WORKING PAPERS SERIES

WP09-04

Spurious Regressions of Stable AR(p) Processes with Structural Breaks

Ba Chu and Roman Kozhan

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

Ba Chu*

Carleton University

Roman Kozhan

Warwick Business School

Abstract

When a pair of *independent* series are 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 finds evidences that spurious regressions also arise in the regressions between stable *AR(p)* processes with structural breaks in the means and the trends. An intuition behind this is that structural breaks can increase the persistence levels in the processes (see, e.g., Granger and Hyung (2004)), which then lead to spurious regressions. These phenomena occur for general distributions and serial dependence of the innovation terms.

Key words: Spurious regression; Weak dependence; Structural breaks; Strictly stationary; Autoregressive processes

JEL Classification: C12, C13

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,

*Address correspondence to Ba Chu, Department of Economics, Carleton University, Ottawa K1S - 5B6, Canada; E-mail: **ba_chu@carleton.ca**. *Other author:* Roman Kozhan, Warwick Business School, University of Warwick, Coventry CV4 - 7AL, UK; E-mail: **Roman.Kozhan@wbs.ac.uk**.

1998) provide an elegant asymptotic framework that vindicates the simulation results. Applied economists become increasingly aware of this problem. For instance, Ferson et al. (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 et al. (2001)).

The goal of the current paper is to investigate the possible existence of spurious relations in a pair of *stable AR(p)* processes with *weakly dependent* innovations and structural breaks in the means and the trends. (It should be noted at this point that the term 'stable' has a specific meaning in probability theory – 'stable' random variable means a random variable with infinite variance. Here we use this term to mean that *AR(p)* processes, after being demeaned and detrended, become strictly stationary.) It is shown, using both simulations and asymptotic theory, that the strength of these spurious relationships varies a little with sample sizes. The proofs of our main results call for the need of developing new strong laws of large numbers (SLLN) for weakly dependent processes and their bivariate products. Therefore, the contribution of this paper is twofold. First, we prove general SLLNs with *minimum* necessary conditions. Second, we apply these SLLNs to derive the asymptotic properties of the well-known statistics for the bivariate regression models with *independent* response and covariate variables.

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. The framework with many break points can be immediately extended without much efforts. 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 with Structural Breaks in Mean

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

$$\begin{aligned} A(L)x_t &= \alpha_x \mathbf{1}\{t > [T\tau_x]\} + u_t, \\ B(L)y_t &= \alpha_y \mathbf{1}\{t > [T\tau_y]\} + v_t, \end{aligned} \tag{2.1}$$

where $A(L)$ and $B(L)$ are lag polynomials with roots outside the unit circle. Without any loss of generality, let us assume that the break points satisfy $\tau_x \leq \tau_y$; and the innovation terms are contemporaneously *independent* and fulfill Assumption 1 stated below.

Assumption 1. *The vector of stochastic sequences, $\boldsymbol{U}_t = (u_t, v_t)$, is strictly stationary with $E[\boldsymbol{U}_0] = 0$ and $E[|\boldsymbol{U}_0|^p] < \infty$ for any $p \geq 2$. Moreover, the mixing coefficient of $\boldsymbol{U}_t$, $\boldsymbol{\phi}_t$* or *$\boldsymbol{\alpha}_t$, satisfies the following* summability *condition:*

$$\sum_{t=1}^{\infty} \boldsymbol{\phi}_t^{\frac{p-1}{p}} < \infty \text{ or } \sum_{t=1}^{\infty} \boldsymbol{\alpha}_t^{\frac{p-2}{p}} < \infty, \text{ respectively.}$$

A typical example of stochastic sequences, satisfying Assumption 1, is a MA process $\sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, where ϵ_j are IID.

Eq. (2.1) can be rewritten as

$$\begin{aligned} x_t &= A^{-1}(L)\alpha_x \mathbf{1}\{t > [T\tau_x]\} + A^{-1}(L)u_t, \\ y_t &= B^{-1}(L)\alpha_y \mathbf{1}\{t > [T\tau_y]\} + B^{-1}(L)v_t. \end{aligned}$$

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

$$\begin{aligned} \hat{\beta} &= \frac{\sum_1^T y_t(x_t - \overline{x})}{\sum_1^T (x_t - \overline{x})^2}, \\ \hat{\gamma} &= \overline{y} - \hat{\beta}\overline{x}, \\ \hat{w}_t &= y_t - \hat{\gamma} - \hat{\beta}x_t. \end{aligned}$$

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

Theorem 1. *As the sample size T goes to infinity, we have*

$$\hat{\beta} \xrightarrow{a.s.} \frac{A^*(1)B^*(1)\alpha_x\alpha_y\tau_x(1-\tau_y)}{A^{*2}(1)\sigma_u^2 + \Omega_{\Delta u^*} + \tau_x(1-\tau_x)A^{*2}(1)\alpha_x^2 + 2A^*(1)\Omega_{u\Delta u^*}} = B.$$

$$\hat{\gamma} \xrightarrow{a.s.} (1-\tau_y)\alpha_y B^*(1) - B(1-\tau_x)\alpha_x A^*(1) = D.$$

$$T^{-1/2}t_{\hat{\beta}} = \frac{\hat{\beta}}{\hat{s}\left(T^{-1}\sum_{t=1}^T (x_t - \overline{x})^2\right)^{-1/2}} \xrightarrow{a.s.} \frac{B}{C},$$

where

$$\hat{s}^2 = T^{-1}\sum_1^T \left[(y_1-\overline{y})-\hat{\beta}(x_t-\overline{x})\right]^2 \quad \xrightarrow{a.s.} \quad B^*(1)^2\sigma_v^2+\sigma_{\Delta v^*}^2+\tau_y(1-\tau_y)\alpha_y^2 B^{*2}(1)+2B^*(1)\Omega_{v\Delta v^*}$$

$$- \quad \frac{\tau_x^2(1-\tau_y)^2\alpha_x^2\alpha_y^2\left[A^*(1)B^*(1)\right]^2}{A^{*2}(1)\sigma_u^2+\Omega_{\Delta u^*}+\tau_x(1-\tau_x)A^{*2}(1)\alpha_x^2+2A^*(1)\Omega_{u\Delta u^*}}$$

$$= \quad S$$

$$C = \left\{\frac{A^{*2}(1)\sigma_u^2+\Omega_{\Delta u^*}+\tau_x(1-\tau_x)A^{*2}(1)\alpha_x^2+2A^*(1)\Omega_{u\Delta u^*}}{S}\right\}^{1/2}.$$

$$T^{-1/2}t_{\hat{\gamma}} = \frac{\hat{\gamma}}{\hat{s}}\frac{\left(T^{-1}\sum_{t=1}^T(x_t-\overline{x})^2\right)^{1/2}}{\left(T^{-1}\sum_{t=1}^T x_t^2\right)^{1/2}} \xrightarrow{a.s.} \frac{D}{E},$$

with

$$E = \left\{\frac{S}{1-\frac{(1-\tau_x)^2\alpha_x A^*(1)}{A^{*2}(1)\sigma_u^2+\Omega_{\Delta u^*}+\tau_x(1-\tau_x)A^{*2}(1)\alpha_x^2+2A^*(1)\Omega_{u\Delta u^*}}}\right\}^{1/2}.$$

Finally, the sample autocorrelation of the residuals is given by

$$\hat{r}_s = \frac{\sum_{s+1}^T \hat{w}_t\hat{w}_{t-s}}{\sum_1^T \hat{w}_t^2} \xrightarrow{a.s.} \frac{F}{S},$$

where

$$\begin{aligned} F &= B^{*2}(1)\alpha_y^2\tau_y(1-\tau_y)+B^{*2}(1)E[v_0v_s]+\Omega^*_{\Delta v^*}+B^*(1)\big(\Omega_{1.v\Delta v^*}+\Omega_{2.v\Delta v^*}\big) \\ &+ B^2\left(A^{*2}(1)\alpha_x^2\tau_x(1-\tau_x)+B^{*2}(1)E[u_0u_s]+\Omega^*_{\Delta u^*}+A^*(1)\left(\Omega_{1.v\Delta v^*}+\Omega_{2.u\Delta u^*}\right)\right) \\ &- B\left(\tau_x(1-\tau_y)\alpha_x\alpha_yA^*(1)B^*(1)+\tau_y(1-\tau_x)\alpha_x\alpha_yA^*(1)B^*(1)\right). \end{aligned}$$

The variance covariance quantities are defined as follows:

$$
\begin{aligned}
\lim_{T\longrightarrow\infty} T^{-1}\sum_{1}^{T} u_t^2 &= \sigma_u^2, \\
\lim_{T\longrightarrow\infty} T^{-1}\sum_{1}^{T} \Delta^2 u_t^* &= \Omega_{\Delta u^*}, \\
\lim_{T\longrightarrow\infty} T^{-1}\sum_{1}^{T} u_t \Delta u_t^* &= \Omega_{u\Delta u^*}, \\
\lim_{T\longrightarrow\infty} T^{-1}\sum_{1}^{T} u_t \Delta u_{t+s}^* &= \Omega_{1.u\Delta u^*}, \\
\lim_{T\longrightarrow\infty} T^{-1}\sum_{1}^{T} u_{t+s} \Delta u_t^* &= \Omega_{2.u\Delta u^*}, \\
\textit{and} \lim_{T\longrightarrow\infty} T^{-1}\sum_{1}^{T} \Delta u_t^* \Delta u_{t+s}^* &= \Omega_{\Delta u^*}^*.
\end{aligned}
$$

Also note that covariance quantities, $\Omega_{\Delta u^*}^*$, $\Omega_{\Delta v^*}^*$, $\Omega_{1.u\Delta u^*}$, $\Omega_{2.u\Delta u^*}$, $\Omega_{1.v\Delta v^*}$, *and* $\Omega_{2.v\Delta v^*}$, *vanish when the time lag,* s, *approaches infinity.*

Proof. The proof is presented in the appendix. □

It is worth noting that, if there are no structural breaks in the means, in view of the previous theorem we obtain

$$
\begin{aligned}
T^{-1}\sum_{t=1}^{T} x_t y_t &\xrightarrow{a.s.} \alpha_x \alpha_y A^*(1) B^*(1), \\
T^{-1}\sum_{t=1}^{T} x_t &\xrightarrow{a.s.} \alpha_x A^*(1), \\
T^{-1}\sum_{t=1}^{T} y_t &\xrightarrow{a.s.} \alpha_y B^*(1).
\end{aligned}
$$

And if x_t has structural breaks whilst y_t does not, we obtain

$$\begin{aligned}
T^{-1}\sum_{1}^{T} x_t y_t &\xrightarrow{a.s.} \alpha_x \alpha_y (1-\tau_x) A^*(1) B^*(1),\\
T^{-1}\sum_{1}^{T} x_t &\xrightarrow{a.s.} (1-\tau_x)\alpha_x A^*(1),\\
T^{-1}\sum_{1}^{T} y_t &\xrightarrow{a.s.} B^*(1)\alpha_y.
\end{aligned}$$

The above limits imply that $\hat{\beta} \xrightarrow{a.s.} 0$. Hence, no spurious regressions occur.

3 Spurious Regression with Structural Breaks in Trend

Let us consider the following DGPs:

$$\begin{aligned}
A(L)x_t &= \alpha_x + \mu_x t + \mu_{x,b}(t-[T\tau_x])\mathbf{1}\{t>[T\tau_x]\} + u_t,\\
B(L)y_t &= \alpha_y + \mu_y t + \mu_{y,b}(t-[T\tau_y])\mathbf{1}\{t>[T\tau_y]\} + v_t, \qquad (3.1)
\end{aligned}$$

where lag polynomials, $A(L)$ and $B(L)$, have roots outside the unit circle. And $\tau_x < \tau_y$ with the innovation terms u_t and v_t, satisfying Assumption 1.

A BN decomposition of the inverse lag operator $A^*(L)$ and $B^*(L)$ yields

$$\begin{aligned}
x_t &= A^*(1)\alpha_x + \mu_x\left(A^*(1)t + A^\diamond(1)\right) + \mu_{x,b}\left(A^*(1)(t-[T\tau_x]) + A^\diamond(1)\right)\mathbf{1}\{t>[T\tau_x]\} + A^*(1)u_t + \Delta u_t^*,\\
y_t &= B^*(1)\alpha_y + \mu_y\left(B^*(1)t + B^\diamond(1)\right) + \mu_{y,b}\left(B^*(1)(t-[T\tau_y]) + B^\diamond(1)\right)\mathbf{1}\{t>[T\tau_y]\} + B^*(1)v_t + \Delta v_t^*,
\end{aligned}$$

where u_t^*, v_t^*, $A^\diamond$, and $B^\diamond$ are defined in Section **2**.

Now we can formulate the following regression equation:

$$y_t = \hat{\gamma}_1 + \hat{\gamma}_2 t + \hat{\beta} x_t + \hat{w}_t,$$

where the OLS estimates are given as

$$\begin{bmatrix} \hat{\gamma}_1 \\ \hat{\beta} \\ \hat{\gamma}_2 \end{bmatrix} = \begin{bmatrix} T & \sum_1^T x_t & \sum_1^T t \\ \sum_1^T x_t & \sum_1^T x_t^2 & \sum_1^T tx_t \\ \sum_1^T t & \sum_1^T tx_t & \sum_1^T t^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum_1^T y_t \\ \sum_1^T x_t y_t \\ \sum_1^T ty_t \end{bmatrix}. \tag{3.2}$$

The limiting behaviors of regression statistics are stated in Theorem 2 below.

Theorem 2. *As the sample size T goes to infinity, we have*

$$\begin{aligned} T^{-1}\hat{\gamma}_1 &\xrightarrow{a.s.} \gamma_1^{(lim)}, \\ \hat{\beta} &\xrightarrow{a.s.} \beta^{(lim)}, \\ \hat{\gamma}_2 &\xrightarrow{a.s.} \gamma_2^{(lim)}, \\ T^{-2}\hat{s}^2 &\xrightarrow{a.s.} b_4 + {\gamma_2^{(lim)}}^2 a_{33} + {\beta^{(lim)}}^2 a_{22} - 2\gamma_1^{(lim)} b_1 - 2\gamma_2^{(lim)} b_3 - 2\beta^{(lim)} b_2 + 2\gamma_2^{(lim)}\gamma_1^{(lim)} a_{13} \\ &+ \quad 2\beta^{(lim)}\gamma_1^{(lim)} a_{12} + 2\gamma_2^{(lim)}\beta^{(lim)} a_{23} = {s^{(lim)}}^2. \end{aligned}$$

The limits of the t *tests are given by*

$$\begin{aligned} T^{-1/2} t_{\hat{\gamma}_1} &\xrightarrow{a.s.} \frac{\gamma_1^{(lim)}}{\sigma_{\gamma_1}^{(lim)}}, \\ T^{-1/2} t_{\hat{\beta}} &\xrightarrow{a.s.} \frac{\beta^{(lim)}}{\sigma_{\beta}^{(lim)}}, \\ T^{-1/2} t_{\hat{\gamma}_2} &\xrightarrow{a.s.} \frac{\gamma_2^{(lim)}}{\sigma_{\gamma_2}^{(lim)}}, \end{aligned}$$

where

$$\begin{bmatrix} \gamma_1^{(lim)} \\ \beta^{(lim)} \\ \gamma_2^{(lim)} \end{bmatrix} = \mathcal{A}^{-1}\mathcal{B},$$

with $\mathcal{A}$ is a nonsingular symmetric matrix, defined as

$$\mathcal{A} = \begin{bmatrix} 1 & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

and

$$\mathcal{B}' = [b_1,\ b_2,\ b_3],$$

(the definitions of $\mathcal{A}$ and $\mathcal{B}$ will be given in the proof.)

$$\begin{aligned} \sigma_{\gamma_1}^{(lim)^2} &= s^{(lim)} \mathbb{I}_{(1)} \mathcal{A}^{-1} \mathbb{I}_{(1)}', \\ \sigma_{\beta}^{(lim)^2} &= s^{(lim)} \mathbb{I}_{(2)} \mathcal{A}^{-1} \mathbb{I}_{(2)}', \\ \sigma_{\gamma_2}^{(lim)^2} &= s^{(lim)} \mathbb{I}_{(3)} \mathcal{A}^{-1} \mathbb{I}_{(3)}', \end{aligned}$$

(the notation $\mathbb{I}_{(i)}$ denotes the vector in which the i-th element is 1 and the other elements are 0.)

The limit of the coefficient of determination is given by

$$R^2 = \hat{\beta}^2 \frac{\sum_1^T (x_t - \overline{x})^2}{\sum_1^T (y_t - \overline{y})^2} \xrightarrow{a.s.} \beta^{(lim)} \left\{ \frac{b_1}{a_{12}} \right\}^2.$$

The limit of the sample autocorrelation of residuals is given by

$$\hat{r}_s = \frac{\sum_{1+s}^T \hat{w}_{t-s} \hat{w}_t}{\sum_1^T \hat{w}_t^2} \xrightarrow{a.s.} \frac{\begin{array}{c} b_4 + \beta^{(lim)^2} a_{22} - 2\big\{\beta^{(lim)} b_2 - \gamma_2^{(lim)} b_3 - \gamma_1^{(lim)} b_1 + \beta^{(lim)} \gamma_1^{(lim)} a_{12} \\ + \beta^{(lim)} \gamma_2^{(lim)} a_{23} + \gamma_1^{(lim)} \gamma_2^{(lim)} \int_0^1 s ds\big\} + \gamma_1^{(lim)^2} \end{array}}{s^{(lim)^2}}.$$

Note that, if x_t has structural breaks whilst y_t does not, the results in Theorem 2 still hold, with

b_1, b_2, b_3 and b_4 being replaced by

$$
\begin{aligned}
b_1 &= \mu_y B^*(1) \int_0^1 s ds, \\
b_2 &= A^*(1)B^*(1) \left(\mu_x \mu_y \int_0^1 s^2 ds + \mu_{xb}\mu_y \left(\int_{\tau_x}^1 s^2 ds - \tau_x \int_{\tau_x}^1 s ds \right) \right), \\
b_3 &= \mu_y B^*(1) \int_0^1 s^2 ds, \\
b_4 &= \mu_y^2 B^{*2}(1) \int_0^1 s^2 ds.
\end{aligned}
$$

4 Simulations

In this section, we present some small simulations to confirm the above theoretical findings. In the sequel, we run two regressions

$$y_t = \gamma + \beta x_t + w_t \tag{4.1}$$

and

$$y_t = \gamma_1 + \gamma_2 t + \beta x_t + w_t, \tag{4.2}$$

where x_t and y_t follow $AR(1)$ processes with structural breaks in the means and the trends, as defined in Eq. (3.1). First, we consider the case with structural breaks in the mean (cf. (4.1)). Then, we consider the case with structural breaks in the trend (cf. (4.2)).

Structural Break in Mean

We generate the two following processes:

$$
\begin{aligned}
x_t &= \phi_x x_{t-1} + \alpha_x \mathbf{1}\{t > [T\tau_x]\} + u_t, \\
y_t &= \phi_y y_{t-1} + \alpha_y \mathbf{1}\{t > [T\tau_y]\} + v_t,
\end{aligned} \tag{4.3}
$$

where $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, and $\phi_y = 0.75$. In this simulation, we assume that the processes u_t and v_t are *independent white noises* for illustration purposes.

This is unfortunately because simulations for general weakly dependent innovations may be quite complicated to implement.

The processes (4.3) are initialized with $x_0 = 10$ and $y_0 = 10$. In order to show the convergence of the estimates of the regression coefficients $\hat{\gamma}$ and $\hat{\beta}$, their t-statistics t_γ and t_β, and serial correlation coefficient $\hat{r}_s$ to their limiting values, we simulate data with increasing sample size, $T = 1,000$ to $2,000,000$, from the processes x_t and y_t. The true limiting values, the simulated mean values of the OLS statistics, and their 95% confidence bounds are presented in Table 1 and Figure 1.

Structural Break in Trend

To illustrate the case with structural breaks in the trends, we simulate the following two processes:

$$
\begin{aligned}
x_t &= \phi_x x_{t-1} + \alpha_x + \mu_x t + \mu_{x,b}(t - [T\tau_x])\mathbf{1}\{t > [T\tau_x]\} + u_t, \\
y_t &= \phi_y y_{t-1} + \alpha_y + \mu_y t + \mu_{y,b}(t - [T\tau_y])\mathbf{1}\{t > [T\tau_y]\} + v_t,
\end{aligned} \tag{4.4}
$$

where $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $\mu_x = 0.2$, $\mu_{x,b} = 0.1$, $\mu_y = 0.3$, $\mu_{y,b} = 0.05$, with u_t and v_t are, as before, *independent white noises*. Using the same initializations as in the previous simulation, $x_0 = 10$ and $y_0 = 10$, we run the regression (4.2) with the data simulated from (4.4) and estimate the OLS statistics. We repeat the simulation 1,000 times to compute the mean values and the 95% confidence bounds of the OLS coefficients $\hat{\gamma}_1$, $\hat{\gamma}_2$ and $\hat{\beta}$, their t-statistics, the autocorrelation coefficient of the residuals, and the determination coefficient R^2. The results are provided in Table 2 and Figure 2.

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 main contributions in this paper are 1) to show the evidences of spurious regressions in the presence of structural breaks by proving the limiting properties of the standard OLS statistics and 2), more importantly, to prove new SLLNs with *minimum* necessary conditions as the tools to tackle the stochastic limits of the OLS statistics.

Tables and Figures

Table 1: Simulation result: structural breaks in the mean

This table presents the simulation results for the OLS statistics convergence under a structural break in the mean. We perform this exercise to demonstrate the convergence of the mean of OLS statistics to their limiting values, computed from Theorem 1. 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. 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	0.373060 ($2.1583 \cdot 10^{-3}$)	0.484633 ($7.7631 \cdot 10^{-3}$)	0.330414 ($2.0393 \cdot 10^{-3}$)	0.101586 ($1.6321 \cdot 10^{-3}$)	0.935153 ($1.9975 \cdot 10^{-4}$)
10,000	0.364988 ($6.6433 \cdot 10^{-4}$)	0.503534 ($2.3915 \cdot 10^{-3}$)	0.321041 ($6.2668 \cdot 10^{-4}$)	0.105044 ($5.0354 \cdot 10^{-4}$)	0.942552 ($6.0014 \cdot 10^{-5}$)
50,000	0.364193 ($2.9574 \cdot 10^{-4}$)	0.506336 ($1.0741 \cdot 10^{-3}$)	0.320185 ($2.7739 \cdot 10^{-4}$)	0.105572 ($2.2595 \cdot 10^{-4}$)	0.943115 ($2.6024 \cdot 10^{-5}$)
100,000	0.364738 ($2.0771 \cdot 10^{-4}$)	0.505706 ($7.3285 \cdot 10^{-4}$)	0.320596 ($1.9269 \cdot 10^{-4}$)	0.105434 ($1.5498 \cdot 10^{-4}$)	0.943121 ($1.8258 \cdot 10^{-5}$)
200,000	0.364592 ($1.5281 \cdot 10^{-4}$)	0.506516 ($5.5929 \cdot 10^{-4}$)	0.320566 ($1.4125 \cdot 10^{-4}$)	0.105626 ($1.1764 \cdot 10^{-4}$)	0.943150 ($1.2725 \cdot 10^{-6}$)
500,000	0.364675 ($9.3489 \cdot 10^{-5}$)	0.506062 ($3.3443 \cdot 10^{-4}$)	0.320572 ($8.6915 \cdot 10^{-5}$)	0.105514 ($7.0744 \cdot 10^{-5}$)	0.943178 ($8.2400 \cdot 10^{-6}$)
1,000,000	0.364507 ($6.7257 \cdot 10^{-5}$)	0.506276 ($2.4530 \cdot 10^{-4}$)	0.320415 ($6.3082 \cdot 10^{-5}$)	0.105554 ($5.1475 \cdot 10^{-5}$)	0.943182 ($5.7439 \cdot 10^{-6}$)
2,000,000	0.364501 ($4.9601 \cdot 10^{-5}$)	0.506574 ($1.7977 \cdot 10^{-4}$)	0.320435 ($4.5505 \cdot 10^{-5}$)	0.105625 ($3.8107 \cdot 10^{-5}$)	0.943186 ($4.1485 \cdot 10^{-6}$)
5,000,000	0.364515 ($2.9319 \cdot 10^{-5}$)	0.506203 ($1.0745 \cdot 10^{-4}$)	0.320436 ($2.7341 \cdot 10^{-5}$)	0.105540 ($2.2555 \cdot 10^{-5}$)	0.943183 ($2.5422 \cdot 10^{-6}$)
Limit Value	**0.364557**	**0.506329**	**0.320468**	**0.105564**	**0.943194**

Figure 1: Convergence of the OLS statistics: structural breaks in the mean

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 (blue dashed line), computed from Theorem 1.

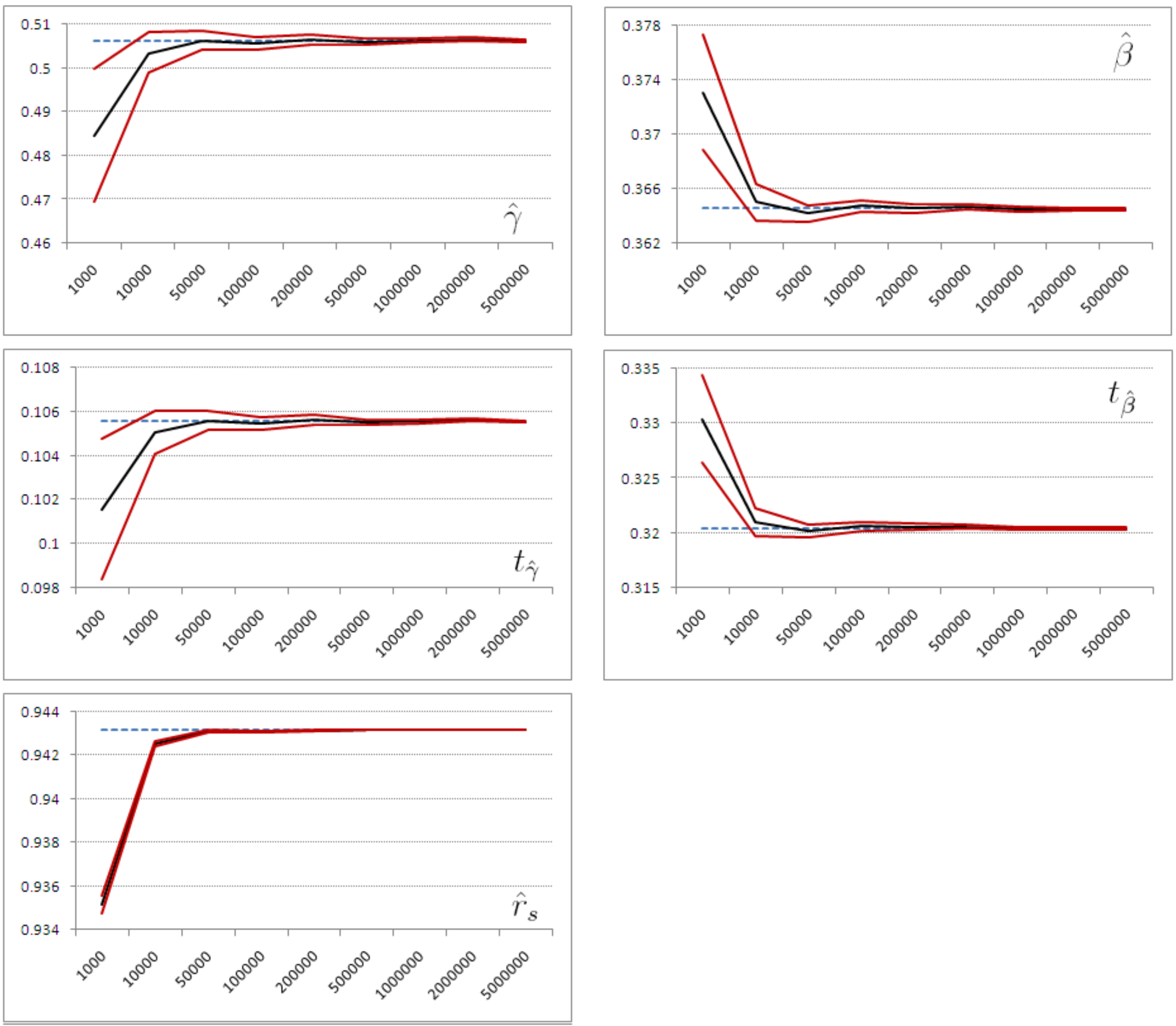


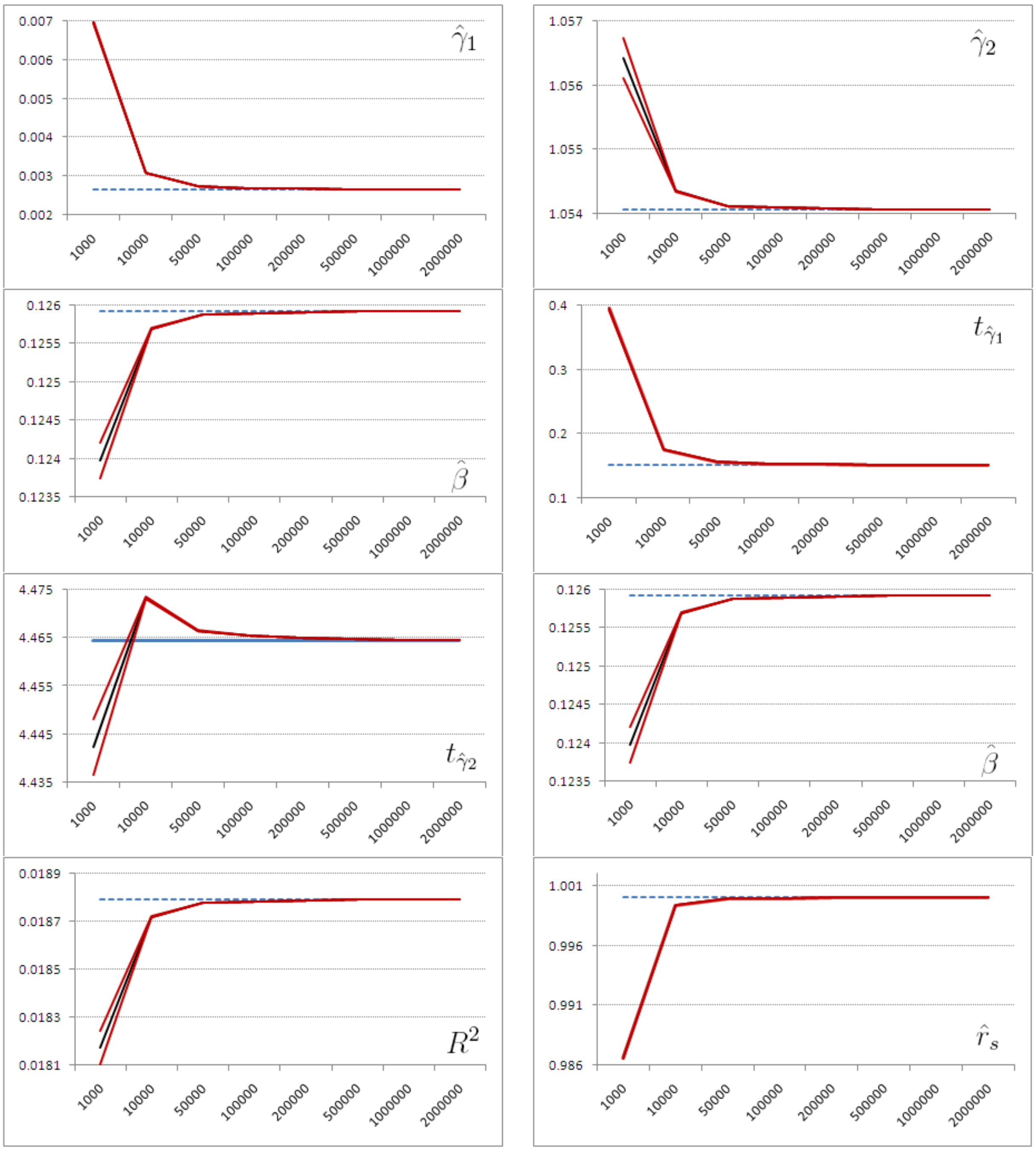
Table 2: Simulation result: structural breaks in the trend

This table presents the simulation results for the OLS statistics convergence under structural breaks in the trend. We perform this exercise to demonstrate the convergence of the mean of the OLS statistics to their limiting values, computed from Theorem 2. First, two *independent* $AR(1)$ processes, defined in Eq. (4.3), are simulated, given $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $\mu_x = 0.2$, $\mu_{x,b} = 0.1$, $\mu_y = 0.3$, $\mu_{y,b} = 0.05$, $x_0 = 10$, and $y_0 = 10$. Then the regression coefficients, their t-statistics, and the serial correlations of regression residuals are estimated from these simulated data. 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 $(1.1497{\cdot}10^{-5})$	1.056422 $(1.5991{\cdot}10^{-4})$	0.123978 $(1.2106{\cdot}10^{-4})$	0.394673 $(7.0295{\cdot}10^{-4})$	4.442476 $(2.9383{\cdot}10^{-3})$	0.691611 $(8.2153{\cdot}10^{-4})$	0.0181748 $(3.5507{\cdot}10^{-5})$	0.986649 $(3.0619{\cdot}10^{-5})$
10,000	0.00309036 $(3.8389{\cdot}10^{-7})$	1.054361 $(5.2202{\cdot}10^{-6})$	0.125696 $(3.9215{\cdot}10^{-6})$	0.176181 $(2.2386{\cdot}10^{-5})$	4.473360 $(9.9254{\cdot}10^{-5})$	0.708730 $(2.7363{\cdot}10^{-5})$	0.0187166 $(1.1685{\cdot}10^{-6})$	0.999308 $(2.4838{\cdot}10^{-7})$
50,000	0.00275121 $(3.4012{\cdot}10^{-8})$	1.054134 $(4.5787{\cdot}10^{-7})$	0.125878 $(3.4380{\cdot}10^{-7})$	0.156599 $(1.9728{\cdot}10^{-6})$	4.466557 $(9.0907{\cdot}10^{-6})$	0.708934 $(2.4598{\cdot}10^{-6})$	0.0187752 $(1.0261{\cdot}10^{-7})$	0.999881 $(9.7210{\cdot}10^{-9})$
100,000	0.00270894 $(1.2323{\cdot}10^{-8})$	1.054104 $(1.6924{\cdot}10^{-7})$	0.125902 $(1.2711{\cdot}10^{-7})$	0.154157 $(7.0885{\cdot}10^{-7})$	4.465557 $(3.0839{\cdot}10^{-6})$	0.708941 $(8.5608{\cdot}10^{-7})$	0.0187828 $(3.7927{\cdot}10^{-8})$	0.999942 $(2.5367{\cdot}10^{-9})$
200,000	0.00268780 $(4.2423{\cdot}10^{-9})$	1.054089 $(5.6266{\cdot}10^{-8})$	0.125914 $(4.2267{\cdot}10^{-8})$	0.152937 $(2.4518{\cdot}10^{-7})$	4.465054 $(1.1190{\cdot}10^{-6})$	0.708946 $(2.9057{\cdot}10^{-7})$	0.0187866 $(1.2613{\cdot}10^{-8})$	0.999971 $(2.4673{\cdot}10^{-10})$
500,000	0.00267512 $(1.1103{\cdot}10^{-9})$	1.054080 $(1.4492{\cdot}10^{-8})$	0.125921 $(1.0822{\cdot}10^{-8})$	0.152205 $(6.3649{\cdot}10^{-8})$	4.464749 $(2.7783{\cdot}10^{-7})$	0.708948 $(7.3034{\cdot}10^{-8})$	0.0187889 $(3.2329{\cdot}10^{-9})$	0.999989 $(9.9200{\cdot}10^{-11})$
1,000,000	0.00267089 $(3.9018{\cdot}10^{-10})$	1.054077 $(5.8109{\cdot}10^{-9})$	0.125924 $(4.0376{\cdot}10^{-9})$	0.151961 $(2.2412{\cdot}10^{-8})$	4.464648 $(9.8737{\cdot}10^{-8})$	0.708949 $(2.6966{\cdot}10^{-8})$	0.0187897 $(1.2056{\cdot}10^{-9})$	0.999994 $(2.5230{\cdot}10^{-11})$
2,000,000	0.00266878 $(1.3448{\cdot}10^{-10})$	1.054076 $(1.9495{\cdot}10^{-9})$	0.125925 $(1.4022{\cdot}10^{-9})$	0.151839 $(7.7272{\cdot}10^{-9})$	4.464597 $(3.6117{\cdot}10^{-8})$	0.708949 $(9.3567{\cdot}10^{-9})$	0.0187901 $(4.1522{\cdot}10^{-10})$	0.999997 $(6.0400{\cdot}10^{-12})$
Limit Value	**0.00266667**	**1.054070**	**0.125926**	**0.151717**	**4.464550**	**0.708949**	**0.0187905**	**1.000000**

Figure 2: Convergence of estimates: stuctural breaks in trend

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. (4.3), with $\tau_x = 0.4$, $\tau_y = 0.8$, $\alpha_x = 1.0$, $\alpha_y = 2.0$, $\phi_x = 0.8$, $\phi_y = 0.75$, $\mu_x = 0.2$, $\mu_{x,b} = 0.1$, $\mu_y = 0.3$, $\mu_{y,b} = 0.05$, $x_0 = 10$, and $y_0 = 10$. As the number of observations, T, increases, the OLS statistics converge to their limiting values (blue dashed line), computed from Theorem 2.



Appendices: Proofs

Useful lemmas

Proofs of Theorems 1 and 2 are based on the lemmas, which we shall state below.

Lemma 1 (SLLN for strictly stationary processes)**.** *Let X_t denote a strictly stationary process. Suppose that*

$$\begin{aligned} E[X_0] &= 0, \\ E[|X_0|^p] &< \infty, \end{aligned}$$

and

$$\sum_{t=1}^{\infty} \gamma_t(p) < \infty \text{ for some } p \geq 2, \tag{5.1}$$

where $\gamma_t(p) = ||E[X_t|\mathcal{F}_0]||_{\frac{p}{p-1}}$. (The notation $||X||_p = \big[E[|X|^p]\big]^{1/p}$ stands for Holder's norm.) Then

$$\frac{1}{T}\sum_{t=1}^{T} X_t \xrightarrow{a.s.} 0.$$

Proof. Let us define the following random sequences:

$$\begin{aligned} \eta_t &= \sum_{s \geq t} E[X_s|\mathcal{F}_t], \\ \zeta_t &= \sum_{s \geq t} \big\{E[X_s|\mathcal{F}_t] - E[X_s|\mathcal{F}_{t-1}]\big\}. \end{aligned}$$

Then,

$$\frac{1}{T}\sum_{t=1}^{T} X_t = \frac{1}{T}\sum_{1}^{T} \zeta_t + \frac{\eta_1 - \eta_{T+1}}{T}.$$

First, we need to show that

$$\eta_1 - \eta_{T+1} = o_{a.s.}(T). \tag{5.2}$$

By Holder's inequality and (5.1), we have

$$||\eta_t||_{\frac{p}{p-1}} = ||\sum_{s\geq 0} E[X_s|\mathcal{F}_{-1}]||_{\frac{p}{p-1}} = ||\sum_{s\geq 0} E\big[E[X_s|\mathcal{F}_{-1}]|\mathcal{F}_0\big]||_{\frac{p}{p-1}}$$
$$\leq \Big(\sum_{s\geq 0} E|E[X_s|\mathcal{F}_{-1}]|^{\frac{p}{p-1}}\Big)^{\frac{p-1}{p}} < \infty.$$

Hence, in view of Lehmann and Romano (2005, Prob. 11.29), we obtain $\lim_{\substack{T\longrightarrow\infty \\ h\longrightarrow\infty}}\{E[\min(|\eta_1 - \eta_{T+1}|, h)] - E|\eta_1 - \eta_{T+1}|\} = 0$. Then, an application of the Markov inequality yields (5.2).

Next, because $E[\zeta_t|\mathcal{F}_{t-1}] = 0$, ζ_t is a martingale difference sequence (mds). We need to show that

$$\frac{1}{T}\sum_{t=1}^{T} \zeta_t \xrightarrow{a.s.} 0. \tag{5.3}$$

By the SLLN for mds (see, e.g., White (1980, p. 58)), we only need to verify that $E|\zeta_1|^2 < \infty$, which is straightforward in view of the following equality:

$$E|\zeta_1|^2 = E[X_0^2] + 2\sum_{t\geq 1} E[X_0 X_t],$$

which is the standard formula for the variance of a strictly stationary stochastic sequence. □

Let us recall definitions of the mixing coefficients ϕ_t, α_t, and ρ_t (see, e.g., White (1980)).

1. The uniform strongly mixing coefficient:

$$\phi_t = \sup_{A\in\mathcal{F}_0^t,\ B\in\mathcal{F}_{t+1}^\infty} |P(B|A) - P(B)|, \tag{5.4}$$

   where $\mathcal{F}_s^t = \sigma\{X_s, X_{s+1}, \ldots, X_t\}$.

2. The strongly mixing coefficient:

$$\alpha_t = \sup_{A\in\mathcal{F}_0^t,\ B\in\mathcal{F}_{t+1}^\infty} |P(A\bigcap B) - P(A)P(B)|. \tag{5.5}$$

3. The maximal correlation coefficient:

$$\rho_t = \sup_{X \in \mathcal{F}_0^t,\ Y \in \mathcal{F}_{t+1}^\infty} \mathrm{Cov}(X, Y) \tag{5.6}$$

Corollary 1. *Let $E|X_0|^p < \infty$. Suppose that one of the following assumptions holds:*

$$\sum_{t=1}^{\infty} \phi_t^{\frac{p-1}{p}} < \infty \text{ for some } p \geq 2. \tag{5.7}$$

$$\sum_{t=1}^{\infty} \alpha_t^{\frac{p-1}{p}} < \infty \text{ for some } p > 2. \tag{5.8}$$

$$\sum_{t=1}^{\infty} \rho_t^{1/2} < \infty. \tag{5.9}$$

Then

$$\frac{1}{T} \sum_{t=1}^{T} X_t \xrightarrow{a.s.} 0.$$

Proof of Corollary 1. The proof follows from Lemma 1 and the following inequalities:

$$\gamma_t(p) \leq 2\phi_t^{\frac{p-1}{p}} ||X_0||_p \text{ for some } p \geq 2, \tag{5.10}$$

$$\gamma_t(p) \leq 2(2^{1/p} + 1)\alpha_t^{\frac{p-2}{p}} ||X_0||_p \text{ for some } p > 2, \tag{5.11}$$

$$\gamma_t(2) \leq \rho_t^{1/2}. \tag{5.12}$$

(See, e.g., McLeish (1975).) □

Corollary 1 is the same as the SLLN established by McLeish (1975). Nevertheless, the *absolute moments* condition (5.1) is much weaker than the *mixing coefficients* conditions (5.7) - (5.9). This is because the conditions (5.7) - (5.9) implies the condition (5.1), and the converse is not so true.

Corollary 2. *Let one of the assumptions stated in Corollary 1 hold. Then, we have*

$$\frac{1}{T} \sum_{t=1}^{T} X_t X_{t+s} \xrightarrow{a.s.} E[X_0 X_s].$$

In addition, suppose that $||A(1)||_1 < \infty$. Then, we have

$$\frac{1}{T}\sum_{t=1}^{T} X_{t+s}X_t^* \xrightarrow{a.s.} A(L)E[X_0X_s],$$

where $X_t^* = A(L)X_t$.

Proof of Corollary 2. Let us define

$$S_T^* = \frac{1}{T}\sum_{t=1}^{T}\{X_tX_{t+s} - E[X_0X_s]\},$$

where, by strict stationarity, $E[X_tX_{t+s}] = E[X_0X_s]$. For the first part, in view of Lemma 1 we need to verify the following two conditions:

$$E[X_0X_s - E[X_0X_s]]^p < \infty \tag{5.13}$$

and

$$\sum_{t=1}^{\infty} \gamma_t^*(p) < \infty, \tag{5.14}$$

where $\gamma_t^*(p) = ||E[X_0X_s|\mathcal{F}_{s-t}] - E[X_0X_s]||_{\frac{p}{p-1}}$.

First, by Minkowski's inequality, we obtain

$$||X_0X_s - E[X_0X_s]||_p \leq E[X_0X_s] + ||X_0X_s||_p.$$

Holder's inequality yields

$$||X_0X_s||_p \leq ||X_0^2||_p < \infty \text{ for any } p \geq 1.$$

Hence, Eq. (5.13) is proved.

Next, by the inequalities for mixingales in (McLeish, 1975), we obtain

$$\begin{aligned}
\sum_{t=1}^{\infty} ||E[X_0X_s|\mathcal{F}_{s-t}] - E[X_0X_s]||_{\frac{p}{p-1}} &\leq 2\hat{\phi}^{\frac{p-1}{p}}(\mathcal{F}_s^{+\infty}, \mathcal{F}_{-\infty}^{s-t})||X_0X_s||_p \text{ for some } p > 2,\\
\sum_{t=1}^{\infty} ||E[X_0X_s|\mathcal{F}_{s-t}] - E[X_0X_s]||_{\frac{p}{p-1}} &\leq 2(2^{1/p}+1)\hat{\alpha}_t^{\frac{p-2}{p}}(\mathcal{F}_s^{+\infty}, \mathcal{F}_{-\infty}^{s-t})||X_0X_s||_p.
\end{aligned}$$

Because X_s is $\mathcal{F}_s$ measurable and $X_0 X_s$ is a finite stretch of X_s, the product process $X_0 X_s$ is also $\mathcal{F}_s$ measurable. Hence, the mixing coefficients $\phi_T \longrightarrow 0$ and $\alpha_T \longrightarrow 0$ imply $\hat{\phi}_T \longrightarrow 0$ and $\hat{\alpha}_T \longrightarrow 0$, respectively. As a result, the summability of $\hat{\phi}_t$ and $\hat{\alpha}_t$ holds. Eq. (5.14) follows.

For the second part, let $A(L) = \sum_{s=0}^{K} A_s L^s$. The second part follows by applying the first part K times. The condition $||A(1)||_2 < \infty$ purports to ensure that $\sum_{s=0}^{K} A_s L^s E[X_0 X_t] < \infty$. □

Corollary 3. *Let all the assumptions in Lemma 1 hold. Then*

$$\frac{1}{T}\sum_{t=1}^{T} X_t X_{t+s} \xrightarrow{a.s.} 0, \ \ \textit{as } s \longrightarrow \infty.$$

Proof of Corollary 3. We need to show that

$$\sum_{t=1}^{\infty} \gamma_t^* < \infty, \text{ as } s \longrightarrow \infty. \tag{5.15}$$

Eq. (5.15) can be rewritten as

$$\sum_{t=1}^{\infty} \gamma_t^*(p) \ \ \leq \ \ \underbrace{\sum_{t=1}^{s} ||E[X_0 X_s|\mathcal{F}_{s-t}] - E[X_0 X_s]||_{\frac{p}{p-1}}}_{\mathbf{I}} + \underbrace{\sum_{t=s+1}^{\infty} ||E[X_0 X_s|\mathcal{F}_{s-t}] - E[X_0 X_s]||_{\frac{p}{p-1}}}_{\mathbf{II}} \tag{5.16}$$

By strictly stationarity, The first term, $\mathbf{I}$, can be rewritten as

$$\mathbf{I} \leq ||X_0||_{\frac{p}{p-1}} \sum_{t=1}^{s} ||E[X_0|\mathcal{F}_{-t}||_p$$

Noting that, under the assumptions of Lemma 1, it is straightforward to verify that $E[X_0|\mathcal{F}_{-\infty}] = \lim_{T\longrightarrow\infty} |E[X_0|\mathcal{F}_{-T}| = 0$ a.s.. Accordingly, we obtain

$$\mathbf{I} < \infty \text{ as } s \longrightarrow \infty.$$

An application of Minkowski's inequality yields

$$||E[X_0 X_s|\mathcal{F}_{s-t}]||_{\frac{p}{p-1}} \leq ||E[X_0 X_s|\mathcal{F}_{s-t}]||_{\frac{p}{p-1}} + E[X_0 X_s].$$

Furthermore, by Holder's inequality we have

$$E[X_0 X_s|\mathcal{F}_{s-t}] \leq \left\{E[|X_0|^{\frac{p}{p-1}}|\mathcal{F}_{s-t}]\right\}^{\frac{p-1}{p}} \left\{E\left[|E[X_s|\mathcal{F}_0]|^p|\mathcal{F}_{s-t}\right]\right\}^{\frac{1}{p}},$$

for $t > s$ and $p > 1$. Hence

$$||E[X_0 X_s|\mathcal{F}_{s-t}]||_{\frac{p}{p-1}} \leq |E[X_s|\mathcal{F}_0]|\left\{E\left|E[|X_0|^{\frac{p}{p-1}}|\mathcal{F}_{s-t}]\right|^{\frac{p-1}{p-2}}\right\}^{\frac{p-2}{p}}.$$

for some $p > 2$. By a similar argument, we have

$$E[X_0 X_s] = E[X_0 E[X_s|\mathcal{F}_0]] \leq ||X_0||_{\frac{p}{p-1}}||E[X_s|\mathcal{F}_0]||_p.$$

It follows that the second term, **II**, satisfies

$$\mathbf{II} \leq (T-s)||X_0||_{\frac{p}{p-1}}||E[X_s|\mathcal{F}_0]||_p + |E[X_s|\mathcal{F}_0]|\sum_{t=1+s}^{\infty}\left\{E\left|E[|X_0|^{\frac{p}{p-1}}|\mathcal{F}_{s-t}]\right|^{\frac{p-1}{p-2}}\right\}^{\frac{p-2}{p}}.$$

By strictly stationarity condition and the assumptions of Lemma 1, we obtain $E[X_0|\mathcal{F}_{-\infty}] = E[X_{+\infty}|\mathcal{F}_0] = 0$. Accordingly,

$$\mathbf{II} = 0$$

Therefore, (5.15) is proved. $\square$

Proof of Theorem 1

The OLS estimates of the regression coefficients can be rewritten as

$$\begin{aligned}
\hat{\beta} &= \frac{T^{-1}\sum_1^T x_t y_t - \overline{xy}}{T^{-1}\sum_1^T (x_t-\overline{x})^2}, && (5.17)\\
\hat{\gamma} &= \overline{y} - \hat{\beta}\overline{x}, && (5.18)\\
\hat{s}^2 &= T^{-1}\sum_1^T (y_t-\overline{y})^2 - \hat{\beta}^2 T^{-1}\sum_1^T (x_t-\overline{x})^2, && (5.19)\\
T^{-1/2}t_{\hat{\beta}} &= \frac{\hat{\beta}}{s\left(T^{-1}\sum_1^T (x_t-\overline{x})^2\right)^{-1/2}}, && (5.20)\\
T^{-1/2}t_{\hat{\gamma}} &= \frac{\hat{\alpha}}{\hat{s}}\frac{\left(T^{-1}\sum_1^T (x_t-\overline{x})^2\right)^{1/2}}{\left(T^{-1}\sum_1^T x_t^2\right)^{1/2}}, && (5.21)\\
\hat{r}_s &= \frac{\sum_{1+s}^T \hat{w}_t\hat{w}_{t-s}}{\sum_1^T \hat{w}_t^2}, && (5.22)
\end{aligned}$$

where

$$T^{-1}\sum_1^T (x_t-\overline{x})^2 = \overline{x^2} - (\overline{x})^2.$$

1.

$$\begin{aligned}
T^{-1}\sum_1^T x_t y_t &= T^{-1}\sum_1^{[T\tau_x]} + \sum_{[T\tau_x]+1}^{[T\tau_y]} + \sum_{[T\tau_y]+1}^{T} x_t y_t = T^{-1}\sum_1^{[T\tau_x]}(A^*(1)u_t + \Delta u_t^*)(B^*(1)v_t + \Delta v_t^*)\\
&+ T^{-1}\sum_{[T\tau_x]+1}^{[T\tau_y]}(A^*(1)\alpha_x + A^*(1)u_t + \Delta u_t^*)(B^*(1)v_t + \Delta v_t^*) + T^{-1}\sum_{[T\tau_y]+1}^{T}(A^*(1)\alpha_x\\
&+ A^*(1)u_t + \Delta u_t^*)(B^*(1)\alpha_y + B^* v_t + \Delta v_t^*) = I + II + III.
\end{aligned}$$

Since u_t and v_t are independent, $||E[u_t v_t|\mathcal{F}_0]||_{\frac{p}{p-1}} = ||E[u_t|\mathcal{F}_0]||_{\frac{p}{p-1}}||E[v_t|\mathcal{F}_0]||_{\frac{p}{p-1}}$. Thus, Assumption 1 implies that $u_t v_t$ satisfies Condition 5.1, i.e. $T^{-1}\sum_1^T u_t v_t \xrightarrow{a.s.} 0$. Similarly, we have

$T^{-1}\sum_1^T \Delta u_t^* \Delta v_t^* \xrightarrow{a.s.} 0$. This results in

$$\begin{aligned} I &\xrightarrow{a.s.} 0, \\ II &\xrightarrow{a.s.} 0, \\ III &\xrightarrow{a.s.} A^*(1)B^*(1)\alpha_x\alpha_y(1-\tau_y). \end{aligned}$$

Thus,

$$T^{-1}\sum_1^T x_t y_t \xrightarrow{a.s.} A^*(1)B^*(1)\alpha_x\alpha_y(1-\tau_y). \tag{5.23}$$

A similar argument yields

$$T^{-1}\sum_1^T x_t \xrightarrow{a.s.} (1-\tau_x)A^*(1)\alpha_x, \tag{5.24}$$

$$T^{-1}\sum_1^T y_t \xrightarrow{a.s.} (1-\tau_y)B^*(1)\alpha_y. \tag{5.25}$$

2. Since

$$\begin{aligned} x_t^2 &= A^{*2}(1)\alpha_x^2\mathbf{1}\{t>[T\tau_x]\} + A^{*2}(1)u_t^2 + \Delta^2 u_t^* \\ &+ 2A^{*2}\alpha_x\mathbf{1}\{t>[T\tau_x]\}u_t + 2A^*(1)\Delta u_t^*\mathbf{1}\{t>[T\tau_x]\} + 2A^*(1)u_t\Delta u_t^* \end{aligned}$$

and

$$\begin{aligned} T^{-1}\sum_1^T u_t^2 &\xrightarrow{a.s.} \sigma_u^2, \\ T^{-1}\sum_1^T \Delta^2 u_t^* &\xrightarrow{a.s.} \Omega_{\Delta u^*}, \\ T^{-1}\sum_1^T u_t\Delta u_t^* &\xrightarrow{a.s.} \Omega_{u\Delta u^*} \end{aligned}$$

by Corollary 1 and 2, we have

$$T^{-1}\sum_{1}^{T} x_t^2 \xrightarrow{a.s.} A^{*2}(1)\sigma_u^2 + \Omega_{\Delta u^*} + \tau_x(1-\tau_x)A^{*2}(1)\alpha_x^2 + 2A^*(1)\Omega_{u\Delta u^*}. \tag{5.26}$$

A similar argument yields

$$T^{-1}\sum_{1}^{T} y_t^2 \xrightarrow{a.s.} B^{*2}(1)\sigma_v^2 + \Omega_{\Delta u^*} + \tau_y(1-\tau_y)B^{*2}(1)\alpha_y^2 + 2B^*(1)\Omega_{u\Delta u^*}. \tag{5.27}$$

3.

$$\begin{aligned}
T^{-1}\sum_{s+1}^{T} \hat{w}_t\hat{w}_{t-s} &= T^{-1}\sum_{s+1}^{T}(y_t-\overline{y})(y_{t-s}-\overline{y}) + \hat{\beta}^2 T^{-1}\sum_{s+1}^{T}(x_t-\overline{x})(x_{t-s}-\overline{x}) \\
&- \hat{\beta}T^{-1}\big[\sum_{s+1}^{T}(y_t-\overline{y})(x_{t-s}-\overline{x}) + \sum_{s+1}^{T}(y_{t-s}-\overline{y})(x_t-\overline{x})\big] \\
&= I + II - \hat{\beta}(III + IV).
\end{aligned}$$

First, we have

$$\begin{aligned}
I &= T^{-1}\sum_{s+1}^{T} y_t y_{t-s} + \overline{y}\left(T^{-1}\sum_{s+1}^{T} y_t + T^{-1}\sum_{s+1}^{T} y_{t-s}\right), \\
II &= T^{-1}\sum_{s+1}^{T} x_t x_{t-s} - \overline{x}\left(T^{-1}\sum_{s+1}^{T} x_t + T^{-1}\sum_{s+1}^{T} y_{t-s}\right),
\end{aligned}$$

$$\begin{aligned}
III &= T^{-1}\sum_{s+1}^{T} x_{t-s}y_t - \overline{x}\ \overline{y}, \\
IV &= T^{-1}\sum_{s+1}^{T} x_t y_{t-s} - \overline{x}\ \overline{y}.
\end{aligned}$$

An application of Corollary 2 yields

$$\begin{aligned}
T^{-1}\sum_{1}^{T} v_t v_{t+s} &\xrightarrow{a.s.} E[v_0 v_s], \\
T^{-1}\sum_{1}^{T} \Delta v_t^* \Delta v_{t+s}^* &\xrightarrow{a.s.} \Omega_{\Delta v^*}^*, \\
T^{-1}\sum_{1}^{T} v_t \Delta v_{t+s}^* &\xrightarrow{a.s.} \Omega_{1.v\Delta v^*}, \\
T^{-1}\sum_{1}^{T} v_{t+s} \Delta v_t^* &\xrightarrow{a.s.} \Omega_{2.v\Delta v^*}.
\end{aligned}$$

Therefore, we have

$$I \xrightarrow{a.s.} B^{*2}(1)\alpha_y^2\tau_y(1-\tau_y) + B^{*2}(1)E[v_0 v_s] + \Omega_{\Delta v^*}^* + B^*\left[\Omega_{1.v\Delta v^*} + \Omega_{2.v\Delta v^*}\right]. \tag{5.28}$$

A similar argument and Corollary 2 yield

$$II \xrightarrow{a.s.} A^{*2}(1)\alpha_x^2\tau_x(1-\tau_x) + B^{*2}(1)E[u_0 u_s] + \Omega_{\Delta u^*}^* + A^*(1)\left[\Omega_{1.u\Delta u^*} + \Omega_{2.u\Delta u^*}\right], \tag{5.29}$$

$$III \xrightarrow{a.s.} A^*(1)B^*(1)\alpha_x\alpha_y(1-\tau_y) - (1-\tau_x)(1-\tau_y)\alpha_x\alpha_y A^*(1)B^*(1), \tag{5.30}$$

$$IV \xrightarrow{a.s.} A^*(1)B^*(1)\alpha_x\alpha_y(1-\tau_x) - (1-\tau_x)(1-\tau_y)\alpha_x\alpha_y A^*(1)B^*(1). \tag{5.31}$$

Substituting (5.23)-(5.31) into (5.17)-(5.22) yields the main results.

Proof of Theorem 2

Eq. (3.2) is equivalent to

$$\begin{aligned}
\begin{bmatrix} T^{-1}\hat{\gamma}_1 \\ \hat{\beta} \\ \hat{\gamma}_2 \end{bmatrix} &= \begin{bmatrix} T & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \hat{\gamma}_1 \\ \hat{\beta} \\ \hat{\gamma}_2 \end{bmatrix} \\
&= \left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & T^{-1} & 0 \\ 0 & 0 & T^{-1} \end{bmatrix} \begin{bmatrix} T & \sum_1^T x_t & \sum_1^T t \\ \sum_1^T x_t & \sum_1^T x_t^2 & \sum_1^T tx_t \\ \sum_1^T t & \sum_1^T tx_t & \sum_1^T t^2 \end{bmatrix} \begin{bmatrix} T^{-1} & 0 & 0 \\ 0 & T^{-2} & 0 \\ 0 & 0 & T^{-2} \end{bmatrix} \right\}^{-1} \\
&\quad \begin{bmatrix} T^{-2} & 0 & 0 \\ 0 & T^{-3} & 0 \\ 0 & 0 & T^{-3} \end{bmatrix} \begin{bmatrix} \sum_1^T y_t \\ \sum_1^T x_t y_t \\ \sum_1^T t y_t \end{bmatrix} \\
&= \begin{bmatrix} 1 & T^{-2}\sum_1^T x_t & T^{-2}\sum_1^T t \\ T^{-2}\sum_1^T x_t & T^{-3}\sum_1^T x_t^2 & T^{-3}\sum_1^T tx_t \\ T^{-2}\sum_1^T t & T^{-3}\sum_1^T tx_t & T^{-3}\sum_1^T t^2 \end{bmatrix}^{-1} \begin{bmatrix} T^{-2}\sum_1^T y_t \\ T^{-3}\sum_1^T x_t y_t \\ T^{-3}\sum_1^T t y_t \end{bmatrix}. \qquad (5.32)
\end{aligned}$$

Moreover, we have

$$\begin{aligned}
T^{-1}\sigma^2_{\hat{\gamma}_1} &= \{T^{-2}\hat{s}^2\}\mathbb{I}_{(1)} \begin{bmatrix} 1 & T^{-2}\sum_1^T x_t & T^{-2}\sum_1^T t \\ T^{-2}\sum_1^T x_t & T^{-3}\sum_1^T x_t^2 & T^{-3}\sum_1^T tx_t \\ T^{-2}\sum_1^T t & T^{-3}\sum_1^T tx_t & T^{-3}\sum_1^T t^2 \end{bmatrix}^{-1} \mathbb{I}'_{(1)}, \qquad (5.33) \\
T\sigma^2_{\hat{\beta}} &= \{T^{-2}\hat{s}^2\}\mathbb{I}_{(2)} \begin{bmatrix} 1 & T^{-2}\sum_1^T x_t & T^{-2}\sum_1^T t \\ T^{-2}\sum_1^T x_t & T^{-3}\sum_1^T x_t^2 & T^{-3}\sum_1^T tx_t \\ T^{-2}\sum_1^T t & T^{-3}\sum_1^T tx_t & T^{-3}\sum_1^T t^2 \end{bmatrix}^{-1} \mathbb{I}'_{(2)}, \qquad (5.34) \\
T\sigma^2_{\hat{\gamma}_2} &= \{T^{-2}\hat{s}^2\}\mathbb{I}_{(3)} \begin{bmatrix} 1 & T^{-2}\sum_1^T x_t & T^{-2}\sum_1^T t \\ T^{-2}\sum_1^T x_t & T^{-3}\sum_1^T x_t^2 & T^{-3}\sum_1^T tx_t \\ T^{-2}\sum_1^T t & T^{-3}\sum_1^T tx_t & T^{-3}\sum_1^T t^2 \end{bmatrix}^{-1} \mathbb{I}'_{(3)}, \qquad (5.35)
\end{aligned}$$

and

$$\begin{aligned}\sum_{1+s}^{T} \hat{w}_{t-s}\hat{w}_t &= \sum_{1+s}^{T} y_t y_{t+s} + \hat{\beta}^2 \sum_{1+s}^{T} x_t x_{t+s} - \hat{\beta}\left(\sum_{1+s}^{T} x_t y_{t-s} + \sum_{1+s}^{T} x_{t-s} y_t\right) \\ &- \hat{\gamma}_2\left(\sum_{1+s}^{T} t y_{t+s} + \sum_{1+s}^{T}(t-s)y_t\right) - \hat{\gamma}_1\left(\sum_{1+s}^{T} y_t + \sum_{1+s}^{T} y_{t-s}\right) + \hat{\gamma}_1\hat{\beta}\left(\sum_{1+s}^{T} x_t + \sum_{1+s}^{T} x_{t-s}\right) \\ &+ \hat{\gamma}_2\hat{\beta}\left(\sum_{1+s}^{T}(t-s)x_t + \sum_{1+s}^{T} t x_{t-s}\right) + \hat{\gamma}_1\hat{\gamma}_2\sum_{1+s}^{T}(2t-s) + T\hat{\gamma}_1^2, \quad (5.36)\end{aligned}$$

$$\sum_{1}^{T} \hat{w}_t^2 \quad (= \quad \hat{s}^2) = \sum_{1}^{T}\left(y_t - \hat{\gamma}_1 - \hat{\gamma}_2 t - \hat{\beta}x_t\right)^2. \quad (5.37)$$

By applying Corollary 1 and Corollary 2, the terms involving the time averages of u_t, v_t, Δu_t^* and Δv_t^* when $t = 1, .., T$ converge almost sure at rate $O(1)$. Therefore, we obtain

$$\begin{aligned}T^{-2}\sum_{1}^{T} x_t &\xrightarrow{a.s.} \mu_x A^*(1)\int_0^{\tau_x} sds + A^*(1)(\mu_x + \mu_{xb})\int_{\tau_x}^1 sds \ (= a_{12}), \quad (5.38) \\ T^{-3}\sum_{1}^{T} x_t^2 &\xrightarrow{a.s.} \mu_x^2 A^{*2}(1)\int_0^{\tau_x} s^2 ds + A^{*2}(1)(\mu_x + \mu_{xb})^2\int_{\tau_x}^1 s^2 ds \\ &- \mu_{xb}A^*(1)\tau_x\int_{\tau_x}^1 sds \ (= a_{22}), \quad (5.39) \\ T^{-2}\sum_{1}^{T} t &\xrightarrow{a.s.} \int_0^1 sds \ (= a_{13}), \quad (5.40) \\ T^{-3}\sum_{1}^{T} t x_t &\xrightarrow{a.s.} \mu_x A^*(1)\int_0^{\tau_x} s^2 ds + A^*(1)(\mu_x + \mu_{xb})\int_{\tau_x}^1 s^2 ds \\ &- \mu_{xb}A^*(1)\tau_x\int_{\tau_x}^1 sds \ (= a_{23}), \quad (5.41) \\ T^{-3}\sum_{1}^{T} t^2 &\xrightarrow{a.s.} \int_0^1 s^2 ds \ (= a_{33}), \quad (5.42) \\ T^{-2}\sum_{1}^{T} y_t &\xrightarrow{a.s.} \mu_y B^*(1)\int_0^{\tau_y} sds + B^*(1)(\mu_y + \mu_{yb})\int_{\tau_y}^1 sds \ (= b_1), \quad (5.43)\end{aligned}$$

$$
\begin{aligned}
T^{-3}\sum_1^T x_t y_t \quad &\xrightarrow{a.s.} \quad \int_0^{\tau_x} s^2 ds + \mu_y(\mu_x+\mu_{xb})A^*(1)B^*(1)\int_{\tau_x}^{\tau_y} s^2 ds - \mu_{xb}\mu_y\tau_x A^*(1)B^*(1)\int_{\tau_x}^{\tau_y} sds \\
&+ \quad A^*(1)B^*(1)(\mu_x+\mu_{xb})(\mu_y+\mu_{yb})\int_{\tau_y}^1 s^2 ds - \mu_{yb}A^*(1)B^*(1)(\mu_x+\mu_{xb})\tau_y\int_{\tau_y}^1 sds \\
&- \quad \mu_{xb}A^*(1)B^*(1)(\mu_y+\mu_{yb})\tau_x\int_{\tau_y}^1 sds \ (=b_2), \qquad (5.44) \\
T^{-2}\sum_1^T ty_t \quad &\xrightarrow{a.s.} \quad \mu_y B^*(1)\int_0^{\tau_y} s^2 ds + B^*(1)(\mu_y+\mu_{yb})\int_{\tau_y}^1 s^2 ds - \mu_{yb}B^*(1)\tau_y\int_{\tau_y}^1 sds \ (=b_3), \qquad (5.45) \\
T^{-2}\sum_1^T y_t^2 \quad &\xrightarrow{a.s.} \quad \mu_y^2 B^{*2}(1)\int_0^{\tau_y} s^2 ds + B^{*2}(1)(\mu_y+\mu_{yb})^2\int_{\tau_y}^1 s^2 ds \\
&- \quad \mu_{yb}B^*(1)\tau_y\int_{\tau_y}^1 sds \ (=b_4). \qquad (5.46)
\end{aligned}
$$

By substituting (5.38)-(5.46) into (5.32)-(5.37), we obtain the results stated.

References

Ferson, W. E., Sarkissian, S. and Simin, T. T. (2003). Spurious regressions in financial economics, *Journal of Finance* **LVIII**(4): 1393–1413.

Granger, C. W. J., Hyung, H. and Jeon, Y. (2001). Spurious regressions with stationary series, *Applied Economics* **33**(7): 899–904.

Granger, C. W. J. and Hyung, N. (2004). Occasional structural breaks and long memory with an application to the s&p500 index absolute returns, *Journal of Empirical Finance* **11**: 399–421.

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

Lehmann, E. L. and Romano, J. P. (2005). *Testing Statistical Hypotheses*, third edn, Springer.

Marmol, F. (1995). Spurious regressions between i(d) processes, *Journal of Time Series Analysis* **16**(3): 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**(6): 1299–1325.

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

Tsay, W. J. and Chung, C. F. (1999). The spurious regression of fractionally integrated processes, *Journal of Econometrics* **96**(1): 155–182.

White, H. (1980). *Asymptotic Theory for Econometricians*, Academic Press.