98-25

UNIVERSITY OF CALIFORNIA, SAN DIEGO

DEPARTMENT OF ECONOMICS

SPURIOUS REGRESSIONS WITH STATIONARY SERIES

BY

CLIVE W.J. GRANGER

NAMWON HYUNG

AND

YONGIL JEON

DISCUSSION PAPER 98-25
OCTOBER 1998

Spurious Regressions With Stationary Series

Clive W.J. Granger, Namwon Hyung, and Yongil Jeon
Department of Economics
University of California, San Diego
La Jolla, CA 92093-0508
USA

September 1998

Abstract: A spurious regression occurs when a pair of independent series, but with strong temporal properties, are found apparently to be related according to standard inference in an OLS regression. Although this is well known to occur with pairs of independent unit root processes, this paper finds evidence that similar results are found with positively autocorrelated autoregressive series on long moving averages. This occurs regardless of the sample size and for various distributions of the error terms.

Keywords: Autoregressions, Spurious Regressions, Inference

JEL Classification: C22

1. Introduction

Suppose that X_t, Y_t are a pair of time series independent of each other and that the regression

$$X_t = \alpha + \beta Y_t + u_t \tag{1.1}$$

is run using a standard least squares program which will also provide a *t*-statistic for the estimate of β. A "spurious relationship" can be said to have been found if the modulus of this *t*-statistic is greater than 1.96. If X_t, Y_t are a pair of independent white noise series, so that $corr\,(X_t, X_{t-k}) = 0$ for all $k \neq 0$, and similarly for Y_t, then basic statistical theory states that $|\hat{t}| \geq 1.96$ will occur approximately 5% of the time. Thus, with the definition being used, in this white noise case, an apparent, and thus spurious, relationship will be found on 5% of occasions. Using a simulation, Granger and Newbold (1974) showed that if X_t and Y_t are each drift-free random walks, then the number of spurious relationships "found" using least-squares estimation and standard inference, greatly increases to 76% of occasions with small sample sizes. A size of $N = 50$ was used in the original simulation. Later, Philips (1986) produced an elegant asymptotic theory that explained the simulation results. These developments generated a lot of interest amongst both theoretical and applied econometricians, although there was little emphasis on the ability of the asymptotic theory to explain results from simulations based on such small samples.

This paper explores the possible existence of spurious relationships between a pair of independent stationary series. It will be shown, using simulations, that spurious regressions can still occur and that the strength of the relationship varies very little with sample size. A theoretical justification for these results is also presented.

It has been known for some time that the spurious regression results do not hold only for independent random walks but also for other persistent processes, such as I(2) (Haldrup (1994)), fractional I(d) (Marmol (1996)), and stochastic unit root processes, defined by Granger and Swanson (1997), as found in unpublished simulations. The results presented here suggest that spurious relationships can be found in cases which do not involve persistent processes. This is in agreement with the original paper by Yule (1926) on nonsense correlations which observed their existence without considering links to persistence.

2. Spurious Regression Between Autoregressive Processes

Suppose that X_t and Y_t are generated by independent AR processes as follows:

$$X_t = \theta_x X_{t-1} + \varepsilon_{xt} \tag{2.1}$$

$$Y_t = \theta_y Y_{t-1} + \varepsilon_{yt} \tag{2.2}$$

where $|\theta_x| < 1$, $|\theta_y| < 1$, and ε_{xt}, ε_{yt} are each iid and zero mean. In the following simulation ε_{xt}, ε_{yt} are drawn from independent $N(0,1)$ populations, but normality is not a particularly relevant feature.

From equation (1.1) the ordinary least squares estimate of β is defined by:

$$\hat{\beta} = \frac{\sum_{t=1}^{T} (X_t - \bar{X})(Y_t - \bar{Y})}{\sum_{t=1}^{T} (X_t - \bar{X})^2} \tag{2.3}$$

where $\bar{X} = T^{-1} \sum_{t=1}^{T} X_t$ and $\bar{Y} = T^{-1} \sum_{t=1}^{T} Y_t$. Then

$$\tilde{t}_{\hat{\beta}} = \frac{\hat{\beta}}{\tilde{\sigma}_{\hat{\beta}}} = \frac{\hat{\beta}}{\hat{\sigma}_u / \sqrt{\sum_{t=1}^{T} (X_t - \bar{X})^2}} \tag{2.4}$$

where $\tilde{\sigma}_{\hat{\beta}}^2 = \hat{\sigma}_u^2 / \sum_{t=1}^{T} (X_t - \bar{X})^2$, $\hat{u}_t = Y_t - \hat{\alpha} - \hat{\beta} X_t$, $\hat{\alpha}$, $\hat{\beta}$ are the least squares estimators and $\hat{\sigma}_u^2 = \frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^2 \xrightarrow{p} \sigma_u^2$, as $T \to \infty$ and $\sigma_u^2 = \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} E(u_t^2)$. Table 1 summarized the results of a simulation using 1,000 iterations in all cases, for which $\theta_x = \theta_y = \theta$. Thus here X_t, Y_t are generated by the same autoregressive model. Sample sizes from 100 to 10,000 are shown plus an asymptotic value derived in the theory presented below. Values are for the percentage of $|t_\beta|$ values using both ordinary least squares (OLS) and with the Bartlett correction described below. Table 2 shows similar results with a variety of values of θ_x, θ_y that need not be the same.

Table 1. Regressing between two independent AR series ($\theta = \theta_x = \theta_y$). Percentage of $|t| > 1.96$.

Method	NOBS	θ=0	θ=0.25	θ=0.5	θ=0.75	θ=0.9	θ=1.0
OLS	100	5.3	7.6	13.3	29.1	51.5	77.0
	500	5.8	7.5	16.3	31.5	51.6	90.0
	2,000	5.8	7.1	13.5	29.4	52.5	94.5
	10,000	4.3	6.6	12.2	30.6	52.3	97.6
	∞	5.0	7.0	13.0	30.0	53.0	100.0
BART	100	7.6	7.7	9.9	16.5	30.6	62.0
	500	6.4	6.8	9.0	14.1	23.9	79.6
	2,000	6.0	5.9	6.1	10.3	16.3	86.4
	10,000	4.6	5.2	5.5	7.7	12.8	92.5
	∞	5.0	5.0	5.0	5.0	5.0	100.0

Notes:

1. The number of iteration = 1,000
2. % of rejection, i.e., absolute value of t-value > 1.96
3. ∞ means asymptotic case
4. To avoid the problem of fixing X_0 and Y_0, 100 pre-samples are generated and let $X_{-100} = Y_{-100} = 0$
5. The number of rejections (BART) depend on the number of lags (l) used to calculate $\hat{v}$. We set l = integer $[4(T/100)^{1/4}]$.

Table 2. Regressing Between Two Independent AR Series ($\theta_x \neq \theta_y$), percentage $|t| > 1.96$.

θ_x=0	NOBS	θ_y=0	θ_y=0.25	θ_y=0.5	θ_y=0.75	θ_y=0.9	θ_y=1.0
OLS	100	5.3	5.1	5.8	6.0	4.4	5.3
	500	5.8	5.8	5.8	6.8	6.6	5.3
	2000	5.8	5.6	5.7	6.2	6.6	4.4
	∞	5.0	5.0	5.0	5.0	5.0	5.0
BART	100	7.6	7.3	7.2	7.5	6.8	6.5
	500	6.4	6.8	7.4	7.1	6.1	5.5
	2000	6.0	5.6	5.7	6.4	6.3	4.1
θ_y=0		θ_x=0	θ_x=0.25	θ_x=0.5	θ_x=0.75	θ_x=0.9	θ_x=1.0
OLS	100	5.3	5.6	5.1	4.6	5.5	5.4
	500	5.8	6.1	7.1	4.6	5.0	5.2
	2000	5.8	5.6	5.6	5.7	6.0	4.6
	∞	5.0	5.0	5.0	5.0	5.0	5.0
BART	100	7.6	7.3	7.6	7.4	9.1	9.4
	500	6.4	6.6	7.1	6.1	5.8	5.7
	2000	6.0	6.0	5.5	5.6	6.0	5.0

Table 2 Continued

θ_x=0.5	**NOBS**	θ_y=0	θ_y=0.25	θ_y=0.5	θ_y=0.75	θ_y=0.9	θ_y=1.0
OLS	100	5.1	8.8	13.3	19.6	21.9	24.8
	500	7.1	9.9	16.3	21.2	23.8	26.2
	2000	5.6	8.4	13.5	20.3	22.5	22.8
	∞	**5.0**	**8.0**	**13.0**	**19.0**	**23.0**	**26.0**
BART	100	7.6	8.7	9.9	12.8	13.9	14.7
	500	7.1	8.1	9.0	10.0	9.8	9.3
	2000	5.5	5.8	6.1	7.3	7.8	6.3
θ_y=0.5		θ_x=0	θ_x=0.25	θ_x=0.5	θ_x=0.75	θ_x=0.9	θ_x=1.0
OLS	100	5.8	9.0	13.3	18.1	21.8	26.4
	500	5.8	10.2	16.3	19.8	22.1	26.4
	2000	5.7	10.2	13.5	20.0	23.3	24.3
	∞	5.0	8.0	13.0	19.0	23.0	26.0
BART	100	7.2	7.7	9.9	12.2	13.4	16.0
	500	7.4	8.2	9.0	9.3	9.7	10.2
	2000	5.7	6.3	6.1	8.3	7.9	6.9

The obvious feature of the simulations in Table 1 are:

(a) Spurious relationships occur quite frequently for $\theta_x = \theta_y < 1$. For example, if θ = 0.75 about 30% of regressions would lead to spurious relationships being found if standard OLS inference is used.

(b) The percentage of spurious relationships does not depend on the sample size.

(c) The Bartlett correction reduces the problem, and is thus clearly helpful, except for θ = 0.9 or larger.

The results in Table 2 illustrate the situation further. An interesting symmetry is seen to occur

(d) The percentages of $|t|$ > 1.96 are similar for ($\theta_x = a$, $\theta_y = b$) and ($\theta_x = b$, $\theta_y = a$) with $a \neq b$, although only the cases where either a or b equals either zero or 0.5 are shown.

Because of the use of an incorrect standard error estimator of $\hat{\beta}$, the *t*-statistic in (2.4) will not converge to a standard normal distribution as T increases. Serial correlation in disturbances requires a different form of consistent estimator for the standard error $\hat{\beta}$ as follows:

$$\sigma^2_{\beta} = M^{-1}\, V M^{-1}$$

where $M \equiv E\,\lfloor (X_t - \bar{X})^2 \rfloor,\ V \equiv var\,\left[T^{-1/2} \sum_{t=1}^{T} (X_t - \bar{X}_t)\, u_t\right].$

Its consistent estimator is

$$\hat{\sigma}^2_{\hat{\beta}} = \hat{M}^{-1}\, \hat{V} \hat{M}^{-1} \xrightarrow{p} \sigma^2_{\beta},\ \text{ as } T \to \infty$$

where $\hat{V} = \frac{1}{T} \sum_{t=1}^{T} (X_t - \bar{X})^2\ \hat{u}_t^2\ +\ \frac{2}{T} \sum_{s=1}^{l}\ w(s, l) \sum_{t=s+1}^{T} (X_t - \bar{X})\ \hat{u}_t\, \hat{u}_{t-s}\ \ (X_{t-s} - \bar{X}) \xrightarrow{p} V,\ \hat{M} = T^{-1}$ $\sum_{t=1}^{T} (X_t - \bar{X}) \xrightarrow{p} M$, as $T \to \infty$. Here $w(s, l)$ is an optimal weighting function that corresponds to the choice of a spectral window. In this case

$$t_{\hat{\beta}} = \frac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}} \xrightarrow{d} N(0, 1)\,. \tag{2.5}$$

In the simulations shown in Table 1 and 2 the Bartlett window

$$w(s, l) = 1 - \frac{s}{(l+1)}$$

is used, as discussed by Newey and West (1987), which guarantees the non-negativity of $\hat{\sigma}^2_{\hat{\beta}}$.

As stated before, one will get a different limiting distribution if use is made of misspecified variance estimator in (2.4) instead of a consistent estimator.

Theorem 1: Lets assume disturbances ε_{xt} and ε_{yt} are each zero mean iid processes, with variances σ^2_x and σ^2_y respectively. furthermore, $E\,|\varepsilon_{xt}|^{\,2+\lambda} < +\infty$ for some $\lambda > 0$. With regression (1.1) and X_t and Y_t generated by (2.1) and (2.2) then the limiting distribution of (2.4) is

$$\tilde{t}_{\hat{\beta}} = \frac{\hat{\beta}}{\tilde{\sigma}_{\hat{\beta}}} = \frac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}} \cdot \frac{\hat{\sigma}_{\hat{\beta}}}{\tilde{\sigma}_{\hat{\beta}}} \xrightarrow{d} N\left(0\,,\ \frac{1 - \theta_x^2 \theta_y^2}{(1 - \theta_x \theta_y)^2}\right),\ \text{ as } T \to \infty\,.$$

Proof of Theorem 1: Given model (2.1) and (2.2) with assumption of theorem, $\hat{u}_t = Y_t - \hat{\alpha} - \hat{\beta} X_t$, where $\hat{\alpha}$ and $\hat{\beta}$ be OLS estimators, we have the following properties:

$$\hat{\sigma}_u^2 = \left[\frac{1}{T} \sum_{t=1}^{T} \hat{u}_t^2 \right] \xrightarrow{p} var\,(Y_t)\ .$$

From $X_t = \sum_{i=0}^{\infty} \theta_x^i \varepsilon_{xt-i}$ and $Y_t = \sum_{i=0}^{\infty} \theta_y^i \varepsilon_{yt-i}$ we can drive $\frac{1}{T} \sum_{t=1}^{T} (X_t - \bar{X})^2 \xrightarrow{p} \frac{1}{1-\theta_x^2} \sigma_x^2$ and $\frac{1}{T} \sum_{t=1}^{T} (Y_t - \bar{Y})^2 \xrightarrow{p} \frac{1}{1-\theta_y^2} \sigma_y^2$.

We can derive the following results by the central limit theorem:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} (X_t - \bar{X})(Y_t - \bar{Y}) \xrightarrow{d} N \left(0\,,\ \frac{1-\theta_x^2 \theta_y^2}{(1-\theta_x^2)(1-\theta_y^2)(1-\theta_x \theta_y)^2} \sigma_x^2 \sigma_y^2 \right) .$$

Some tedious algebra is required to get this.

This result, together with the fact that the limiting distribution is quickly reached as the sample size increases explains all of the features (a) to (d) observed in the simulation.

Two extensions to the simulations were considered but will just be discussed here rather than showing the complete results. The effectiveness of using a *t*-statistic based on White's heteroskedastic robust procedure found that misleading inferences would not have been reduced if the procedure was used. It is not designed to be robust against autocorrelated errors.

The simulations were reproduced using different distributions for the shocks than normal. Each pair of shocks remains iid and independent of each other, but the distribution used was exponential. Cauchy, Laplace, uniform and *t*-distribution with 5 degrees of freedom. All gave very similar results for table 1, with the exception of the Cauchy, where the N = 100 results were similar but now the extent of the problem declines with sample size. Table 3 shows OLS results for the exponential (with mean = 1) and Laplace (or double exponential) distribution with mean = 0 and variance = 2.

3. Spurious Regressions Using Independent MA(k) Processes

In this section it will be shown that spurious relationships may be found between other pairs of independent stationary series that are somewhat persistent. If X_t and Y_t are generated by independent MA(k) processes

$$X_t = \sum_{j=0}^{k} e_{x,t-j}\ , \quad Y_t = \sum_{j=0}^{k} e_{y,t-j} \tag{3.1}$$

where e_{xt}, e_{yt} are each zero mean iid processes. Table 4 shows the percentage of times $|t_\beta| \geq 1.96$ from OLS regressions (3.1) for *k* taking the values 0, 1, 2, 5, 10, 20, or 50. Sample sizes vary from 100 to 10,000 and corresponding results are shown when the Bartlett correction is used. The results agree with those of the previous section. Apparent or spurious regressions occur frequently, roughly one third of the time when *k* is only 5; the values in the table increase steadily with *k* but do not vary with sample size, and the Bartlett correction is generally useful.

Table 3. Spurious Regression With Non-normal Distributions

OLS	NOB	θ=0	θ=0.25	θ=0.5	θ=0.75	θ=0.9	θ=1.0
Case I	100	6.0	8.3	15.7	32.7	55.8	78.5
Laplace	500	3.7	5.9	11.9	29.6	53.0	90.4
	2,000	5.5	6.7	13.9	30.2	52.0	94.9
	10,000	4.8	5.6	11.8	29.1	50.5	97.3
Case II	100	5.1	6.6	11.2	23.1	49.3	79.4
Cauchy	500	2.9	3.4	6.3	12.5	36.0	92.6
	2,000	2.2	3.1	5.5	10.7	26.9	94.2
	10,000	0.5	0.9	2.3	5.5	15.1	98.0
Case III	100	6.9	8.2	14.7	29.9	51.2	100.0
Exponential	500	4.8	6.9	14.2	30.8	52.9	100.0
	2,000	4.4	5.8	13.2	29.0	53.3	100.0
	10,000	4.3	5.9	11.5	32.0	53.7	100.0

Table 4. Regressing Between Two Indpendent MA Series

Method	Number of data	White Noise	MA(1)	MA(2)	MA(5)	MA(10)	MA(20)	MA(50)
OLS	100	6.4	9.3	17.1	30.6	45.3	60.2	76.0
	500	5.8	12.8	20.7	35.4	48.0	59.4	71.9
	2,000	6.3	12.1	20.2	33.8	46.1	59.4	75.7
	10,000	4.4	9.9	17.4	33.5	46.3	59.4	74.1
BART	100	7.1	8.1	10.6	13.8	22.5	34.4	58.2
	500	6.1	8.2	10.6	11.7	11.5	12.1	19.7
	2,000	6.3	8.3	9.0	10.1	11.4	10.3	14.3
	10,000	4.6	6.8	8.2	9.2	9.9	10.1	10.4

1. The number of iteration = 1,000
2. % of rejection, i.e., absolute value of *t*-value > 1.96

Theorem 2: The disturbance e_{xt}, e_{yt} are each zero mean iid processes, and variances σ_x^2, σ_y^2 respectively. Furthermore, $E|e_{xt}|^{2+\lambda} < +\infty$ for some $\lambda < 0$. With regression (1.1) and X_t, Y_t generated by (3.1) then as $T \to \infty$:

(a) if $k = O(T)$, $\hat{\beta} = O_p(1)$

(b) if $k = O(1)$, $\hat{\beta} = O_p(T^{-1/2})$

(c) if $k = O(T^l)$ with $0 < l < 1$, $\hat{\beta} = O_p(T^{-1/2} k^{1/2}) = O_p(T^{1/2(l-1)})$ and

furthermore, $\hat{u}_t = Y_t - \hat{\alpha} - \hat{\beta} X_t$, where $\hat{\alpha}$, $\hat{\beta}$ be OLS estimators, and if $k = O(T^l)$ with $0 < l < 1$,

(d) $\tilde{t}_{\hat{\beta}} = \dfrac{\hat{\beta}}{\tilde{\sigma}_{\beta}} = O_p(k^{1/2})$

(e) $t_{\hat{\beta}} = \dfrac{\hat{\beta}}{\hat{\sigma}_{\hat{\beta}}} = O_p(1)$

Lemma 1: Given model (3.1) with assumption of Theorem 2, we have the following properties:

(a) $E\left(\frac{1}{Tk}\sum_{t=1}^{T}(X_t - \bar{X})^2\right) = O(1)$ and $var\left(\frac{1}{Tk}\sum_{t=1}^{T}(X_t - \bar{X})^2\right) = O(1)$

(b) $E\left(\frac{1}{T^{1/2}k^{3/2}}\sum_{t=1}^{T}(X_t-\bar{X})(Y_t-\bar{Y})\right)=0$ and $Var\left(\frac{1}{T^{1/2}k^{3/2}}\sum_{t=1}^{T}(X_t-\bar{X})(Y_t-\bar{Y})\right)=O(1)$.

Let $\xi_{xx}=\sum_{t=1}^{T}(X_t-\bar{X})^2$ and $\xi_{xy}=\sum_{t=1}^{T}(X_t-\bar{X})(Y_t-\bar{Y})$.

Proof of Theorem (a): If $k=O(T)$ and $e_{xt}=e_{yt}=0$ for all $t\le 0$, by assumption of Theorem 2 and Lemma 1 (a), then $T^{-2}\xi_{xy}=O_p(1)$ and $T^{-2}\xi_{xx}=O_p(1)$. Then formula (a) of theorem follows directly. The result shows a spurious regression which was mentioned by Granger and Newbold (1974).

Proof of Theorem (b): If $k=O(1)$ then $T^{-1/2}\xi_{xy}=O_p(1)$ and $T^{-1}\xi_{xx}\to M<\infty$ in probability for some positive number M. Then $\hat{\beta} = T^{-1/2}T^{-1/2}\xi_{xy}/(T^{-1}\xi_{xx})=O_p(T^{-1/2})$.

Proof of Theorem (c): If $k=O(T^l)$ with $0<l<1$ and $e_{xt}=e_{yt}=0$ for all $t\le 0$, then $T^{-1/2}k^{-3/2}\xi_{xy}=O_p(1)$ and $T^{-1}k^{-1}\xi_{xx}\to M'<\infty$ in probability for some positive number M'. Then $\hat{\beta}=T^{-1/2}k^{1/2}T^{-1/2}k^{3/2}\xi_{xy}/T^{-1}k^{-1}\xi_{xx}=O_p(T^{-1/2}k^{1/2})$.

Lemma 2: Given model (3.1) with Assumption 1, $\hat{u}_t=Y_t-\hat{\alpha}-\hat{\beta}X_t$, where $\hat{\alpha}$, $\hat{\beta}$ be OLS estimators, and if $k=O(T^l)$ with $0<l<1$ and $e_{xt}=e_{yt}=0$ for all $t\le 0$, we have the following properties:

(a) $\tilde{\sigma}_{\hat{\beta}}^2=\left(\frac{1}{T}\sum_{t=1}^{T}\hat{u}_t^2\right)\left(\sum_{t=1}^{T}(X_t-\bar{X})^2\right)^{-1}=O_p(T^{-1})$.

(b) $\hat{\sigma}_{\hat{\beta}}^2=\hat{M}^{-1}\hat{V}\hat{M}^{-1}=O_p(T^{-1}k)$.

Proof of Theorem (d): By formula (c), lemma 2 (a)

$$k^{-1/2}\tilde{t}_{\hat{\beta}}=\frac{T^{1/2}k^{-1/2}\hat{\beta}}{T^{1/2}\tilde{\sigma}_{\hat{\beta}}}=O_p(1)\,.$$

Proof of Theorem (e): Similarly, by formula (c), lemma 2 (b)

$$t_{\hat{\beta}}=\frac{T^{1/2}k^{-1/2}\hat{\beta}}{T^{1/2}k^{-1/2}\hat{\sigma}_{\hat{\beta}}}=O_p(1)\,.$$

The theorem shows that the conventional t-ratio does not have limiting distributions if $l > 0$. In fact, the distribution of $t_{\hat{\beta}}$ diverges as k increases, not decreases as T increases if a misspecified variance estimator is used.

4. Conclusions

This is a paper that could have been written 25 years ago, but it still tries to issue a warning. Concerns about spurious regressions have been widely discussed for some time, but largely in the context of a pair of unit root or persistent processes. It is here shown that the problem arises in a wider context; although less seriously. Spurious relationships occur from a misspecified model, under the null. If X_t is serially correlated in (1.1) and the true β is zero, this implies that the residuals cannot be white noise. The problem is thus that (1.1) is a misspecified model. It is resolved by improving the specification; that is by adding more lagged dependent, and possibly lagged independent, variables. Patch-work procedures, such as the Cochrane-Orcutt correction, will be inefficient compared to using a wider specification. Strategies that have been suggested, such as testing for I(1) versus I(0) and then possibly building (1.1) with differenced data, are irrelevant when stationary variables are involved.

There are a number of consequences of these results. Clearly care has to be taken when interpreting regressions, however estimated, if insufficient care is taken with serial correlations in the residuals. This can occur if an overly simple specification is chosen, or possibly one based on a model selection criterion, such as BIC. Some estimated error-correction model equations could provide examples. The results also provide a simple explanation of why interpretation of important coefficients change when alternative specifications are considered of a dynamic model.

References

Granger, C.W.J. and P. Newbold (1974): "Spurious Regressions in Economics." *Journal of Econometrics* 2, 111-120.

Granger, C.W.J. and N. Swanson (1997): "An Introduction to Stochastic Unit-Root Processes." *Journal of Econometrics* 80, 35-62.

Haldrup, N. (1994): "The Asymptotics of Single-Equation cointegration Regressions with I(1) and I(2) Variables." *Journal of Econometrics* 63, 153-181.

Marmol, F. (1996): "Nonsense Regressions Between Integrated Processes of Different Orders." *Oxford Bulletin of Econometics and Statistics* 58, 525-536.

Newey, W.K. and K.D. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55, 703-708.

Phillips, P.C.B. (1986): "Understanding Spurious Regressions in Econometrics." *Journal of Econometrics* 33, 311-340.

Yule, G.U. (1926): "Why Do We Sometimes Get Nonsense Correlations Between Time Series? A Study in Sampling and the Nature of Time Series." *Journal of the Royal Statistical Society* 89, 1-64.