Unit Roots and Co-integration

Topic 4: Spurious Regressions

by Vanessa Berenguer-Rico

University of Oxford

January 2014

Spurious Regressions

- Spurious regressions have a long tradition in statistics
- Yule (1926): "Why do we sometimes get nonsense-correlations between time-series?"
- *"It is fairly familiar knowledge that we sometimes obtain between quantities varying with the time (time-variables) quite high correlations to which we cannot attach any physical significance whatever, although under the ordinary test the correlation would be held to be certainly "significant""*

Spurious Regressions

- Real Data Example: Yule (1926)

- Proportion of Church of England marriages to all marriages for the years 1866-1911 inclusive and standarized mortality per 1000 persons for the same years

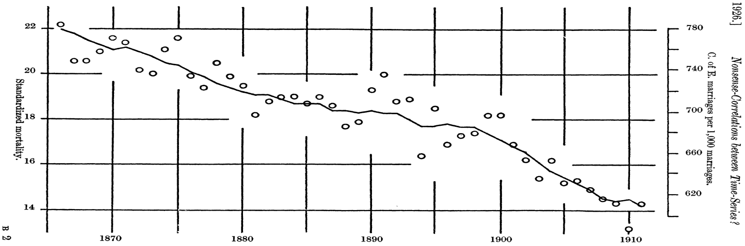

FIG. 1.—Correlation between standardized mortality per 1,000 persons in England and Wales (circles), and the proportion of Church of England marriages per 1,000 of all marriages (line), 1866–1911. $r = +0{\cdot}9512$.

3

- Correlation coefficient = 0.9512

Spurious Regressions

- *"As the occurrence of such "nonsense-correlations" makes one mistrust the serious arguments that are sometimes put forward on the basis of correlations between time series (...) it is important to clear up the problem* ***how they arise and in what special cases****."*

- *"When we find that a theoretical formula applied to a particular case gives results which common sense judges to be incorrect, it is generally as well to* ***examine the particular assumptions*** *from which it was deduced, and see which of them are inapplicable to the case in point."*

- *"The* ***special case*** *considered in the next section will suffice to show that when the successive x's and y's in a sample no longer form a random series, but a series in which* ***succesive terms are closely related to one another****, the usual conceptions to which we are accustomed fail totally and entirely to apply."*

Spurious Regressions

- Yule (1926): Experimental excercise
- He simulated the distribution of the empirical correlation coefficient calculated from two independent i.i.d. processes, from two independent random walks, and from two independent cumulated random walks

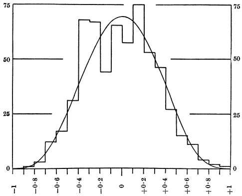

Fig. 16.—Frequency-distribution of 600 correlations between samples of 10 observations from random series (Table VIII).

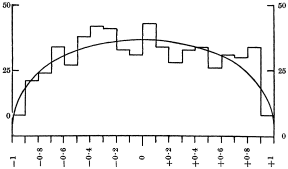

Fig. 17.—Frequency-distribution of 600 correlations between samples of 10 observations from conjunct series with random differences (Series X_1, Table IX).

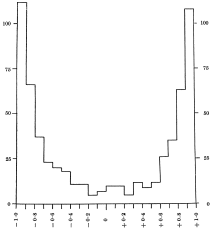

Fig. 18.—Frequency-distribution of 600 correlations between samples of 10 observations from conjunct series with conjunct differences (Table XI).

Spurious Regressions

- Granger and Newbold (1974): "Spurious Regressions in Econometrics"

- *"It is very common to see reported in applied econometric literature time series regression equations with an apparently high degree of fit, measured by the coefficient of multiple correlation R^2 or the corrected coefficient $\bar{R}^2$, but with an extremely low value for the Durbin-Watson statistic."*

- How nonsense regressions can arise? A simulation study:

- Let $y_t = y_{t-1} + v_t$ and $x_t = x_{t-1} + w_t$ with $y_0 = O_p(1)$, $x_0 = O_p(1)$, and $v_t \sim i.i.d.\,(0,1)$ independent of $v_t \sim i.i.d.\,(0,1)$ and consider the least squares regression

$$y_t = \hat{\alpha} + \hat{\beta} x_t + \hat{u}_t, \quad t = 1, ..., T$$

Spurious Regressions

Let $y_t = y_{t-1} + v_t$ and $x_t = x_{t-1} + w_t$ with $y_0 = O_p(1)$, $x_0 = O_p(1)$, and $v_t \sim i.i.d.(0,1)$ independent of $v_t \sim i.i.d.(0,1)$ and consider the least squares regression $y_t = \hat{\alpha} + \hat{\beta} x_t + \hat{u}_t, \quad t = 1, ..., T$

A draw	$T = 99$	$T = 999$
$\hat{\alpha}$	-0102	-11.658
$t_{\hat{\alpha}}$	-0.170	-26.032
$\hat{\beta}$	0.605	0.610
$t_{\hat{\beta}}$	10.821	65.532
R^2	0.546	0.811
DW	0.151	0.032

Spurious Regressions

A Monte Carlo		percentiles			
		5%	50%	95%	
$T = 99$	$t_{\hat{\alpha}}$	-17.384	0.361	18.955	
	$t_{\hat{\beta}}$	-10.987	0.012	11.247	$rf_{1.96}$=0.756
	R^2	0.001	0.168	0.659	$m(R^2)$=0.232
	DW	0.047	0.144	0.372	$m(DW)$=0.171
$T = 999$	$t_{\hat{\alpha}}$	-56.073	-1.656	56.769	
	$t_{\hat{\beta}}$	-37.819	0.516	39.271	$rf_{1.96}$=0.927
	R^2	0.001	0.171	0.695	$m(R^2)$=0.240
	DW	0.004	0.016	0.038	$m(DW)$=0.018

Spurious Regressions

- Phillips (1986): "Understanding Spurious Regressions in Econometrics"
- *"The present paper develops an **asympotitc theory** for regressions that relate quite general integrated random processes. This includes **spurious regressions** of the Granger-Newbold type as a special case. It turns out that the correct asymptotic theory goes a long way towards explaining the experimental results that these authors obtained. In many cases their findings are quite predictable from the true asymptotic behavior of the relevant statistics."*

Spurious Regressions

- *"Thus, our theory demonstrate that in the Granger-Newbold regressions of independent random walks the usual t-ratio significance test does not possess a limiting distribution, but actually diverges as the sample size* $T \to \infty$*."*
- *"Inevitably, therefore, the bias in this test towards the rejection of no relationship (based on a nominal critical value of 1.96) will increase with* T*."*
- *"We.also show that the Durbin-Watson statistic actually converges in probability to zero, while the regression* R^2 *has a non-degenerate limiting distribution as* $T \to \infty$*."*

Spurious Regressions

- Let $\{\xi_t\}_1^\infty$ be a sequence of random n-vectors defined on a probability space $(\Omega, \mathcal{F}, \Upsilon)$ and $S_t = \sum_{j=1}^t \xi_j$ be the partial sum process and set $S_0 = 0$
- <u>*Assumption 1*</u>

  (a) $E(\xi_t) = 0$ for all t;
  (b) $\sup_{i,t} E|\xi_{it}|^{\beta+\varepsilon} < \infty$ for some $\beta > 2$ and $\varepsilon > 0$;
  (c) $\Sigma = \lim_{T\to\infty} T^{-1} E(S_T S_T')$ exists and is positive definite;
  (d) $\{\xi_t\}_1^\infty$ is strong mixing with mixing numbers α_m satisfying

$$\sum_{1}^{\infty} \alpha_m^{1-2/\beta} < \infty.$$

- If $\xi_t' = (v_t, w_t)$, then the Granger-Newbold framework applies

Remark: Mixing: Let $\{X_t\}_{-\infty}^{\infty}$ be a stochastic process defined in a probability space $(\Omega, \mathcal{F}, P)$ and let $\mathcal{F}_{-\infty}^{t} = \sigma\left(..., X_{t-2}, X_{t-1}, X_t\right)$ and $\mathcal{F}_{-\infty}^{t} = \sigma\left(X_{t+m}, X_{t+m+1}, X_{t+m+2}...\right)$. The sequence is said to be α-mixing (or strong mixing) if

$$\lim_{m\to\infty} \alpha_m = 0,$$

where

$$\alpha_m = \sup_t \alpha\left(\mathcal{F}_{-\infty}^{t}, \mathcal{F}_{t+m}^{\infty}\right)$$

and

$$\alpha\left(\mathcal{F}_{-\infty}^{t}, \mathcal{F}_{t+m}^{\infty}\right) = \sup_{G\in\mathcal{F}_{-\infty}^{t}, H\in\mathcal{F}_{t+m}^{\infty}} \left|P\left(G\cap H\right) - P\left(G\right)P\left(H\right)\right|.$$

Spurious Regressions

- Moreover, when $\xi_t' = (v_t, w_t)$ and v_t and w_t are independent, as in the spurious regressions context, we have

$$\Sigma = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_w^2 \end{bmatrix},$$

where

$$\sigma_v^2 = \lim_{T\to\infty} T^{-1} E\left[\left(\sum_1^t v_j\right)^2\right] \text{ and } \sigma_w^2 = \lim_{T\to\infty} T^{-1} E\left[\left(\sum_1^t w_j\right)^2\right]$$

Spurious Regressions

Lemma SR. Suppose $y_t = y_{t-1} + v_t$ and $x_t = x_{t-1} + w_t$. If the innovation sequences $\{v_t\}_1^\infty$ and $\{w_t\}_1^\infty$ are independent and if $\{(v_t, w_t)\}_1^\infty$ satisfies Assumption 1 then, as $T \to \infty$,

(a)

$$T^{-3/2} \sum_1^T x_t \Longrightarrow \sigma_w \int_0^1 W(r)\, dr \text{ and } T^{-3/2} \sum_1^T y_t \Longrightarrow \sigma_v \int_0^1 V(r)\, dr,$$

(b)

$$T^{-2} \sum_1^T x_t^2 \Longrightarrow \sigma_w^2 \int_0^1 W(r)^2\, dr \text{ and } T^{-2} \sum_1^T y_t^2 \Longrightarrow \sigma_v^2 \int_0^1 V(r)^2\, dr,$$

(c)

$$T^{-2}\sum_{1}^{T}(x_t-\bar{x})^2 \Longrightarrow \sigma_w^2\left[\int_0^1 W(r)^2\,dr-\left\{\int_0^1 W(r)\,dr\right\}^2\right]$$

$$T^{-2}\sum_{1}^{T}(y_t-\bar{y})^2 \Longrightarrow \sigma_v^2\left[\int_0^1 V(r)^2\,dr-\left\{\int_0^1 V(r)\,dr\right\}^2\right],$$

(d)

$$T^{-2}\sum_{1}^{T}y_t x_t \Longrightarrow \sigma_v\sigma_w\int_0^1 V(r)\,W(r)\,dr$$

(e)

$$T^{-1}\sum_{r}^{T} y_t\left(y_t - y_{t-r}\right) \Longrightarrow (r/2)\left\{\sigma_v^2 V(1)^2 + \Omega_{v0}\right\} + \sum_{j=1}^{r}(r-j)\,\Omega_{vj},$$

$$T^{-1}\sum_{r}^{T} x_t\left(x_t - x_{t-r}\right) \Longrightarrow (r/2)\left\{\sigma_w^2 W(1)^2 + \Omega_{w0}\right\} + \sum_{j=1}^{r}(r-j)\,\Omega_{wj},$$

(f)

$$T^{-1}\sum_{r}^{T} y_t\left(x_t - x_{t-r}\right) + T^{-1}\sum_{r}^{T} x_t\left(y_t - y_{t-r}\right) \Longrightarrow r\sigma_v\sigma_w V(1)\,W(1)\,,$$

Spurious Regressions

where $W(r)$ and $V(r)$ are independent Wiener processes on $C[0,1]$ and where for $j = 0, 1, ...$

$$\Omega_{vj} = \lim_{T\to\infty} T^{-1} \sum_{j+1}^{T} E\left(v_t v_{t-j}\right) \text{ and } \Omega_{wj} = \lim_{T\to\infty} T^{-1} \sum_{j+1}^{T} E\left(w_t w_{t-j}\right).$$

Moreover (a)-(f) hold irrespective of the initial conditions assigned to y_0 and x_0.

Theorem SR. Suppose that

$$y_t = \hat{\alpha} + \hat{\beta} x_t + \hat{u}_t, \quad t = 1, ..., T$$

is estimated by least squares regression and the conditions of Lemma 1 are satisfied. Then, as $T \to \infty$,

(a)

$$\hat{\beta} \Longrightarrow \frac{\sigma_v \left\{ \int_0^1 V(r) W(r) dr - \int_0^1 V(r) dr \int_0^1 W(r) dr \right\}}{\sigma_w \left\{ \int_0^1 W(r)^2 dr - \left(\int_0^1 W(r) dr \right)^2 \right\}} = (\sigma_v / \sigma_w) \zeta,$$

(b)

$$T^{-1/2} \hat{\alpha} \Longrightarrow \sigma_v \left\{ \int_0^1 V(r) dr - \zeta \int_0^1 W(r) dr \right\},$$

(c)

$$T^{-1/2} t_{\beta} \Longrightarrow \frac{\mu}{v^{1/2}},$$

where

$$\mu = \int_0^1 V(r) W(r) \, dr - \int_0^1 V(r) \, dr \int_0^1 W(r) \, dr,$$

and

$$\begin{aligned} v &= \left\{ \int_0^1 V(r)^2 dr - \left(\int_0^1 V(r) \, dr \right)^2 \right\} \left\{ \int_0^1 W(r)^2 dr - \left(\int_0^1 W(r) \, dr \right)^2 \right\} \\ &\quad - \left\{ \int_0^1 V(r) W(r) \, dr - \int_0^1 V(r) \, dr \int_0^1 W(r) \, dr \right\}^2 . \end{aligned}$$

Spurious Regressions

(d)

$$T^{-1/2}t_{\alpha} \Longrightarrow \frac{\left\{\int_0^1 V(r)\,dr - \zeta\int_0^1 W(r)\,dr\right\}\left\{\int_0^1 W(r)^2\,dr - \left(\int_0^1 W(r)\,dr\right)^2\right\}}{\left[v\int_0^1 W(r)^2\,dr\right]^{1/2}}$$

(e)

$$R^2 \Longrightarrow \frac{\zeta^2\left\{\int_0^1 W(r)^2\,dr - \left(\int_0^1 W(r)\,dr\right)^2\right\}}{\int_0^1 W(r)^2\,dr - \left(\int_0^1 W(r)\,dr\right)^2}$$

(f)

$$DW \xrightarrow{p} 0 \quad \text{and} \quad TDW = O_p(1)\,.$$