CSCanada
Advances in Natural Science
Vol. 7, No. 4, 2014, pp. 1-7
DOI: 10.3968/6210

ISSN 1715-7862 [PRINT]
ISSN 1715-7870 [ONLINE]
www.cscanada.net
www.cscanada.org

Spurious Relationship of Long Memory Sequences in Presence of Trends Breaks

ZHANG Danshi[a],*; JIN Hao[a],[b]; ZHANG Dan[a]; CAO Wenhua[a]

[a]College of Science, Xi'an University of Science and Technology, Xi'an, China.
[b]College of Management, Xi'an University of Science and Technology, Xi'an, China.
*Corresponding author.

Supported by National Natural Science Foundation of China (71103143, 71473194); Natural Science Foundation of Shannxi Province (2013KJXX-40)

Received 30 September 2014; accepted 5 November 2014
Published online 26 December 2014

Abstract

This article extends the theoretical analysis of spurious relationship and considers the situation where the deterministic components of the processes generating the individual series are long memory sequences with structural changes. Show it by using the ordinary least squares estimator, the *t*-statistics become divergent and pseudo correlation. However, two long memory time series having change points can produce spurious regression. In the presence of structural change points, confirm the rate of t-statistic tends to infinity increased with the increase in sample size. Numerical simulation results show that when structural changes are a feature of the data, the presence of spurious relationship is unambiguous. And the spurious regression not only depends on long memory indexes, but also for trend of model is also very sensitive.

Key words: Spurious relationship; Long memory sequences; Structural changes; *t*-statistics; Numerical simulation

Zhang, D. S., Jin, H., Zhang, D., Cao, W. H. (2014). Spurious Relationship of Long Memory Sequences in Presence of Trends Breaks. *Advances in Natural Science*, *7*(4), 1-7. Available from: http://www.cscanada.net/index.php/ans/article/view/6210
DOI: http://dx.doi.org/10.3968/6210

INTRODUCTION

Many years ago, economists had been found that there exists spurious regression phenomenon in economic variables. However, in what circumstances will be spurious regression phenomenon, there was no unified recognition for a long time. As the early work of Monte Carlo study of Chow (1960) and Quant (1960), more and more effort has been taken to understand the nature of spurious regression and numerous studies have been undertaken an upsurge of interest in various models with an unknown change point. Issues about the distributional properties of the estimates in particular those of the break date, have been considered by Bai (1997). These tests and inference issue have also been addressed in the context of multiple change points in Bai and Perron (2003). Recent contributions include Gombay (2010), Jin et al. (2011), Beran and Shumeyko (2012) and Aue and Horváth (2013).

Moreover, the least squares estimates from the conventional spurious regression are inconsistent and have random limits. Numerical simulation has been used to study the spurious regression phenomena caused by the independent I (1) variables, such as Granger and Newbold (1974).Spurious relationships are shown to occur in random walks and linear trends by Durlauf and Phillips (1988). For the case of a random walk with drift, Entorf (1997) showed that t-statistic diverges to infinity as the sample size grows. Recently, Stewart (2006) argued that spurious relationship would generally occur in a regression of an I (1) dependent variable on an I (0) regressor, with or without another I (1) regressor. On the other hand, Noriega and Ventosa-Santaulària (2007) presented an analysis of the spurious phenomenon when there are a mix of deterministic and stochastic nonstationarity among the dependent and the explanatory variables in a linear regression model. More research in spurious relationship is appropriate include Kim and Lee (2011) and García-Belmonte.

However, there are more and more evidences showing that many economic and financial time series have the characteristics of long memory series. And some studies reveal that spurious regression do not only for independent random walks but also for long memory processes, such as Marmol (1998) and Tsay and Chung (2000). It demonstrated that spurious correlations are evident in regression involving combinations of long memory I (d) series, with d being a fractional number. However, it has been acknowledged that structural changes in the mean of time series can easily be confused with long memory dependence. Hence, the purpose of this paper is to investigate the possible existence of spurious relationship with a pair of stationary long memory processes with mis-specified change points, which could take place at different points in time.

1. THE MODELS AND ASSUMPTIONS

Similarly to these previous studies, we examine the regression, but assuming that the series are independent long memory processes with mis-specified change points. In order to investigate the spurious regression effects caused by structural changes, we specify the DGPs of three variables, and y_t,t=1...,T,as

$$x_t = \mu_x + \delta_x 1_{\{t>[T\tau_x]\}} + \varepsilon_t \ , \quad (1)$$

$$y_t = \theta_y t + \gamma_y(t - [T\tau_y])1_{\{t>[T\tau_y]\}} + \xi_t \ , \quad (2)$$

where x_t contain structural breaks in means, and y_t involves trends changes; μ_x and δ_x are, respectively, the permanent means and the transitory intercepts, resulting from a change, of the process x_t;θ_y and γ_y are , respectively, the permanent trends and the transitory trends, resulting from a change, of the process x_t ;τ_y and τ_y are change points of x_t and y_t; $1_{\{\cdot\}}$is the indicator function. These sequences ε_t and ξ_t are the stochastic part of the processes with long memory.

Before introducing our models, we would briefly review some basic properties of the long memory processes. The series ε_t is a fractionally integrated process of order, denoted as $I(d)$ if $(1\text{-}L)^d\varepsilon_t=\zeta_t$ is white noise with zero mean and finite variance, where L is the back-shift operator. The fractional differencing operator $(1\text{-}L)^d$ is defined as follows:

$$(1-L)^d = \sum_{j=0}^{\infty}\left(\prod_{k=1}^{j}\frac{k-1-d}{k}\right)L^j.$$

The $I(d)$ processes are natural generalization of the $I(1)$ processes that exhibit a broader long memory characteristics. The main feature of $I(d)$ process is that its autocovariance function declines at a slower hyperbolic rate (instead of the geometric rate found in the conventional ARMA models)

$$\gamma(j)=o(j^{2d-1}).$$

Where $\gamma(j)$ is the autocovariance function at lag . In particularly when 0<d<0.5, the stationary $I(d)$ process is said to have long memory since it exhibits long-range dependence in the sense that $\Sigma_{j=-\infty}^{\infty}\ \gamma(j)=\infty$ For -0.5<d<0, the process has short memory. We have independence or standard short memory when d=0. If 0.5<d<1, the process is nonstationary but still mean reverting. The relevant lines of research on long memory models may be found in Horváth and Kokoszka (2008) and McElroy and Politis (2011).

We focus on the case that ε_t is a stationary $I(d)$ process with 0<d<0.5, and require it obeying a functional central limit theorem, as stated in the lemma below. Let "$\Rightarrow$" and "$\xrightarrow{p}$" stand respectively for the weak convergence and convergence in probability. $B_d(\tau)$ denotes the fractional Brownian motion (of "type I" in the terminology of Marinucci and Robinson [1999]) with

$$\mathrm{B}_d(\tau) \equiv \frac{1}{\Gamma(1+d)}\left\{\int_0^{\tau}(\tau-s)^d W(s) + \int_{-\infty}^{0}[(\tau-s)^d - (-s)^d]dW(s)\right\} .$$

Where $\Gamma(\cdot)$ is the Gamma function and $W(s)$ is a standard Brownian motion.

Assumption 1.1 *The strictly stationary symmetrical innovations ε_t and ξ_t are assumed to be mean zero with long memory indexes d_ε, $d_\xi \in (0,0.5)$.*

Our method relies on the results derived by Davidson and Jong (2000).

Lemma 1.1 *If Assumption 1.1 holds, then as $T\rightarrow\infty$,*

$$\left(T^{-(0.5+d_\varepsilon)}\sum_{t=1}^{[T\tau]}\varepsilon_t \, , T^{-1}\sum_{t=1}^{T}\varepsilon_t^2\right) \Rightarrow \left(\kappa_\varepsilon B_{d_\varepsilon}(\tau), \sigma_\varepsilon^2\right) \ ,$$

$$\left(T^{-(0.5+d_\xi)}\sum_{t=1}^{[T\tau]}\xi_t \, , T^{-1}\sum_{t=1}^{T}\xi_t^2\right) \Rightarrow \left(\kappa_\xi B_{d_\xi}(\tau), \sigma_\xi^2\right),$$

Where $\tau\in[0,1]$ and κ_ε *and* κ_ξ are positive real numbers.

We need these results to derive the asymptotic validity of our linear relationship testing procedures. Furthermore, the fractional invariance principle could be extended to cover all possible long memory indexes d> -0.5, following Marinucci and Robinson (2000) who work with a "type II"fractional Brownian motion as limiting process.

Lemma 1.2 *Suppose x_t and y_t are respectively generated by (1) and (2) with break fraction τ_x,$\tau_y\in(0,1)$. The long memory series ε_t and ξ_t satisfy Assumption 1.1. Then as $T\rightarrow\infty$,*

a) $T^{-1}\sum x_t \Rightarrow \mu_x + (1-\tau_x)\delta_x \equiv \Gamma_x$

b) $T^{-1}\sum y_t \Rightarrow \frac{1}{2}[\theta_y + (1-\tau_y)^2\gamma_y] \equiv \mathit{\Gamma}_y$

c) $T^{-2}\sum tx_t \Rightarrow \frac{1}{2}[\mu_x + (1-\tau_x^2)\delta_x] \equiv \Gamma_{tx}$

d) $T^{-2}\sum ty_t \Rightarrow \frac{1}{6}[2\theta_y + (1-\tau_y)^2(2+\tau_y)\gamma_y \equiv \Gamma_{ty}$

e) $T^{-1}\sum x_t^2 \Rightarrow \mu_x^2 + (1-\tau_x)(2\mu_x\delta_x + \delta_x^2) + \sigma_\varepsilon^2 \equiv \Gamma_{xx}$

f) $T^{-1}\sum y_t^2 \Rightarrow \frac{1}{3}[\theta_y^2 + (1-\tau_y)^3\gamma_y^2 + (1-\tau_y)^2(2+\tau_y)\theta_y\gamma_y] \equiv \Gamma_{yy}$

g) $T^{-1}\sum x_t y_t \Rightarrow \frac{1}{2}[\mu_x\theta_y + \delta_x\theta_y(1-\tau_x)^2 + \mu_x\gamma_y(1-\tau_y)^2 + \delta_x\gamma_y((1-\tau_M)^2 + \tau_D(1-\tau_M))] \equiv \Gamma_{xy}$

Proof From the GDP (1), to prove that item 1, we have:

$$\sum x_t = \sum \mu_x + \sum \delta_x 1_{\{t>[T\tau_x]\}} + \sum \varepsilon_t,$$

where $\Sigma\mu_x = T\mu_x$ and $\Sigma\delta_x 1_{\{t>[T\tau x]\}} = T(1-\tau_x)\delta_x$. In the view of $\Sigma\varepsilon_\tau = O_p(T^{0.5+d_\varepsilon})$ and $0<d_\varepsilon<0.5$, then

$$T^{-1}\sum x_t \Rightarrow \mu_x + (1-\tau_x)\delta_x \equiv \Gamma_x.$$

To prove item 3, note that

$$\Sigma t x_t = \Sigma t\mu_x + \Sigma t\delta_x 1_{\{t>[T\tau_x]\}} + \Sigma t\varepsilon_t,$$

where $\Sigma t\mu_x = \frac{1}{2}(T^2+T)\mu_x$ and $\Sigma t\delta_x 1_{\{t>[T\tau_x]\}} = ((1-\tau_x^2)T^2 + T(1-\tau_x))\delta_x$. The result of Lemma 2.1 supplements that $\Sigma t\varepsilon_t = O_p(T\cdot T^{0.5+d_\varepsilon})$, one obtains

$$T^{-2}\sum t x_t \Rightarrow \frac{1}{2}[\mu_x + (1-\tau_x^2)\delta_x] \equiv \Gamma_{tx}.$$

It is straightforward to prove that

$$\sum [\mu_x + \delta_x 1_{\{t>[T\tau_x]\}}]\cdot \varepsilon_t = O_p(T^{0.5+d_\varepsilon}),$$

and

$$\sum [\mu_x + \delta_x 1_{\{t>[T\tau_x]\}}]^2 = T(\mu_x^2 + (1-\tau_x)(2\delta_x\mu_x + \delta_x^2)).$$

These results, together with ,imply

$$T^{-1}\sum x_t^2 = T^{-1}\sum\left\{[\mu_x + \delta_x 1_{\{t>[T\tau_x]\}}]^2 + 2(\mu_x + \delta_x 1_{\{t>[T\tau_x]\}}\varepsilon_t + \varepsilon_t^2)\right\}$$

$$\Rightarrow \mu_x^2 + (1-\tau_x)(2\mu_x\delta_x + \delta_x^2) + \sigma_\varepsilon^2 \equiv \Gamma_{xx}.$$

Therefore, the proof of item 5 is completed.

To avoid any cumbersome mathematical expression, we summarize only the terms of order $O_p(1)$ as follows. Note that

$$\sum (t-[T\tau_y])1_{\{t>[T\tau_y]\}} = \sum_1^{(1-\tau_y)T} t = \frac{1}{2}\left[(1-\tau_y)^2T^2 + (1-\tau_y)T\right]$$

and

$$\sum \xi_t = O_p(T^{0.5+d_\xi}),$$

From the DGP (2), one could have

$$T^{-2}\sum y_t = T^{-2}\sum\left\{\theta_y t + \gamma_y(t-[T\tau_y])1_{\{t>[T\tau_y]\}} + \xi_t\right\} + o_p(1)$$

$$\Rightarrow \frac{1}{2}\left[\theta_y + (1-\tau_y)^2\gamma_y\right] \equiv \Gamma_y.$$

Therefore, the proof of item 2 is completed.

In the view of , $\Sigma t\xi_t = O_p(T\cdot T^{0.5+d_\xi})$, one can obtain that

$$T^{-3}\sum t y_t = T^{-3}\sum\left\{\theta_y t^2 + \gamma_y t(t-[T\tau_y])1_{\{t>[T\tau_y]\}} + t\xi_t\right\} + o_p(1)$$

$$\Rightarrow \frac{1}{6}\left[2\theta_y + (1-\tau_y)^2(2+\tau_y)\gamma_y\right] \equiv \Gamma_{ty}.$$

Note that $\Sigma\xi^2_t = O_p(T)$, one should have

$$T^{-3}\sum y_t^2 = T^{-3}\sum\left\{\theta_y^2 t^2 + [\gamma_y^2(t-[T\tau_y])^2 + 2\theta_y\gamma_y t(t-[T\tau_y])1_{\{t>[T\tau_y]\}}\right\} + o_p(1)$$

$$\Rightarrow \frac{1}{3}\left[\theta_y^2 + (1-\tau_y)^3\gamma_y^2 + (1-\tau_y)^2(2+\tau_y)\theta_y\gamma_y\right] \equiv \Gamma_{yy}.$$

Thus the proof of item 6 follows.

Now it remains to prove item 9. We assume $\tau_x<\tau_y$, then the last part can be proven as follows:

$$\sum x_t y_t = \sum\left\{\mu_x\theta_y t + \delta_x\gamma_y 1_{\{t>[T\tau_x]\}}(t-[T\tau_y])1_{\{t>[T\tau_y]\}}\right\}$$

$$+\sum\left\{\mu_x\gamma_y 1_{\{t>[T\tau_x]\}}(t-[T\tau_y])1_{\{t>[T\tau_y]\}} + \delta_x\theta_y t 1_{\{t>[T\tau_x]\}}\right\} + O_p(T^2)$$

Therefore:

$$T^{-1}\sum x_t y_t$$

$$\Rightarrow \frac{1}{2}\Big[\mu_x\theta_y + \delta_x\theta_y(1-\tau_x)^2 + \mu_x\gamma_y(1-\tau_y)^2 + \delta_x\gamma_y$$

$$\left((1-\tau_M)^2 + \tau_D(1-\tau_M)\right)\Big] \equiv \Gamma_{xy}.$$

Where, $\tau_M=\max(\tau_x,\tau_y)$, $\tau_m=\min(\tau_x,\tau_y)$, and $\tau_D=\tau_M-\tau_m$.

Hence, the proof of Lemma 2.2 is completed.

2. MAIN RESULTS

We consider the regression model given by

$$yt=\alpha+\varphi t+\beta x_t+\eta_t, \qquad (3)$$

where x_t and y_t are respectively the regressand and regressor, and μ_t is the regression error. Let $\hat{\alpha}$, $\hat{\varphi}$ and $\hat{\beta}$ denote the ordinary least squares estimates from a regression of y_t on a constant,the trend t and x_t respectively. Their respective '*variances*' are estimated by $s^2_{\hat{\alpha}}$, $s^2_{\hat{\varphi}}$ and $s^2_{\hat{\beta}}$, from which we have the t-ratios $t_{\hat{\alpha}}=\hat{\alpha}/s_{\hat{\varphi}}$, and $t_{\hat{\beta}}=\hat{\beta}/s_{\hat{\beta}}$.

As stated the following theorems, we could denote $\tau_M=\max\{\tau_x,\tau_y\}$ and $\tau_D=\tau_M-\tau_y$. The proofs of these theorems are omitted since they are similar to the line of proofs in Tasy (2000). But full proof is available on request.

The next theorem interprets the asymptotic behavior of the estimated parameters, associated

t-ratios and R^2 in model (3).

Theorem 2.1 *Suppose x_t and y_t are respectively generated by (1) and (2) with break fraction $\tau_x,\tau_y\in(0,1)$. The long memory series ε_t and ξ_t satisfy Assumption 2.1, then as $T\to\infty$*

a) $T^{-1}\hat{\alpha} \Rightarrow D_X^{-1}\left\{\left(\frac{1}{3}\Gamma_{xx} - \Gamma_{tx}^2\right)\Gamma_y + \left(\Gamma_x\Gamma_{tx} - \frac{1}{2}\Gamma_{xx}\right)\Gamma_{ty} + \left(\frac{1}{2}\Gamma_{tx} - \frac{1}{3}\Gamma_x\right)\Gamma_{xy}\right\} \equiv \alpha,$

b) $\hat{\varphi} \Rightarrow D_X^{-1}\cdot\left\{\left(\Gamma_x\Gamma_{tx} - \frac{1}{2}\Gamma_{xx}\right)\Gamma_y + (\Gamma_{xx} - \Gamma_x^2)\Gamma_{tx} + \left(\frac{1}{2}\Gamma_x - \Gamma_{tx}\right)\Gamma_{xy}\right\} \equiv \varphi,$

c) $T^{-1}\hat{\beta} \Rightarrow D_X^{-1}\cdot\left\{\left(\frac{1}{2}\Gamma_{tx} - \frac{1}{3}\Gamma_x\right)\Gamma_y + \left(\frac{1}{2}\Gamma_x - \Gamma_{tx}\right)\Gamma_{ty} + \frac{1}{12}\Gamma_{xy}\right\} \equiv \beta,$

d) $T^{-1/2}t_{\hat{\alpha}} \Rightarrow \alpha\cdot\left(\sigma^2\cdot\frac{\frac{1}{3}\Gamma_{xx}-\Gamma_{tx}^2}{D_X}\right)^{-1/2},$

e) $T^{-1/2}t_{\hat{\varphi}} \Rightarrow \varphi\cdot\left(\sigma^2\cdot\frac{\Gamma_{xx}-\Gamma_x^2}{D_X}\right)^{-1/2},$

f) $T^{-1/2}t_{\hat{\beta}} \Rightarrow \beta\cdot\left(\frac{\sigma^2}{12D_X}\right)^{-1/2},$

g) $R^2 \Rightarrow \beta^2\cdot\frac{\Gamma_{xx}-\Gamma_x^2}{\Gamma_{yy}-\Gamma_y^2},$

where

$$D_X = \Gamma_x\Gamma_{tx} - \tfrac{1}{3}\Gamma_x^2 - \Gamma_{\text{tx}}^2 + \tfrac{1}{12}\Gamma_{xx} \ ,$$

$$\sigma^2 = \Gamma_{yy} + \alpha^2 + \tfrac{1}{3}\varphi^2 + \beta^2\Gamma_{xx} - 2\big(\alpha\Gamma_y + \varphi\Gamma_{ty} + \beta\Gamma_{xy}\big) + \alpha\varphi + 2\alpha\beta\Gamma_x + 2\varphi\beta\Gamma_{tx}.$$

Proof Write the regression model $y_t=\alpha+\varphi t+\beta x_t+\eta_t$ inmatrix from:

$$Y=X\theta+U.$$

The vector of OLS estimators is $\hat{\theta}=(X'X)^{-1}X'Y$, and we define

$$X'X = \begin{Bmatrix} \sum 1 & \sum t & \sum x_t \\ \sum t & \sum t^2 & \sum tx_t \\ \sum x_t & \sum tx_t & \sum x_t^2 \end{Bmatrix} = \begin{Bmatrix} a & b & c \\ b & d & e \\ c & e & f \end{Bmatrix}.$$

Hence, one could rewrite

$$\hat{\theta} = [det(X'X)]^{-1} \begin{Bmatrix} df-e^2 & ce-bf & be-cd \\ ce-bf & af-c^2 & bc-ae \\ be-cd & bc-ae & ad-b^2 \end{Bmatrix} \begin{Bmatrix} \sum y_t \\ \sum ty_t \\ \sum x_t y_t \end{Bmatrix}$$

Where

$$\det(X'X) = 2bce + adf - c^2d - ae^2 - b^2$$
$$= 2\sum t\sum x_t\sum tx_t + T\sum t^2\sum x_t^2 - \sum t^2\Big(\sum x_t\Big)^2 - T\Big(\sum tx_t\Big)^2 - \Big(\sum t\Big)^2\sum x_t^2.$$

Using the results from Lemma 2.2, one could get

$$T^{-5}\cdot \det(X'X) \Rightarrow \Gamma_x\Gamma_{tx} - \frac{1}{3}\Gamma_x^2 - \Gamma_{tx}^2 + \frac{1}{12}\Gamma_{xx} \equiv D_X.$$

To prove that a) - c), one can derive expression for $\hat{\alpha}$, $\hat{\varphi}$ and $\hat{\beta}$ as follows:

$$\hat{\alpha} = [\det(X'X)]^{-1}\cdot\Big\{(df-e^2)\sum y_t + (ce-bf)\sum ty_t + (be-cd)\sum x_t y_t\Big\}$$
$$\Rightarrow D_X^{-1}\cdot\Big\{\Big(\tfrac{1}{3}\Gamma_{xx} - \Gamma_{tx}^2\Big) + \Big(\Gamma_x\Gamma_{tx} - \tfrac{1}{2}\Gamma_{xx}\Big)\Gamma_{ty} + \Big(\tfrac{1}{2}\Gamma_{tx} - \tfrac{1}{3}\Gamma_x\Big)\Gamma_{xy}\Big\} \equiv \alpha \ ,$$

$$T\hat{\varphi} = [\det(X'X)]^{-1}\cdot T\Big\{(ce-bf)\sum y_t + (af-c_2)\sum ty_t + (bc-ae)\sum x_t y_t\Big\}$$
$$\Rightarrow D_X^{-1}\cdot\Big\{\Big(\Gamma_x\Gamma_{tx} - \tfrac{1}{2}\Gamma_{xx}\Big)\Gamma_y + (\Gamma_{xx} - \Gamma_x^2)\Gamma_{ty} + \Big(\tfrac{1}{2}\Gamma_x - \Gamma_{tx}\Big)\Gamma_{xy}\Big\} \equiv \varphi \ ,$$

$$\hat{\beta} = [\det(X'X)]^{-1}\cdot\Big\{(be-cd)\sum y_t + (bc-ae)\sum ty_t + (ad-b^2)\sum x_t y_t\Big\}$$
$$\Rightarrow D_X^{-1}\cdot\Big\{\Big(\tfrac{1}{2}\Gamma_{tx} - \tfrac{1}{3}\Gamma_x\Big)\Gamma_y + \Big(\tfrac{1}{2}\Gamma_x - \Gamma_{tx}\Big)\Gamma_{ty} + \tfrac{1}{12}\Gamma_{xy}\Big\} \equiv \beta \ .$$

To prove d) - f), one could write the *t*-statistics as:

$$t_{\hat{\alpha}} = \hat{\alpha}\cdot[s^2\cdot(X'X)_{11}^{-1}]^{-1/2} \ ,$$
$$t_{\hat{\varphi}} = \hat{\varphi}\cdot[s^2\cdot(X'X)_{22}^{-1}]^{-1/2} \ ,$$
$$t_{\hat{\beta}} = \hat{\beta}\cdot[s^2\cdot(X'X)_{33}^{-1}]^{-1/2} \ .$$

Where

$$s^2 = T^{-1}\sum\big(y_t - \hat{\alpha} - \hat{\varphi}t - \hat{\beta}x_t\big)^2$$
$$= T^{-1}\sum\big\{y_t^2 + \hat{\alpha}^2 + \hat{\varphi}^2t^2 + \hat{\beta}^2x_t^2 - 2\hat{\alpha}y_t - 2\hat{\varphi}ty_t - 2\hat{\beta}x_ty_t\big\}$$
$$+ T^{-1}\sum\big\{2\hat{\alpha}\hat{\varphi}t + 2\hat{\alpha}\hat{\beta}x_t + 2\hat{\varphi}\hat{\beta}tx_t\big\}$$
$$\Rightarrow \Gamma_{yy} + \alpha^2 + \frac{1}{3}\varphi^2 + \beta^2\Gamma_{xx} - 2\big(\alpha\Gamma_y + \varphi\Gamma_{ty} + \beta\Gamma_{xy}\big) + \alpha\varphi + 2\alpha\beta\Gamma_x + 2\varphi\beta\Gamma_{tx} \equiv \sigma_u^2.$$

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

$$T\cdot(X'X)_{11}^{-1} = \frac{T^{-4}\cdot(\sum t^2\sum x_t^2 - \sum tx_t\sum tx_t)}{T^{-5}\cdot\det(X'X)} \Rightarrow \frac{\frac{1}{3}\Gamma_{xx} - \Gamma_{tx}^2}{D_X} \ ,$$

$$T^3\cdot(X'X)_{22}^{-1} = \frac{T^{-2}\cdot(\sum 1\sum x_t^2 - \sum x_t\sum x_t)}{T^{-5}\cdot\det(X'X)} \Rightarrow \frac{\Gamma_{xx} - \Gamma_x^2}{D_X} \ ,$$

$$T\cdot(X'X)_{33}^{-1} = \frac{T^{-4}\cdot(\sum 1\sum t^2 - \sum t\sum t)}{T^{-5}\cdot\det(X'X)} \Rightarrow \frac{1}{12D_X} \ .$$

Then, one can show that

$$T^{-1/2}\cdot t_{\hat{\alpha}} = \hat{\alpha}\cdot[s^2\cdot T(X'X)_{11}^{-1}]^{-1/2} \Rightarrow \alpha\cdot\left(\sigma_u^2\cdot\frac{\frac{1}{3}\Gamma_{xx} - \Gamma_{tx}^2}{D_X}\right)^{-1/2},$$

$$T^{-1/2}\cdot t_{\hat{\varphi}} = \hat{\varphi}\cdot[s^2\cdot T^3(X'X)_{22}^{-1}]^{-1/2} \Rightarrow \varphi\cdot\left(\sigma_u^2\cdot\frac{\Gamma_{xx} - \Gamma_x^2}{D_X}\right)^{-1/2},$$

and

$$T^{-1/2}\cdot t_{\hat{\beta}} = \hat{\beta}\cdot[s^2\cdot T(X'X)_{33}^{-1}]^{-1/2} \Rightarrow \beta\cdot\left(\frac{\sigma_u^2}{12D_X}\right)^{-1/2}.$$

The limits of these expressions yield the formula in d)-f). Similarly, g) can be proved that

$$R^2 = \hat{\beta}\frac{T^{-1}\sum x_t^2 - (T^{-1}\sum x_t)^2}{T^{-1}\sum y_t^2 - (T^{-1}\sum y_t)^2} \Rightarrow \beta^2\cdot\frac{\Gamma_{xx} - \Gamma_x^2}{\Gamma_{yy} - \Gamma_y^2}.$$

Hence, the proof of Theorem 2.2 is completed.

3. SIMULATION

This section adopts the method of Monte Carlo simulation to verify the correctness of the theoretical derivation. In the simulated data, the sample size respectively by T=100, 500, 1,000, 10,000. For each simulated sample, the percentage of rejection are obtained by 1, 000 replications at 5% nominal level, i.e., the percentage of *t*-ratios such that $|t|>1.96$.

We consider the process with mean change points (1) and with trend change points process (2),

$$x_t = \mu_x + \delta_x 1_{\{t>[T\tau_x]\}} + \varepsilon_t,$$

$$y_t = \theta_y t + \gamma_y\left(t - \left[T\tau_y\right]\right)1_{\{t>[T\tau_y]\}} + \xi_t,$$

Where, μ_x=0.8, θy=0.3 we give δ_x=0.6, γ_y=0.2, other cases have similar results. Information process of ε_t and ξ_t are independent of each other, the long memory index d_ε and d_ξ adopted the value in {0.1, 0.2, 0.3, 0.4}.

Table 1
Regressing Between Two Long Memory Sequences With Structural Changes, $|t_{\hat\beta}|>1.96$

d_ε=0.2		τ_x=0		τ_x=0.1		τ_x=0.5		τ_x=0.9		τ_x=1	
τ_y	T/d_ξ	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4
0	100	5.86	6.70	5.26	5.98	5.52	5.56	5.54	6.08	6.26	6.06
	500	5.54	4.98	5.24	5.38	4.84	4.58	5.58	5.46	5.36	5.10
	1000	5.08	5.24	5.08	5.28	5.68	4.78	5.10	4.94	5.66	5.56
	5000	4.76	4.96	4.62	4.52	4.86	5.34	5.14	4.68	5.06	5.44
0.1	100	5.52	7.68	8.04	10.08	6.24	5.56	5.76	5.78	5.28	6.20
	500	6.16	6.44	68.92	60.48	5.42	5.22	5.30	4.80	6.58	5.14
	1000	5.30	5.88	96.18	95.90	4.82	12.96	9.62	7.88	14.18	5.44
	5000	4.80	5.44	100	100	5.02	69.70	45.28	45.80	69.30	5.24
0.5	100	5.92	8.04	12.88	14.48	6.16	5.52	7.38	5.62	5.18	5.56
	500	5.58	5.48	45.26	45.48	5.34	5.26	38.06	37.12	4.86	4.98
	1000	5.40	4.98	73.46	73.26	5.08	4.96	66.66	66.90	5.34	5.40
	5000	5.08	5.02	100	100	4.60	4.42	100	100	4.80	5.44
0.9	100	5.84	7.32	5.62	7.16	5.40	6.64	7.94	5.24	5.74	5.56
	500	5.44	5.80	8.36	8.64	4.80	10.44	61.14	46.60	10.48	5.28
	1000	5.12	4.78	13.34	13.46	4.84	19.00	94.96	91.48	19.50	4.84
	5000	5.04	5.22	50.00	50.08	4.60	73.74	100	100	73.40	5.14
1	100	5.30	6.28	5.34	5.86	5.10	5.28	5.42	5.84	5.44	6.18
	500	5.26	4.88	5.24	4.84	5.34	4.80	5.66	4.98	5.32	5.52
	1000	4.74	4.98	5.40	5.10	5.04	4.56	5.12	5.20	5.00	5.28
	5000	5.32	5.46	5.50	4.98	5.32	5.24	5.32	5.00	4.60	4.86

In order to study the asymptotic properties of the *t*-statistics. We set the change point of $\tau_x,\tau_y \in$ {0,0.1 0.5 0.9 1}, When $\tau_x=\tau_y$=0 and $\tau_x=\tau_y$=1, x_t and y_t series doesn't exist change points. Table 1-2 provide the percentage of rejection of $t_{\hat\beta}$ convergence for each change points and long memory indexes profile, on the basis of the combination of change point and long memory index. We denote Mean-Trend representative of sequences generated by (x_t;y_t) in models (1) and (2). We can see the *t*-statistics rejection rate increased with the increase in sample size *T*. However, Table 1-2 show that, if there are no changes, the rejection frequency of $t_{\hat\beta}$ does not support the existence of spurious relationship even for large long memory indexes and sample size. For example, when $d_\varepsilon=d_\xi$=0.2 and $\tau_x=\tau_y$=0, the rejection rate are 5.86%, 5.54%, 5.08% and 4.76% for *T*=100, 500, 1000, 5000. Particularly, if one of these two series does not exist change points, the rejection power is always closed to its nominal level. What is surprising is that, the spurious relationship are very strong if the observed sequence is driven by a linear trend. And the rejection frequency significantly increases as the sample size growing. However, it is not true for $\tau_x=\tau_y$=0.5 case, its rejection frequency is closed to the nominal level. Finally, the Spurious regression relationship still is sensitive to the long memory index d_ε, but its effects are negligible for long memory index d_ξ. For example, when $\tau_x=\tau_y$=0.1, d_ξ=0.2 and *T*=1,000, the rejection rate are 96.18% and 87.16% for d_ε=0.2, 0.4. When if $\tau_x=\tau_y$=0.1, d_ε=0.4 and T=1,000 the rejection frequency are 87.16% and 86.8% for d_ξ=0.2, 0.4. Therefore, the presence of spurious regression in Mean-Trend case is dominated by the long memory index d_ε.

Table 2
Regressing Between Two Long Memory Sequences With Structural Changes, $|t_{\hat{\beta}}|$>1.96

d_ε=0.4		τ_x=0		τ_x=0.1		τ_x=0.5		τ_x=0.9		τ_x=1	
τ_y	T/d_ξ	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4	0.2	0.4
0	100	4.98	9.22	4.44	8.56	5.00	8.24	5.20	8.56	4.62	8.58
	500	4.74	5.74	4.86	6.04	5.38	5.46	5.06	5.82	5.36	5.10
	1000	5.20	5.30	4.56	5.72	5.64	5.20	5.50	5.60	5.00	5.84
	5000	5.16	5.64	4.88	5.50	5.06	4.50	5.04	5.34	4.78	5.26
0.1	100	6.66	16.54	7.86	19.64	5.22	12.12	5.56	13.98	6.40	14.82
	500	9.68	12.68	60.14	57.98	4.46	4.58	4.92	7.04	9.22	10.86
	1000	7.70	9.60	87.16	86.80	5.52	5.20	4.42	4.78	7.60	8.24
	5000	5.66	5.64	100	100	37.02	36.98	22.16	21.32	5.84	5.64
0.5	100	6.42	14.44	12.04	23.36	6.54	12.34	4.58	5.84	6.00	12.96
	500	4.86	6.40	32.84	33.90	5.42	6.02	16.84	16.00	5.52	6.24
	1000	5.92	5.28	51.58	51.72	4.52	5.06	36.26	35.94	4.46	5.16
	5000	4.48	6.02	98.68	98.68	4.76	5.14	97.84	97.68	5.00	4.72
0.9	100	4.30	11.18	4.62	10.86	4.48	10.52	5.72	6.44	4.52	8.62
	500	4.48	6.02	7.08	8.80	8.22	10.26	35.44	20.44	4.90	5.88
	1000	5.54	5.42	9.52	10.78	14.06	14.54	72.48	63.66	5.06	5.20
	5000	4.38	5.22	31.32	31.24	48.30	47.84	100	100	4.60	4.30
1	100	4.32	9.48	5.02	9.48	5.06	8.30	4.90	8.78	4.56	9.60
	500	5.50	6.38	4.94	5.90	5.06	5.44	5.14	5.74	5.10	5.50
	1000	5.14	5.08	5.20	5.14	5.24	5.60	4.84	5.46	5.34	5.42
	5000	4.80	5.08	5.46	5.46	4.94	4.52	4.96	4.64	4.88	5.14

To give intuitive idea for the influence of stable indexes and breaks fractions, we provide the rejection frequency with sample size T=1,000 in Figures 1-2. As expected, Figures 1-2 clearly show that spurious relationship is present when the long memory sequences contains breaks. There are two peaks, and at the same time there are high rejection power. On the other hand, there are higher rejection frequency occurs near $\tau_x=\tau_y=0.9$ than other cases; while the rejection rate is the lowest at the center and edges of the sample. We can get a conclusion that the smaller long memory index provide higher rejection rate.

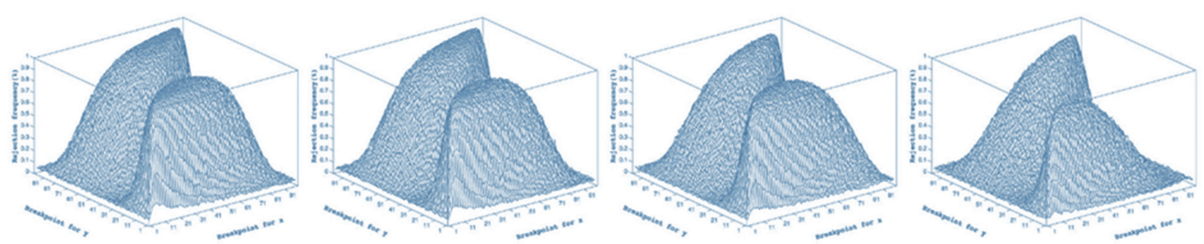

Figure 1
The Rejection Frequency as T = 1,000, d_ξ=0.2 and d_ε=0.1,0.2,0.3,0.4, Respectively

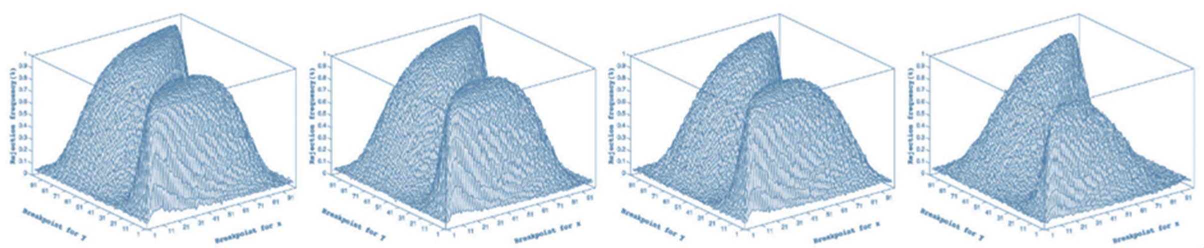

Figure 2
The Rejection Frequency as T = 1,000, d_ξ=0.4 and d_ε=0.1,0.2,0.3,0.4, Respectively

In general, it is clear from these simulations that if the sequence undergoes the structural breaks, the spurious relationship is present.

CONCLUSION

In this paper, we use the least squares estimation to study the spurious relationship of long memory time series. Our model includes two different model, mean model and trend model. In the presence of structural change points, confirm the rate of t-statistic tends to infinity increased with the increase in sample size. Numerical simulation results show that when structural changes are a feature of the data, the presence of spurious relationship is unambiguous. The presence of spurious relationship will always be more severe if the sequences involve multiplies breaks. And the spurious regression not only depends on long memory indexes, but also for trend of model and the sample size are also very sensitive.

REFERENCES

Aue, A., & Horváth L. (2013). Structural breaks in time series. *Journal of Time Series Analysis, 34,* 1-16.

Bai, J. (1997). Estimation of a change point in multiple regression models. *Time Series Analysis, 15,* 453-472.

Bai, J., & Perron, P. (2003). Computation and estimation of multiple structural change models. *Applied Econometrics, 18,* 1-22.

Beran, J., & Shumeyko, F. (2012). Bootstrap testing for discontinuities under long-range dependence. *Journal of Multivariate Analysis, 105,* 322-347.

Chow, G. C., (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica, 28,* 591-605.

Davidson, J., & de Jong, R. M. (2000). The functional central limit theorem and weak convergence to stochastic integral II: fractionally integrated processes. *Economic Theory, 16,* 643-666.

Durlauf, S. N., & Phillips, P. C. B. (1988). Trends versus random walks in time series analysis. *Econometrica, 56,* 1333-1354.

Entorf, H., 1997. Random walks with drifts: Nonsense regression and spurious fixed-effect estimation. *Journal of Econometrics, 80,* 287-296.

García-Belmonte, L., & Venttosa-Santaulμaria, D. (2011). Spurious regression and lurking variables. *Stat. Probab. Lett, 81,* 2004-2011.

Gombay, E. (2010). Change detection in linear regression with time series errors. *The Canadian Journal of Statistics, 38,* 65-79.

Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics, 2,* 111-120.

Horváth, L., & Kokoszka, P. (2008). Sample autocovariances of long memory time series. *Bernoull, 14,* 405-418.

Jin, H., & Zhang, J. (2011). Modified tests for variance changes in autoregressive regression. *Mathematics and Computers in Simulation, 81,* 1099-1109.

Kim, C. S., & Lee, S. (2011). Spurious regressions driven by excessive volatility. *Econom. Lett*, 292-297.

Marmol, F. (1998). Spurious regression theory with nonstationary fractionally integrated processes. *Journal of Econometrics, 84,* 223-250.

Mayoral, L. (2011). Testing for fractional integration versus short memory with trends and structural breaks. *Oxford Bulletin of Economics and Statistics, 74,* 278-305.

McElroy, T., & Politis, D. N. (2011). Fixed-b asymptotic for the studentized mean from time series with short, long or negative memory. *Econometric Theory, 28,* 1-11.

Noriega, A. E., & Ventosa-Santaulμaria, D. (2006). Spurious regression under broken-trend stationarity. *Journal of Time Series Analysis, 27,* 671-684.

Quandt, R. E., (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *American Statistical Association, 50,* 320-330.

Robinson, P. M., & Robinson, P. M. (1999). Alternative forms of fractional Brownian motion. *Journal of Statistical Planning and Inference, 80,* 111-122.

Robinson, P. M., & Robinson, P. M. (2000). Weak convergence of multivariate fractional processes. *Stochastic Processes and Their Applications, 86,* 103-120.

Stewart, C. (2006). Spurious correlation of $I(0)$ regressors in regressions with an $I(1)$ dependent variable. *Econom.Lett, 91,* 184-189.

Tsay, W.-J. & Chung, C.-F. (2000). The spurious regression of fractionally integrated processes. *Journal of Economics, 96,* 155-182.

