*Econometric Theory*, **15**, 1999, 622–628. Printed in the United States of America.

# SPURIOUS REGRESSION BETWEEN $I(1)$ PROCESSES WITH INFINITE VARIANCE ERRORS

WEN-JEN TSAY
*National Chi Nan University*

This paper considers spurious regression between integrated processes with stable errors. Our results show that the *t*-ratios diverge at the rate of $\sqrt{T}$, which is identical to what Phillips (1986, *Journal of Econometrics* 33, 311–340) has obtained for the Gaussian case. Therefore, it is the long memory in the dependent variable and regressors, instead of the moment conditions of the error terms, that causes the spurious regression.

## 1. INTRODUCTION

Spurious regression was first studied by Granger and Newbold (1974). They showed that the regression of an integrated process of order one, or $I(1)$, on another independent $I(1)$ process will result in spurious effects, i.e., the null hypothesis of no relationship among the unrelated $I(1)$ processes will be rejected too often. Granger and Newbold's simulation results were later proved by Phillips (1986). Phillips showed that the usual *t*-statistic in a spurious regression does not have a limiting distribution but actually diverges as the sample size $(T)$ approaches infinity. Marmol (1995) generalized the work of Phillips to $I(d)$ processes with $d$ being integer. His results reveal that the *t*-ratios also diverge at the $\sqrt{T}$ rate independent of the integer $d$. The case of long memory processes was considered by Tsay and Chung (1995). They found that spurious regression can arise among a wide range of long memory processes, even in cases where both the dependent variable and regressor are stationary.

There has been fast growing interest in modeling time series by ARIMA models with heavy-tailed errors, i.e., the error process belongs to the domain of attraction of a stable law with index $\alpha (0 < \alpha < 2)$. This process was first used by Mandelbrot (1963) to model the heavy-tailed nature of stock returns. Examples where such models are appropriate include Fama (1965), Fama and Roll (1968), So (1987), and Chan (1990). Therefore, we wonder whether the presence of stable errors affects the divergence rate of the *t*-statistics in the regression among unrelated $I(1)$ processes. The nonstandard asymptotic distributions of various

I am grateful for the helpful comments by two anonymous referees. The comments of the Co-Editor B.E. Hansen greatly improved the exposition of the paper. Address correspondence to: Wen-Jen Tsay, Department of Economics, National Chi Nan University, 1, University Road, Puli, Nantou Hsien, Taiwan, R.O.C.; e-mail: wjtsay@ncnu.edu.tw.

 © 1999 Cambridge University Press 0266-4666/99 $9.50

https://doi.org/10.1017/S0266466699154069 Published online by Cambridge University Press

coefficient estimators and test statistics are derived to answer the preceding question. Instead of Brownian motions, weak convergence involving Lévy processes is needed to handle the infinite variance errors. Our results show that the *t*-ratios diverge at the rate of $\sqrt{T}$, which is independent of the index parameters of the error terms. Moreover, the divergence rate of the *t*-ratios is exactly the same as what Phillips (1986) has obtained for the Gaussian case where $\alpha = 2$. Consequently, spurious regression is a phenomenon unaffected by whether or not the errors are heavy-tailed.

## 2. A GENERAL THEORY OF SPURIOUS REGRESSION

Our analysis is based on a simple linear regression model in which the dependent variable and the regressor are independent $I(1)$ processes with different heavy-tailed errors. Before presenting these models, let us first briefly review some basic properties of the stable processes.

A stable distribution is characterized by four parameters: $\alpha \in (0,2]$ representing the index of stability, $\sigma \in (0,\infty)$ representing the scale, $\beta \in [-1,1]$ representing the skewness, and $\delta \in (-\infty,\infty)$ representing the location. The most important parameter in determining the tail shape is $\alpha$, which is the maximal moment exponent of the distribution. Absolute moments of the stable process of order less than $\alpha$ are finite, whereas all higher-order moments are infinite. Therefore, stable processes have infinite variance except in the special case when $\alpha = 2$ (Gaussian). Moreover, the stable process is closed under temporal aggregation. When $\alpha < 2$, the tail of the distribution decays like a power function.

Throughout this paper, we assume the following conditions for the stable process.

Assumption 1. The process $\epsilon_t$ is an independently and identically distributed (i.i.d.) sequence with distribution $F$. The distribution $F$ is in the domain of attraction of a stable law with index $\alpha \in (0,2]$. If $\alpha > 1$ then we require $E(\epsilon_t) = 0$. If $\alpha = 1$ then we require $\epsilon_t$ to be symmetric at zero.

Following the discussion in Chan and Tran (1989), we define $a_\epsilon(T) = T^{1/\alpha} l(T)$, where $l(T)$ is a slowly varying function that satisfies

$$\lim_{t\to\infty} \frac{l(tx)}{l(t)} = 1 \quad \text{for all } x > 0.$$

As $\epsilon_t$ satisfies the conditions in Assumption 1, Resnick (1986) shows that as $T \to \infty$,

$$\left( a_\epsilon(T)^{-1} \sum_{t=1}^{[\mathrm{Tr}]} \epsilon_t, a_\epsilon(T)^{-2} \sum_{t=1}^{[\mathrm{Tr}]} \epsilon_t^2 \right) \Rightarrow (U_\alpha(r), V_\alpha(r))$$

in $D[0,1]^2$, where $r \in [0,1]$, $[\mathrm{Tr}]$ denotes the integer part of Tr, $\Rightarrow$ denotes weak convergence, and $(U_\alpha, V_\alpha)$ is a Lévy process with index $\alpha$. When $\alpha = 2$, which

https://doi.org/10.1017/S0266466699154069 Published online by Cambridge University Press

corresponds to the finite variance case, we have $U_\alpha(r) = B(r)$, a standard Brownian motion, and $V_\alpha(r) = r$.

To examine the spurious effects, we first define two independent stable processes, $v_t$, $w_t$, which both satisfy the conditions in Assumption 1 with indices $\alpha_1$ and $\alpha_2$, respectively. We can then define two $I(1)$ processes by integrating $v_t$ and $w_t$:

$$y_t = y_{t-1} + v_t \quad \text{and} \quad x_t = x_{t-1} + w_t.$$

We also assume, without loss of generality, that the initial values of the stable processes $v_0$, $w_0$, $y_0$, and $x_0$ are all zero. Hence, $y_t$ and $x_t$ can be viewed as the partial sums of $v_t$ and $w_t$, respectively; i.e., $y_T = \sum_{t=1}^{T} v_t$ and $x_T = \sum_{t=1}^{T} w_t$. Thus, if we define $a_y(T) = T^{1/\alpha_1} l_y(T), a_x(T) = T^{1/\alpha_2} l_x(T)$, then we have

$$a_y(T)^{-1} \sum_{t=1}^{[\mathrm{Tr}]} v_t \Rightarrow U_{\alpha_1}(r) \quad \text{and} \quad a_x(T)^{-1} \sum_{t=1}^{[\mathrm{Tr}]} w_t \Rightarrow U_{\alpha_2}(r).$$

Given the preceding discussion, we consider the following simple linear regression model:

$$y_t = \beta_0 + \beta_1 x_t + u_t.$$

Let $\hat{\beta}_0$ and $\hat{\beta}_1$ denote the usual ordinary least squares (OLS) estimators from a regression of $y_t$ on a constant and $x_t$, respectively. Their respective variances are estimated by $s^2_{\hat{\beta}_0}$ and $s^2_{\hat{\beta}_1}$, from which we have the $t$-ratios $t_{\beta_0} = \hat{\beta}_0 / s_{\beta_0}$ and $t_{\beta_1} = \hat{\beta}_1 / s_{\beta_1}$.

THEOREM 1. *Given that Assumption* 1 *holds, then, as* $T \to \infty$, *we have the following results*:

1. $$\frac{a_x(T)}{a_y(T)} \hat{\beta}_1 \Rightarrow \frac{\int_0^1 U_{\alpha_1}(s) \cdot U_{\alpha_2}(s)\, ds - \left[\int_0^1 U_{\alpha_1}(s)\, ds\right]\left[\int_0^1 U_{\alpha_2}(s)\, ds\right]}{\int_0^1 [U_{\alpha_2}(s)]^2\, ds - \left[\int_0^1 U_{\alpha_2}(s)\, ds\right]^2} \equiv \beta_1^*.$$

   *Note that* $a_y(T)/a_x(T) = O(T^{1/\alpha_1 - 1/\alpha_2})$.

2. $$\frac{1}{a_y(T)} \hat{\beta}_0 \Rightarrow \int_0^1 U_{\alpha_1}(s)\, ds - \beta_1^* \int_0^1 U_{\alpha_2}(s)\, ds \equiv \beta_0^*,$$

   *where* $\beta_1^*$ *is defined in* 1. *Note that* $a_y(T) = O(T^{1/\alpha_1})$.

3. $$\frac{1}{\sqrt{T}} t_{\beta_1} \Rightarrow \frac{\beta_1^*}{\sigma_{*\beta_1}},$$

   *where* $\beta_1^*$ *is defined in* 1 *and* $\sigma^2_{*\beta_1}$ *is defined in the Appendix*.

https://doi.org/10.1017/S0266466699154069 Published online by Cambridge University Press

4. $\dfrac{1}{\sqrt{T}}\, t_{\beta_0} \Rightarrow \dfrac{\beta_0^*}{\sigma_{*\beta_0}}$,

*where* $\beta_0^*$ *is defined in* 2 *and* $\sigma^2_{*\beta_0}$ *is defined in the Appendix*.

*Here,* $U_{\alpha_1}(s)$ *and* $U_{\alpha_2}(s)$ *are two independent Lévy processes*.

The proof of Theorem 1 is in the Appendix. Theorem 1 shows that, as the sample size $T$ increases, the two $t$-ratios $t_{\beta_1}$ and $t_{\beta_0}$ diverge at the $\sqrt{T}$ rate, which is independent of the index parameters, $\alpha_1$ and $\alpha_2$. The divergence rate of these two $t$-ratios is exactly the same as what Phillips (1986) has obtained for the case where $\alpha_1 = \alpha_2 = 2$. Thus, even when $v_t$ and $w_t$ are stable processes, the spurious regression occurs, i.e., the probability of rejecting the null hypothesis of $\beta = 0$ or $\alpha = 0$ based on $t$-tests increases monotonically as the sample size approaches infinity.

The limiting distributions of the $t$-ratios, after being normalized by $\sqrt{T}$, are direct generalization of those derived by Phillips. In other words, when we compare our results with those of Phillips, we observe a common feature in these two statistics, i.e., the index parameters, $\alpha_1$ and $\alpha_2$, do not affect their convergence rates whereas the effects on their limiting distributions are quite straightforward: all the standard Brownian motions in Phillips's theory are replaced by Lévy processes. That the stability parameters $\alpha_1$ and $\alpha_2$ play a relatively minor role here is mainly because the two statistics are all ratios so that the effects of $\alpha_1$ and $\alpha_2$ are canceled out. In contrast, the results on the OLS estimators $\hat{\beta}_1$ and $\hat{\beta}_0$ are a different story. In Phillips's theory both $\hat{\beta}_1$ and $\hat{\beta}_0/\sqrt{T}$ converge to some nonnormal degenerate limiting distributions. But for the present model concerning the stable processes, the orders of $\hat{\beta}_1$ and $\hat{\beta}_0$ are $T^{1/\alpha_1 - 1/\alpha_2}$ and $T^{1/\alpha_1}$, respectively. Therefore, the divergence rate of $\hat{\beta}_0$ decreases with $\alpha_1$, which is intuitively reasonable because the driving force of the dependent variable becomes less volatile.

## 3. CONCLUSION

In our analysis of spurious regression between integrated processes with stable errors, we find that the $t$-ratios diverge at the rate of $\sqrt{T}$, which is identical to what Phillips (1986) has obtained for the Gaussian case. So it is the long memory in the dependent variable and regressor, instead of the moment conditions of the error terms, that causes the spurious regression.

A few generalizations of our work are worthy of further consideration. As suggested by Chan (1990), the errors $v_t$ and $w_t$ can be generated by the more general ARMA process driven by i.i.d. stable process, respectively. Based on the work of Phillips (1990), we expect most of the results in this paper to hold when $v_t$ and $w_t$ are both ARMA processes with stable errors. Another extension of this paper is to generalize the investigation of Tsay and Chung (1995) that $v_t$ and $w_t$ are generated by the fractionally integrated processes with stable errors. This process is first introduced in Kokoszka and Taqqu (1996), which will serve as the building block for future studies.

https://doi.org/10.1017/S0266466699154069 Published online by Cambridge University Press

*REFERENCES*

Chan, N.H. (1990) Inference for near-integrated time series with infinite variance. *Journal of the American Statistical Association* 85, 1069–1074.

Chan, N.H. & L.T. Tran (1989) On the first order autoregressive process with infinite variance. *Econometric Theory* 5, 354–362.

Fama, E. (1965) The behavior of stock market prices. *Journal of Business* 38, 34–105.

Fama, E. & R. Roll (1968) Some properties of symmetric stable distributions. *Journal of the American Statistical Association* 63, 817–836.

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

Kokoszka, P.S. & M.S. Taqqu (1996) Infinite variance stable moving averages with long memory. *Journal of Econometrics* 73, 79–99.

Mandelbrot, B.B. (1963) The variation of certain speculative prices. *Journal of Business* 36, 394–419.

Marmol, F. (1995) Spurious regression between $I(d)$ processes. *Journal of Time Series Analysis* 16, 313–321.

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

Phillips, P.C.B. (1990) Time series regression with a unit root and infinite variance errors. *Econometric Theory* 6, 44–62.

Resnick, S.I. (1986) Point processes, regular variation and weak convergence. *Advances in Applied Probability* 18, 66–138.

So, J. (1987) The sub-Gaussian distribution of currency futures: Stable Paretian or nonstationary? *Review of Economics and Statistics* 69, 100–107.

Tsay, W.J. & C.-F. Chung (1995) The spurious regression of fractionally integrated processes. *Journal of Econometrics*, forthcoming.

# APPENDIX

To prove Theorem 1, first we note the following result.

THEOREM A.1. *Given that Assumption* 1 *holds, then, as* $T \to \infty$, *the following joint weak convergences hold*:

1. $\dfrac{1}{T}\sum_{t=1}^{T}\dfrac{y_t}{a_y(T)} \Rightarrow \int_0^1 U_{\alpha_1}(s)\,ds \quad \text{and} \quad \dfrac{1}{T}\sum_{t=1}^{T}\dfrac{x_t}{a_x(T)} \Rightarrow \int_0^1 U_{\alpha_2}(s)\,ds.$

2. $\dfrac{1}{T}\sum_{t=1}^{T}\dfrac{y_t^2}{a_y(T)^2} \Rightarrow \int_0^1 [U_{\alpha_1}(s)]^2\,ds \quad \text{and} \quad \dfrac{1}{T}\sum_{t=1}^{T}\dfrac{x_t^2}{a_x(T)^2} \Rightarrow \int_0^1 [U_{\alpha_2}(s)]^2\,ds.$

3. $\dfrac{1}{T}\sum_{t=1}^{T}\dfrac{y_t}{a_y(T)}\cdot\dfrac{x_t}{a_x(T)} \Rightarrow \int_0^1 U_{\alpha_1}(s)\cdot U_{\alpha_2}(s)\,ds.$

*Here*, $U_{\alpha_1}(s)$ *and* $U_{\alpha_2}(s)$ *are two independent Lévy processes.*

https://doi.org/10.1017/S0266466699154069 Published online by Cambridge University Press

The preceding results are straightforward applications of the results in Lemma 1 of Chan (1990) and the continuous mapping theorem.

To prove item 1 of Theorem 1, we note

$$\hat{\beta}_1 = \frac{\sum_{t=1}^{T} y_t x_t - \frac{1}{T}\sum_{t=1}^{T} y_t \sum_{t=1}^{T} x_t}{\sum_{t=1}^{T} (x_t - \bar{x})^2}.$$

Therefore, we have

$$\frac{a_x(T)}{a_y(T)}\hat{\beta}_1 = \frac{\frac{1}{Ta_y(T)a_x(T)}\sum_{t=1}^{T} y_t x_t - \left(\frac{1}{Ta_y(T)}\sum_{t=1}^{T} y_t\right)\left(\frac{1}{Ta_x(T)}\sum_{t=1}^{T} x_t\right)}{\frac{1}{Ta_x(T)^2}\sum_{t=1}^{T} (x_t - \bar{x})^2} \Rightarrow \beta_1^*,$$

where the weak convergence is based on items 1–3 of Theorem A.1.

To prove item 2 of Theorem 1, we note

$$\hat{\beta}_0 = \frac{1}{T}\sum_{t=1}^{T} y_t - \hat{\beta}_1 \cdot \frac{1}{T}\sum_{t=1}^{T} x_t,$$

so we have

$$\frac{1}{a_y(T)}\hat{\beta}_0 = \frac{1}{Ta_y(T)}\sum_{t=1}^{T} y_t - \frac{a_x(T)}{a_y(T)}\hat{\beta}_1 \cdot \frac{1}{Ta_x(T)}\sum_{t=1}^{T} x_t \Rightarrow \beta_0^*,$$

where the weak convergence is based on item 1 of Theorem 1 and item 1 of Theorem A.1.

To prove item 3 of Theorem 1, first we note the estimated variance of the OLS residuals. The term $s^2$ is defined as

$$s^2 = \frac{1}{T}\sum_{t=1}^{T} (y_t - \hat{\beta}_0 - \hat{\beta}_1 x_t)^2 = \frac{1}{T}\sum_{t=1}^{T} (y_t - \bar{y})^2 - \hat{\beta}_1^2 \cdot \frac{1}{T}\sum_{t=1}^{T} (x_t - \bar{x})^2.$$

So we have

$$\begin{aligned}\frac{1}{a_y(T)^2}s^2 &= \frac{1}{Ta_y(T)^2}\sum_{t=1}^{T} (y_t - \bar{y})^2 - \left(\frac{a_x(T)}{a_y(T)}\hat{\beta}_1\right)^2 \frac{1}{Ta_x(T)^2}\sum_{t=1}^{T} (x_t - \bar{x})^2 \\ &\Rightarrow \int_0^1 [U_{\alpha_1}(s)]^2\, ds - \left[\int_0^1 U_{\alpha_1}(s)\, ds\right]^2 \\ &\quad - (\beta_1^*)^2\left\{\int_0^1 [U_{\alpha_2}(s)]^2\, ds - \left[\int_0^1 U_{\alpha_2}(s)\, ds\right]^2\right\} \\ &\equiv \sigma_*^2,\end{aligned}$$

https://doi.org/10.1017/S0266466699154069 Published online by Cambridge University Press

where the weak convergence is based on item 1 of Theorem 1 and items 1 and 2 of Theorem A.1. Furthermore, we note that the variance of $\hat{\beta}_1, s_{\beta_1}^2$ is defined as

$$s_{\beta_1}^2 = \frac{s^2}{\sum_{t=1}^{T} (x_t - \bar{x})^2},$$

and we have

$$\frac{Ta_x(T)^2}{a_y(T)^2} s_{\beta_1}^2 = \frac{\frac{1}{a_y(T)^2} s^2}{\frac{1}{Ta_x(T)^2} \sum_{t=1}^{T} (x_t - \bar{x})^2}$$

$$\Rightarrow \frac{\sigma_*^2}{\int_0^1 [U_{\alpha_2}(s)]^2 \, ds - \left[ \int_0^1 U_{\alpha_2}(s) \, ds \right]^2} \equiv \sigma_{*\beta_1},$$

where the weak convergence is based on items 1 and 2 of Theorem A.1 and the asymptotic behavior of $s^2$. Combining the asymptotic behaviors of $\hat{\beta}_1$ and $s_{\beta_1}$, we prove item 3 of Theorem 1.

To prove item 4 of Theorem 1, we note that the variance of $\hat{\beta}_0, s_{\beta_0}^2$ is defined as

$$s_{\beta_0}^2 = s^2 \left[ \frac{1}{T} + \frac{(\bar{x})^2}{\sum_{t=1}^{T} (x_t - \bar{x})^2} \right],$$

so we have

$$\frac{T}{a_y(T)^2} s_{\beta_0}^2 = \frac{1}{a_y(T)^2} s^2 \left[ 1 + \frac{\left( \frac{1}{Ta_x(T)} \sum_{t=1}^{T} x_t \right)^2}{\frac{1}{Ta_x(T)^2} \sum_{t=1}^{T} (x_t - \bar{x})^2} \right]$$

$$\Rightarrow \sigma_*^2 \left\{ 1 + \frac{\left[ \int_0^1 U_{\alpha_2}(s) \, ds \right]^2}{\int_0^1 [U_{\alpha_2}(s)]^2 \, ds - \left[ \int_0^1 U_{\alpha_2}(s) \, ds \right]^2} \right\} \equiv \sigma_{*\beta_0}^2,$$

where the weak convergence is based on the asymptotic behavior of $s^2$ and items 1 and 2 of Theorem A.1. Combining the asymptotic behaviors of $\hat{\beta}_0$ and $s_{\beta_0}$, we prove item 4 of Theorem 1.

https://doi.org/10.1017/S0266466699154069 Published online by Cambridge University Press