CSCanada
International Business and Management
Vol. 11, No. 2, 2015, pp. 69-73
DOI:10.3968/7627

ISSN 1923-841X [Print]
ISSN 1923-8428 [Online]
www.cscanada.net
www.cscanada.org

The Spurious Regression of Fractionally Integrated Processes With Change Points

ZHANG Si[a],*; SUN Rui[a]

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

Received 22 July 2015; accepted 26 September 2015
Published online 31 October 2015

Abstract

This paper considers the situation where two independent fractionally integrated processes in presence of change points are used in a linear spurious regression model. By using ordinary least squares regression, the order in probability of t-ratio is derived and its asymptotic distribution is not invariant to shifts in level. Simulation results indicate that, the extent of spurious regression not only seriously depends on the fractionally integrated indexes and sample size, but also is sensitive to the relative location of change points with the sample.

Key words: Spurious regression; Fractionally integrated processes; Change points; Linear spurious regression model

Zhang, S., & Sun, R. (2015). The Spurious Regression of Fractionally Integrated Processes With Change Points. *International Business and Management, 11*(2), 69-73. Available from: http://www.cscanada.net/index.php/ibm/article/view/7627
DOI: http://dx.doi.org/10.3968/7627

INTRODUCTION

The spurious regression phenomenon takes place when there is a statistically significant relationship between two independent fractionally integrated processes. Since the contribution of Granger and Newbold (1974) on the issue of spurious regressions in econometrics, several articles have investigated the phenomenon under a variety of structures for the data generating processes. Phillips (1986) assumes that the individual series in a spurious regression are driftless random walks, while Marmol (1995) extends this to the general $I(d)$ case, with d being an integer number. Marmol (1996) studies the spurious regression problem with different orders of integration of the dependent and independent variables. Entorf (1997) analyzes random walks with drifts, given the relevance of such models, as argued by Nelson and Plosser (1982). Granger, Hyung and Jeon (2001) extend the analysis to positively autocorrelated autoregressive series on long moving averages, and Marmol (1998) and Tsay and Chung (2000) to long memory fractionally integrated processes. For more details about asymptotic properties of estimates and test statistics for regression coefficients of these spurious regression, we refer the reader to McCallum (2010) and Martinez-Rivera and Ventosa-Santaularia (2012) among others.

This paper explores the possible existence of spurious relationship between two independent fractionally integrated processes. The popular fractionally integrated processes are capable of capturing long-range dependence, as well as a periodic cyclical pattern. It has been shown that the autocovariances of fractionally integrated processes decay hyperbolically and sinusoidally, a feature that is manifested in a number of financial and economic time series. However, it is well known that structural changes in the mean of time series can easily be confused with long memory dependence. Shifts in mean can heavily bias estimates for long memory index and therefore create misleading results. See Banerjee and Urga (2005) and Mayoral (2011) for excellent reviews of fractionally integrated and structural breaks from an econometric standpoint. Therefore, the purpose of this paper is to investigate the possible existence of spurious relationship in a pair of stationary fractionally integrated processes with breaks.

After this introduction, Section 2 is dedicated to the requisite test procedures we use and develop the asymptotic results for fractionally integrated series involving structural breaks. In Section 3, we undertake Monte Carlo simulation to assess the sample performance of our asymptotic results. Conclusions are drawn in the final section.

1. MAIN RESULTS

In order to investigate the spurious regression effects caused by structural breaks in level, we would specify the data generating processes (DGP) of two independent variables, x_t and y_t, with sample size T as:

$$x_t = \mu + \delta \cdot 1_{\{t>[T\tau]\}} + \varepsilon_t, \tag{1}$$

$$y_t = \phi + \varphi \cdot 1_{\{t>[T\lambda]\}} + \xi_t, \tag{2}$$

where x_t and y_t are stationary with structural breaks in level respectively; $\mu(\phi)$ and $\delta(\varphi)$ are respective the permanent means and the transitory intercepts, resulting from a change, of the process $x_t(y_t)$; τ and λ are the 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 fractionally integrated processes with differencing parameters d_ε and d_ξ: $(1-L)^{d_\varepsilon}\varepsilon_t = u_t$ and $(1-L)^{d_\xi}\xi_t = v_t$, where both u_t and v_t are independent standard normal variables.

We focus on the 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 "$\xrightarrow{w}$" stands for the weak convergence and $M_T \sim N_T$ mean $M_T/N_T \to 1$ as $T \to \infty$. The random variable $B_d(r)$ denotes the fractional Brownian motion with

$$B_d(r) \equiv \frac{1}{\Gamma(1+d)}\left\{\int_0^r (r-s)^d\, dW(s) + \int_{-\infty}^0 [(r-s)^d - (-s)^d]\, dW(s)\right\},$$

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

Assumption 2.1 *The stationary series* ε_t and ξ_t *are assumed to mean zero with fractionally integrated indexes* $d_\varepsilon, d_\xi \in (0, 0.5)$.

Our method relies on the specific regularity conditions ensuring following lemma, which has been considered by Davidson and Jong (2000).

Lemma 2.1 *If Assumption 2.1 holds, then as* $T \to \infty$,

$$\left(\lambda_\varepsilon^{-1}\sum_{t=1}^{[Tr]}\varepsilon_t,\ T^{-1}\sum_{t=1}^{T}\varepsilon_t^2\right) \xrightarrow{w} \left(B_{d_\varepsilon}(r),\ \sigma_\varepsilon^2\right),$$

$$\left(\lambda_\xi^{-1}\sum_{t=1}^{[Tr]}\xi_t,\ T^{-1}\sum_{t=1}^{T}\xi_t^2\right) \xrightarrow{w} \left(B_{d_\xi}(r),\ \sigma_\xi^2\right),$$

where

$$\lambda_\varepsilon^2 \sim T^{1+2d_\varepsilon}\sigma_\varepsilon^2, \qquad \sigma_\varepsilon^2 = \frac{\Gamma(1-2d_\varepsilon)}{\Gamma(1-d_\varepsilon)^2}\sigma_u^2,$$

$$\lambda_\xi^2 \sim T^{1+2d_\xi}\sigma_\xi^2, \qquad \sigma_\xi^2 = \frac{\Gamma(1-2d_\xi)}{\Gamma(1-d_\xi)^2}\sigma_v^2.$$

Lemma 2.1 indicates these results that $\sum_{t=1}^{[Tr]}\varepsilon_t = O_p(T^{1/2+d_\varepsilon})$ and $\sum_{t=1}^{[Tr]}\xi_t = O_p(T^{1/2+d_\xi})$ Furthermore, the fractional invariance principle could be extensively extended to cover all possible fractionally integrated indexes $d>-0.5$, following Marinucci and Robinson (2000).

The simple regression model is given by

$$y_t = \alpha + \beta x_t + \eta_t, \tag{3}$$

where y_t and x_t are the regressand and regressor, and η_t is the regression error. The general null hypothesis on β is routinely formulated as $H_0: \beta = 0$, and tested against the alternative hypothesis $H_1: \beta \neq 0$. We need the following lemma to complete our theorem.

Lemma 2.2 *If the conditions of Lemma 2.1 are satisfied, then as* $T \to \infty$,

1)
$$\frac{1}{T}\sum_{t=1}^{T} x_t \xrightarrow{w} \mu + (1-\tau)\delta \equiv \Pi_x, \qquad \frac{1}{T}\sum_{t=1}^{T} x_t^2 \xrightarrow{w} \mu^2 + (1-\tau)(2\mu\delta + \delta^2) + \sigma_\varepsilon^2 \equiv \Pi_{xx},$$

$$\frac{1}{T}\sum_{t=1}^{T} y_t \xrightarrow{w} \phi + (1-\tau)\varphi \equiv \Pi_y, \qquad \frac{1}{T}\sum_{t=1}^{T} y_t^2 \xrightarrow{w} \phi^2 + (1-\lambda)(2\phi\varphi + \varphi^2) + \sigma_\xi^2 \equiv \Pi_{yy},$$

2)
$$\frac{1}{T}\sum_{t=1}^{T} x_t y_t \xrightarrow{w} \mu\phi + \mu\varphi(1-\lambda) + \delta\phi(1-\tau) + \delta\varphi(1-\max(\tau,\lambda)) \equiv \Pi_{xy}.$$

On routine applications of standard t-ratio test in regression models, we demonstrate that spurious regression can happen due to neglected of change points in level. The theorem collects the asymptotic behavior of the estimated parameters, $\hat{\beta}$ and associated $t_{\hat{\beta}}$ in model (3).

Theorem 2.1 *Suppose* x_t *and* y_t *are respectively generated by* (1)-(2) *with break fraction* τ, $\lambda \in (0, 1)$. *If the conditions of Lemma 2.1 are satisfied, then as* $T \to \infty$,

a) $$\hat{\beta} \xrightarrow{w} \frac{\Pi_{xy} - \Pi_x\Pi_y}{\Pi_{xx} - \Pi_x^2} \equiv \beta,$$

b) $$T^{-1/2}t_{\hat{\beta}} \xrightarrow{w} \beta \cdot \left(\frac{\Pi_{yy} - \Pi_y^2 + \beta^2(\Pi_{xx} - \Pi_x^2)}{\Pi_{xx} - \Pi_x^2}\right)^{-1/2}$$

Theorem 3.1 presents that the asymptotic distribution of $\hat{\beta}$ is not standard normal and more complicated, not yet to its true value zero. It is interesting that $\hat{\beta}$ diverges at the rate of $T^{1/2}$, wrongly indicating spurious relationship. The divergence of $t_{\hat{\beta}}$ implies that the rejection probability for testing null of no relationship increases as sample size approaches infinity. It seems the rate of $t_{\hat{\beta}}$ is always divergent regardless of the fractionally ingerated indexes, but the simulation in Section 3 shows that the rejection probability for testing null of no relationship is serious sensitive to them. Hence, it visually shows that the presence of spurious regression is unambiguous if the observations are fractionally integrated with structural breaks.

3. SIMULATION

The aim of the Monte Carlo simulation is to examine the sample performance of our theoretical results in Sections 2. We calculate and construct the t-ratios in models (3) for the samples of size T=500,1000,10000. For each simulated sample, the percentage of rejection are obtained by 3000 replications at 5% nominal level, i.e., the percentage of t-ratios such that $|t_{\hat{\beta}}|$>1.96. In order to weaken the initialization effect, we would generate these series of length T+200 and trim the first 200 observations to get the simulated series. To save space, we just report the results for $\mu=\phi=0$, and break transition $\delta=\varphi=0.5$, the other cases have similar results. These series ε_t and ξ_t satisfy Assumption 2.1 with fractionally integrated indexes varying among {0.2, 0.4}. The break fraction τ and λ belong to {0.2, 0.5, 0.8}.

Table 1-3 provide the percentages of rejection of $t_{\hat{\beta}}$ convergence for each structural breaks, fractionally integrated indexes, and sample size profile. As expected, the rejection percentage of $t_{\hat{\beta}}$ increases as sample size approaching to infinity. For example, when $d_\varepsilon=d_\xi=0.2$ and $\tau=\lambda=0.8$, the rejection percentage are 21.80%, 73.52% and 96.25% for T=500,1000,10000. Another impressive finding is that, as the break fractions approach to each other, the rejection rate increases; the further away the breaks are from each other, the more observations are required to generate the spurious correlation. If $d_\varepsilon=d_\xi=0.4$, T=10000 and $\tau=0.5$, the rejection rate are 18.68%, 81.17% and 18.76% for λ=0.2, 0.5, 0.8. All confirm that the spurious relationship is sensitive to the position of structural breaks.

What is surprising is that, the rate of divergence of $t_{\hat{\beta}}$ varies according to the degree of fractionally integrated indexes d_ε and d_ξ, which is an interesting result for our proposed test procedures. When $\tau=0.5$, $\lambda=0.8$, T=1000, the rejection rate are 36.22% and 22.26% for d_ε=0.2, 0.4. The smaller fractionally integrated index d_ε does induce the larger rejection percentages. The same results hold true for another fractionally integrated index d_ξ. The spurious regression could occur and the extent of that phenomenon depends on the closeness of the fractionally integrated index d_ξ to its lower bound 0.2. If $\tau=\lambda=0.5$, d_ε=0.4 and T=500, the rejection rate are 45.52% and 22.24 for d_ξ=0.2, 0.4.

Table 1
Rejection Percentage of the $|t_{\hat{\beta}}|$>1.96, T=500

		τ=0.2 d_ε		τ=0.2 d_ε		τ=0.2 d_ε	
λ	d_ξ	0.2	0.4	0.2	0.4	0.2	0.4
0.2	0.2	21.84	14.47	12.04	8.72	6.49	6.07
	0.4	10.48	6.48	6.68	5.26	5.12	4.63
0.5	0.2	12.20	9.75	45.52	25.68	10.93	9.04
	0.4	6.36	4.92	22.24	11.12	6.44	5.63
0.8	0.2	6.04	5.92	11.32	8.64	21.80	16.12
	0.4	5.06	4.89	5.92	6.14	11.76	6.80

Table 2
Rejection Percentage of the $|t_{\hat{\beta}}|$>1.96, T=1000

		τ=0.2 d_ε		τ=0.2 d_ε		τ=0.2 d_ε	
λ	d_ξ	0.2	0.4	0.2	0.4	0.2	0.4
0.2	0.2	73.61	48.03	35.36	22.56	9.36	6.61
	0.4	43.60	24.37	17.24	11.60	8.72	5.47
0.5	0.2	37.31	22.24	97.88	83.36	36.36	22.29
	0.4	17.92	11.56	79.04	51.60	20.08	12.12
0.8	0.2	10.24	7.28	36.22	22.26	73.52	48.57
	0.4	7.32	5.12	19.84	11.12	42.42	23.96

Table 3
Rejection Percentage of the $|t_{\hat{\beta}}|$>1.96, T=10000

		τ=0.2 d_ε		τ=0.2 d_ε		τ=0.2 d_ε	
λ	d_ξ	0.2	0.4	0.2	0.4	0.2	0.4
0.2	0.2	95.15	77.15	61.92	38.22	15.64	10.22
	0.4	71.72	46.43	35.04	18.68	9.16	8.29
0.5	0.2	62.16	39.53	100.0	98.81	63.12	38.41
	0.4	34.87	20.84	97.43	81.17	33.24	19.56
0.8	0.2	15.09	10.64	63.88	38.08	96.25	78.15
	0.4	8.82	6.56	35.28	18.76	73.52	47.81

In order to give intuitive idea for the influence of fractionally integrated sequences, we provide the rejection frequency with sample size T=1000 in Figure 1. As expected, all these figures clearly show that spurious relationship is present when the fractionally integrated series involves change points. This confirms the conclusion that the smaller fractionally integrated indexes provides higher rejection rate. Moreover, a surprising finding which could not provided by Theorem 3.1 is that there is one peak. More special, the highest rejection frequency occurs near $\tau=\lambda=0.5$. Hence, it visually shows that the presence of spurious regression is unambiguous if the observations are fractionally integrated with structural breaks.

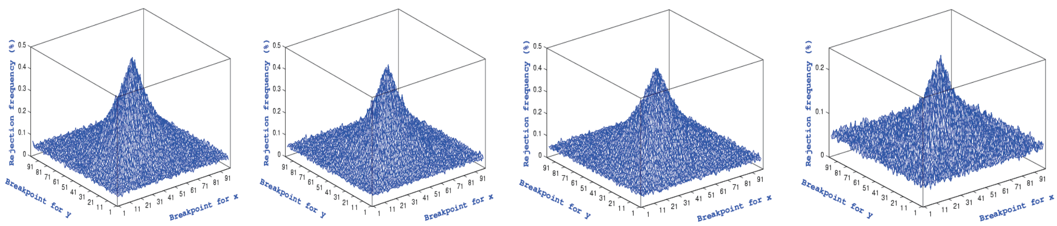

Figure 1
Note. The rejection frequency when T=1000, d_ξ=0.4 and d_ε=0.1, 0.2, 0.3, 0.4.

CONCLUSION

In this paper, we consider spurious relationship between fractionally integrated sequences driven by structural changes in regressions model. The t-ratios are always divergent with rate of $\sqrt{T}$ and wrongly indicate the fractionally integrated sequences with structural breaks have a significant linear relationship. Furthermore, the simulation results reveal the phenomenon of spurious regression significantly depends on the fractionally integrated indexes, the location of breaks and sample size. Given that many important economic and financial time series show some strong evidence for fractionally integrated volatilities with changes, the potential to encounter such spurious relationship in practical applications seems to be very high.

REFERENCES

Banerjee, A., & Urga, G. (2005). Modelling structural breaks, long memory and stock market volatility: An overview. *Journal of Econometrics*, *129*(5), 1-34.

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*(4), 643-666.

Entorf, H. (1997). Random walks with drifts: Nonsense regression and spurious fixed-effect estimation. *Journal of Econometrics*, *80*(1), 287-296.

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

Granger, C. W. J., Hyung, N., & Jeon, Y. (2001). Spurious regressions with stationary series. *Applied Economics*, *33*(4), 899-904.

Marinucci, D., & Robinson, P. M. (1999). Alternative forms of fractional Brownian motion. *Journal of Statistical Planning and Inference*, *80*(3), 111-122.

Marmol, F. (1996). Correlation theory for spuriously related higher order integrated processes. *Journal of Time Series Analysis*, *16*(6), 313-321.

Marmol, F. (1998). Spurious regression theory with nonstationary fractionally integrated processes. *Journal of Econometrics*, *84*(5), 233-250.

Marmol, F., (1995). Spurious regressions for I(d) processes. *Journal of Time Series Analysis*, *16*(12), 313-321.

Martinez-Rivera, B., & Ventosa-Santaularia, D. (2012). A comment on "Is the spurious regression problem spurious?" *Economics Letters*, *115*(2), 229-231.

Mayoral, L. (2011). Testing for fractional integration versus short memory with trends and structural breaks. *Oxford Bulletin of Economics and Statistics*, *74*(3), 278-305.

McCallum, B. (2010). Is the spurious problem spurious? *Economics Letters*, *107*(7), 321-323.

Nelsonl, C., & Plosser, C. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, *8*(13), 139-162.

Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, *33*(8), 311-340.

Tsay, W. -J., & Chung, C. -F. (2000). The spurious regression of fractionally integrated processes. *Journal of Economics*, *96*(5), 155-182.

APPENDIX

Proof of Lemma 2.2 By Lemma 2.1, we obtain

$$\sum_{t=1}^{T} x_t = T\mu + \sum_{t=1}^{T} \delta \cdot 1_{\{t>[T\tau]\}} + \sum_{t=1}^{T} \varepsilon_t,$$

where $\sum_{t=1}^{T} \delta \cdot 1_{\{t>[T\tau]\}} = T\delta(1-\tau)$

and $\sum_{t=1}^{T} \varepsilon_t = O_p(T^{1/2+d_\varepsilon})$, $d_\varepsilon \in (0, 0.5)$,

then $\dfrac{1}{T}\sum_{t=1}^{T} x_t \xrightarrow{w} \mu + \delta(1-\tau) \equiv \Pi_x$.

Because of $T^{-1}\sum_{t=1}^{T} x_t^2 \xrightarrow{w} \sigma_\varepsilon^2$, we can get

$$\frac{1}{T}\sum_{t=1}^{T} x_t^2 = \frac{1}{T}\sum_{t=1}^{T}\left\{(\mu + \delta \cdot 1_{\{t>[T\tau]\}})^2 + 2(\mu + \delta \cdot 1_{\{t>[T\tau]\}})\varepsilon_t + \varepsilon_t^2\right\}$$

$$\xrightarrow{w} \mu^2 + (1-\tau)(2\mu\delta + \delta^2) \equiv \Pi_{xx}.$$

The proof of item 1 in Lemma 2.2 is established. Thus we can get the same argument for item 2 in Lemma 2.2

Some algebra yields

$$x_t y_t = (\mu + \delta \cdot 1_{\{t>[T\tau]\}})(\phi + \varphi \cdot 1_{\{t>[T\tau]\}}) +$$
$$(\mu + \delta \cdot 1_{\{t>[T\tau]\}})\xi_t + (\phi + \varphi \cdot 1_{\{t>[T\tau]\}})\varepsilon_t + \varepsilon_t \xi_t.$$

Since ε_t and ξ_t are independent, applying the results of Tsay and Chung (2000), thus we have that $T^{-1}\sum_{t=1}^{T}\varepsilon_t\xi_t = o_p(1)$. Hence, we could together results with $\sum_{t=1}^{T}\varepsilon_t = O_p(T^{1/2+d_\varepsilon})$ and $\sum \quad _t \quad O_p(T^{1/2} \quad)$, then

$$\frac{1}{T}\sum_{t=1}^{T}x_t y_t \xrightarrow{w} \mu\phi + \mu\varphi(1-\lambda) + \delta\varphi(1-\tau) + \delta\varphi(1-\max(\tau,\lambda)).$$

Hence the item 3 is proved, and we complete the proof of Lemma 2.2.

Proof of Theorem 3.1 The OLS statistics are given by

$$\hat{} \quad \frac{\sum_{t=1}^{T}x_t y_t \quad T \quad \sum_{t=1}^{T}x_t \sum_{t=1}^{T}y_t}{\sum \quad (x \quad \bar{x})}$$

$$s^2 = \frac{1}{T}\sum_{t=1}^{T}\hat{\eta}_t^2 = \frac{1}{T}\sum_{t=1}^{T}(y_t-\bar{y})^2 - \hat{\beta}^2\frac{1}{T}\sum_{t=1}^{T}(x_t-\bar{x})^2,$$

$$s_{\hat{\beta}}^2 = \frac{s^2}{\sum_{t=1}^{T}(x_t-\bar{x})^2}, \qquad t_{\hat{\beta}} = \frac{\hat{\beta}}{s_{\hat{\beta}}}.$$

It is note that

$$\hat{\beta} = \frac{T^{-1}\sum_{t=1}^{T}x_t y_t - \left(T^{-1}\sum_{t=1}^{T}x_t\right)\left(T^{-1}\sum_{t=1}^{T}y_t\right)}{T^{-1}\sum_{t=1}^{T}x_t^2 - \left(T^{-1}\sum_{t=1}^{T}x_t\right)\left(T^{-1}\sum_{t=1}^{T}x_t\right)} \xrightarrow{w} \frac{\Pi_{xy} - \Pi_x\Pi_y}{\Pi_{xx}},$$

where the convergence is based on Lemma 2.2.

Applying the Lemma 2.2 and the asymptotic behavior of $\hat{\beta}$, we have

$$s^2 = \frac{1}{T}\sum_{t=1}^{T}y_t^2 - \left(\frac{1}{T}\sum_{t=1}^{T}y_t\right)^2 - \hat{\beta}^2\left(\frac{1}{T}\sum_{t=1}^{T}x_t^2 - \left(\frac{1}{T}\sum_{t=1}^{T}x_t\right)^2\right) \xrightarrow{w} \Pi_{yy} - \Pi_y^2 + \beta^2(\Pi_{xx} - \Pi_x^2),$$

and

$$Ts_{\hat{\beta}}^2 = \frac{s^2}{T^{-1}\sum_{t=1}^{T}x_t^2 - \left(T^{-1}\sum_{t=1}^{T}x_t\right)^2} \xrightarrow{w} \frac{\Pi_{yy} - \Pi_y^2 + \beta^2(\Pi_{xx} - \Pi_x^2)}{\Pi_{xx} - \Pi_x^2}.$$

then,

$$T^{-1/2}t_{\hat{\beta}} = \frac{\hat{\beta}}{\left(Ts_{\hat{\beta}}^2\right)^{1/2}} \xrightarrow{w} \beta\cdot\left(\frac{\Pi_{yy} - \Pi_y^2 + \beta^2(\Pi_{xx} - \Pi_x^2)}{\Pi_{xx} - \Pi_x^2}\right)^{-1/2}.$$

Hence, the proof of Theorem 2.1 is completed.

