}\widehat{\gamma}T^{-1}\sum_{t=s+1}^{T}(X_t+X_{t-s})+\widehat{\gamma}^2+\widehat{\beta}^2T^{-1}\sum_{t=s+1}^{T}X_tX_{t-s}; \end{aligned}$$

and

$$\begin{aligned} T^{-1}\sum_{t=1}^{T}\widehat{w}_t^2 &= T^{-1}\sum_{t=1}^{T}Y_t^2+\widehat{\gamma}+\widehat{\beta}^2T^{-1}\sum_{t=1}^{T}X_t^2 \\ &- 2\widehat{\gamma}T^{-1}\sum_{t=1}^{T}Y_t-2\widehat{\beta}T^{-1}\sum_{t=1}^{T}X_tY_t+2\widehat{\gamma}\widehat{\beta}T^{-1}\sum_{t=1}^{T}X_t. \end{aligned}$$

Using the formula: $T^{-1}\sum_{t=1}^{T}(\xi_t-\overline{\xi})^2 = T^{-1}\sum_{t=1}^{T}\xi_t^2-\overline{\xi}^2$ and

$$\begin{aligned} T^{-1}\sum_{t=s+1}^{T}Y_tY_{t-s} &= T^{-1}\sum_{t=s+1}^{T}(Y_t-\overline{Y})(Y_{t-s}-\overline{Y})+\overline{Y}T^{-1}\sum_{t=s+1}^{T}(Y_t+Y_{t-s})+\frac{T-s}{T}\overline{Y}^2 \\ T^{-1}\sum_{t=s+1}^{T}X_{t-s}Y_t &= T^{-1}\sum_{t=s+1}^{T}(Y_t-\overline{Y})(X_{t-s}-\overline{X})+\overline{X}\,\overline{Y} \\ T^{-1}\sum_{t=s+1}^{T}X_tY_{t-s} &= T^{-1}\sum_{t=s+1}^{T}(X_t-\overline{X})(Y_{t-s}-\overline{Y})+\overline{X}\,\overline{Y}. \end{aligned}$$

Lemma 1 yields the main results.

Proof of Lemma 2

First, we shall prove Eq. (3.3). We shall note at the outset that roman numbers, used to indicate mathematical expressions, are specific to each equation of Lemma 2.

$$\begin{aligned} T^{-2}\sum_{t=1}^{T}X_t &= T^{-1}\left(A^*(1)c^{(x)}+\mu^{(x)}A^{\diamond}(1)\right)+\mu^{(x)}A^*(1)T^{-2}\sum_{t=1}^{T}t \\ &+ \mu_1^{(x)}A^*(1)T^{-2}\sum_{t=1}^{T}(t-[T\tau^{(x)}])\mathbf{1}\{t>[T\tau^{(x)}]\} \\ &+ \mu_1^{(x)}A^{\diamond}(1)T^{-2}\sum_{t=1}^{T}\mathbf{1}\{t>[T\tau^{(x)}]\}+T^{-2}\sum_{t=1}^{T}A^*(L)u_t \\ &= I+II+III+IV+V. \end{aligned}$$

One can see that

$$\begin{aligned}
&I \Longrightarrow 0;\\
&II \Longrightarrow \mu^{(x)}A^*(1)\int_0^1 sds;\\
&III \Longrightarrow \mu_1^{(x)}A^*(1)\int_{\tau^{(x)}}^1 (s-\tau^{(x)})ds;\\
&IV \longrightarrow 0
\end{aligned}$$

and an application of Lemma 3 yields $V = o_{a.s.}(T^{-1})$.

To prove Eq. (3.4), we shall note that

$$\begin{aligned}
T^{-3}\sum_{t=1}^{T} X_t^2 &= \{\mu^{(x)}A^*(1)\}^2 T^{-3}\sum_{t=1}^{T} t^2\\
&+ \left(\{\mu_1^{(x)}A^*(1)\}\right)^2 T^{-3}\sum_{t=1}^{T}\left(t-[T\tau^{(x)}]\right)^2 \mathbf{1}\{t>[T\tau^{(x)}]\}\\
&+ 2A^{*2}(1)\mu^{(x)}\mu_1^{(x)}T^{-3}\sum_{t=1}^{T} t(t-[T\tau^{(x)}])\mathbf{1}\{t>[T\tau^{(x)}]\}\\
&+ T^{-3}\sum_{t=1}^{T} A^{*2}(L)u_t^2 + \mathscr{O}\left(T^{-1}\right) = I+II+III+IV.
\end{aligned}$$

One can verify that

$$\begin{aligned}
&I \Longrightarrow \{\mu^{(x)}A^*(1)\}^2\int_0^1 s^2 ds;\\
&II \Longrightarrow \{\mu_1^{(x)}A^*(1)\}^2\int_{\tau^{(x)}}^1 (s-\tau^{(x)})^2 ds;\\
&III \Longrightarrow 2A^{*2}(1)\mu^{(x)}\mu_1^{(x)}\int_{\tau^{(x)}}^1 s\left(s-\tau^{(x)}\right)ds.
\end{aligned}$$

In view of Lemma 3, under Assumption 1, $IV = o_{a.s.}(T^{-2})$. Some algebra manipulations yields Eq. (3.4).

Let us prove Eq. (3.9). (Eqs. (3.5), (3.6), (3.7), (3.8), (3.10) and (3.11) can be similarly proved.) To avoid any cumbersome mathematical expression, we shall summarize only the terms of order $\mathscr{O}(1)$ as follows:

$$\begin{aligned}
T^{-3}\sum_{t=1}^{T} X_tY_t &= T^{-3}\sum_{t=1}^{T}\Big[\left\{\mu^{(x)}A^*(1)t+\mu_1^{(x)}A^*(1)\left(t-[T\tau^{(x)}]\right)\mathbf{1}\left(t>[T\tau^{(x)}]\right)\right\}\\
&\quad\left\{\mu^{(y)}B^*(1)t+\mu_1^{(x)}B^*(1)\left(t-[T\tau^{(y)}]\right)\mathbf{1}\left(t>[T\tau^{(y)}]\right)\right\}\Big]\\
&+ T^{-3}\sum_{t=1}^{T} A^*(L)B^*(L)u_tv_t + \mathscr{O}(T^{-1}) = I+II.
\end{aligned}$$

One can verify that $I \Longrightarrow \int_0^1\left\{\mu^{(x)}A^*(1)s+\mu_1^{(x)}A^*(1)\left(s-\tau^{(x)}\right)\mathbf{1}\{s>\tau^{(x)}\}\right\}$ $\left\{\mu^{(y)}B^*(1)s+\mu_1^{(y)}B^*(1)\left(s-\tau^{(y)}\right)\mathbf{1}\{s>\tau^{(y)}\}\right\}ds$ and $II \overset{a.s.}{\Longrightarrow} 0$ in view of Lemma 3. Some algebra manipulations yields Eq. (2.9).

Proof of Theorem 2

Eq. (3.2) is equivalent to

$$\begin{aligned}
\begin{bmatrix} T^{-1}\hat{\gamma}_1 \\ \hat{\beta} \\ \hat{\gamma}_2 \end{bmatrix} &= \begin{bmatrix} T & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\gamma}_1 \\ \hat{\beta} \\ \hat{\gamma}_2 \end{bmatrix} \\
&= \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & T^{-1} & 0 \\ 0 & 0 & T^{-1} \end{bmatrix} \begin{bmatrix} T & \sum_1^T X_t & \sum_1^T t \\ \sum_1^T X_t & \sum_1^T X_t^2 & \sum_1^T tX_t \\ \sum_1^T t & \sum_1^T tX_t & \sum_1^T t^2 \end{bmatrix} \right. \\
&\quad \left. \begin{bmatrix} T^{-1} & 0 & 0 \\ 0 & T^{-2} & 0 \\ 0 & 0 & T^{-2} \end{bmatrix} \right\}^{-1} \begin{bmatrix} T^{-2} & 0 & 0 \\ 0 & T^{-3} & 0 \\ 0 & 0 & T^{-3} \end{bmatrix} \begin{bmatrix} \sum_1^T Y_t \\ \sum_1^T X_t Y_t \\ \sum_1^T tY_t \end{bmatrix} \\
&= \begin{bmatrix} 1 & T^{-2}\sum_1^T X_t & T^{-2}\sum_1^T t \\ T^{-2}\sum_1^T X_t & T^{-3}\sum_1^T X_t^2 & T^{-3}\sum_1^T tX_t \\ T^{-2}\sum_1^T t & T^{-3}\sum_1^T tX_t & T^{-3}\sum_1^T t^2 \end{bmatrix}^{-1} \begin{bmatrix} T^{-2}\sum_1^T Y_t \\ T^{-3}\sum_1^T X_t Y_t \\ T^{-3}\sum_1^T tY_t \end{bmatrix}.
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
\frac{\mathrm{var}(\hat{\gamma}_1)}{T} &= \{T^{-2}\hat{s}^2\}\mathbb{I}_{(1)} \begin{bmatrix} 1 & T^{-2}\sum_1^T X_t & T^{-2}\sum_1^T t \\ T^{-2}\sum_1^T X_t & T^{-3}\sum_1^T X_t^2 & T^{-3}\sum_1^T tx_t \\ T^{-2}\sum_1^T t & T^{-3}\sum_1^T tX_t & T^{-3}\sum_1^T t^2 \end{bmatrix}^{-1} \mathbb{I}'_{(1)}; \\
T\mathrm{var}(\widehat{\beta}) &= \{T^{-2}\hat{s}^2\}\mathbb{I}_{(2)} \begin{bmatrix} 1 & T^{-2}\sum_1^T X_t & T^{-2}\sum_1^T t \\ T^{-2}\sum_1^T X_t & T^{-3}\sum_1^T X_t^2 & T^{-3}\sum_1^T tx_t \\ T^{-2}\sum_1^T t & T^{-3}\sum_1^T tX_t & T^{-3}\sum_1^T t^2 \end{bmatrix}^{-1} \mathbb{I}'_{(2)}; \\
T\mathrm{var}(\widehat{\gamma}_2) &= \{T^{-2}\hat{s}^2\}\mathbb{I}_{(3)} \begin{bmatrix} 1 & T^{-2}\sum_1^T X_t & T^{-2}\sum_1^T t \\ T^{-2}\sum_1^T X_t & T^{-3}\sum_1^T X_t^2 & T^{-3}\sum_1^T tX_t \\ T^{-2}\sum_1^T t & T^{-3}\sum_1^T tX_t & T^{-3}\sum_1^T t^2 \end{bmatrix}^{-1} \mathbb{I}'_{(3)}; \\
\frac{\sum_{t=1+s}^T \hat{w}_{t-s}\hat{w}_t}{T^3} &= T^{-3}\left\{ \sum_{t=1+s}^T Y_t Y_{t-s} + \hat{\beta}^2 \sum_{t=1+s}^T X_t X_{t-s} - \hat{\beta} \sum_{t=1+s}^T (X_t Y_{t-s} + X_{t-s} Y_t) \right. \\
&- \hat{\gamma}_2 \sum_{1+s}^T (tY_{t+s} + (t-s)Y_t) - \hat{\gamma}_1 \sum_{t=1+s}^T (Y_t + Y_{t-s}) + \widehat{\gamma}_1 \widehat{\beta} \sum_{1+s}^T (X_t + X_{t-s}) \\
&+ \left. \hat{\gamma}_2 \hat{\beta} \sum_{1+s}^T ((t-s)X_t + tX_{t-s}) + \hat{\gamma}_1 \hat{\gamma}_2 \sum_{1+s}^T (2t-s) + T\hat{\gamma}_1^2 \right\}; \\
T^{-2}\hat{s}^2 &= T^{-3}\sum_1^T \hat{w}_t^2 = T^{-3}\left\{ \sum_1^T Y_t^2 + \widehat{\beta}^2 \sum_1^T X_t^2 + \widehat{\gamma}_1^2 T + \widehat{\gamma}_2^2 \sum_1^T t^2 - 2\widehat{\gamma}_1 \sum_1^T Y_t \right. \\
&- \left. 2\widehat{\gamma}_2 \sum_1^T tY_t - 2\widehat{\beta} \sum_1^T X_t Y_t + 2\widehat{\gamma}_1 \widehat{\gamma}_2 \sum_1^T t + 2\widehat{\beta}\widehat{\gamma}_1 \sum_1^T X_t + 2\widehat{\gamma}_2 \widehat{\beta} \sum_1^T tX_t \right\}.
\end{aligned}$$

By applying Lemma 2, one can immediately derive the limits of individual terms in the above equations. Theorem 2 has been proved.

References

Davidson, J. (2002): *Stochastic Limit Theory*, Oxford University Press, second edition.

Ferson, W. E., S. Sarkissian, and T. T. Simin (2003): "Spurious regressions in financial economics," *Journal of Finance*, LVIII, 1393–1413.

Gorodetskii, V. V. (1977): "On the strong mixing property for linear sequences," *Teor. Veroyatnost. i Primenen.*, 22, 421–423.

Granger, C. W. J., H. Hyung, and Y. Jeon (2001): "Spurious regressions with stationary series," *Applied Economics*, 33, 899–904.

Granger, C. W. J. and P. Newbold (1974): "Spurious regressions in econometrics," *Journal of Econometrics*, 74, 111–120.

Marmol, F. (1995): "Spurious regressions between i(d) processes," *Journal of Time Series Analysis*, 16, 313–321.

McLeish, D. L. (1975): "A maximal inequality and dependent strong laws," *Annals of Probability*, 3, 829–839.

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

Phillips, P. C. B. (1998): "New tools for understanding spurious regressions," *Econometrica*, 66, 1299–1325.

Rosenblatt, M. (1956): "A central limit theorem and a strong mixing condition," *Procedure of the National Academy of Science USA*, 42, 43–47.

Tsay, W. J. (1999): "Spurious regression between i(1) processes with infinite variance errors," *Econometric Theory*, 15, 622–628.

Tsay, W. J. and C. F. Chung (1999): "The spurious regression of fractionally integrated processes," *Journal of Econometrics*, 96, 155–182.

Withers, C. S. (1981): "Conditions for linear processes to be strong-mixing," *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 57, 477–480.