A comment on ‘resolving spurious regressions and serially correlated errors’

Ventosa-Santaulària, Daniel; Vera-Valdés, J. Eduardo; Martínez-Olmos, Alejandra I.

Published in:
Empirical Economics

DOI (link to publication from Publisher):
10.1007/s00181-015-1035-7

Publication date:
2016

Document Version
Accepted author manuscript, peer reviewed version

A comment on ‘Resolving spurious regressions and serially correlated errors’

Daniel Ventosa-Santaulària* J. Eduardo Vera-Valdés† Alejandra I. Martínez Olmos‡

Abstract

In order to diminish size distortions of the t-test in a time-series linear specification, Agiakloglou (2013) [Empir Econ (2013) 45:1361-1366] proposed to (1) include the first lag of the dependent variable as a regressor, or (2) estimate it using the first differences of the variables. He provided finite-sample evidence to support his proposal. In this paper we extend the Monte Carlo experiment to different data-generating processes, and calculate the asymptotic behavior of the modified specifications. We show that including the lag of the dependent variable as a regressor reduces size distortions when the variables are driftless unit roots, but this approach does not hold under the presence of long memory, nonlinearities or structural breaks.

Keywords: Spurious regression, autocorrelation corrective methods, unit roots, shifts, fractional integration
JEL clasification: C22

The spurious regression phenomenon occurs when a statistically significant relationship between two or more variables (stationary or not) is inferred when in fact there is none. It has been an ongoing topic in time series analysis since its discovery in the late twenties (Yule, 1926), its rediscovery in time-series econometrics by Granger and Newbold (1974), and the elegant asymptotic explanation of this phenomenon (assuming driftless unit root processes as the data-generating process, DGP) by Phillips (1986). In the last three decades many authors extended the *realms* of spurious regression by proposing more elaborate DGPs and/or estimators (see Ventosa-Santaulària 2009 for a review), the main insight being relatively simple: Econometricians must be cautious when making empirical studies given the ubiquity of the phenomenon of spurious regression when modelling economic times series.

* *Corresponding Author*: Centro de Investigación y Docencia Económicas (CIDE) Carretera México-Toluca 3655 Col. Lomas de Santa Fe 01210 México City, Mexico.
Email: daniel.ventosa@cide.edu
Phone: +52 01 800 021 2433.

† CREATES and Aarhus University. The author acknowledges support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation.

‡ Universidad de Guanajuato, Departamento de Economía y Finanzas

Nevertheless, several authors, such as McCallum (2010), Kolev (2011), and Agiakloglou (2013), *inter alia*, consider that the spurious regression phenomenon is rather due to a poorly-controlled autocorrelation problem. They further propose that an estimation procedure that controls for its presence eliminates the spurious regression phenomenon. McCallum (2010), for instance, proposed a Cochrane-Orcutt procedure to avoid nonsense inference, but Martínez-Rivera and Ventosa-Santaulària (2012) provided evidence that his proposal does not always work.

Recently, Agiakloglou (2013) suggested that (i) adding the first lag of the dependent variable as a regressor (equation 3) or, (ii) using the variables in first difference (equation 2), instead of the original regression (equation 1), reduces the size distortion of t-statistics (under both, stationarity and nonstationarity). Agiakloglou shows the effectiveness of his proposal through a Monte Carlo experiment. The latter is carried out simulating 1,000 replications, each of which generates two independent series, x_t and y_t, which follow an AR(1) process with different parametric settings. The average rejection rate of the t-ratio (as well as the average value of the Durbin-Watson statistic and the average rejection rate of an ARCH(1) test), shows that his autocorrelation-correcting procedure works well in finite samples.

$$y_t = \alpha + \beta x_t + \epsilon_t, \tag{1}$$

$$\Delta y_t = \alpha + \beta \Delta x_t + \epsilon_t, \tag{2}$$

$$y_t = \alpha + \beta x_t + \delta y_{t-1} + \epsilon_t. \tag{3}$$

Agiakloglou concludes that the model with lags (3) eliminates, in most cases (and for rather small values of the autoregressive parameter), the evidence of spurious results and serially correlated errors, while the model with differences (2) eliminates the size distortions when the AR parameter is large.

Although Agiakloglou's proposal is interesting and valuable, its scope remains limited given that it is only based on Monte Carlo simulations (no theoretical arguments are provided), the employed DGP, an AR(1), is rather simple, and it remains unclear whether the latter can be considered as representative of most economic time series. Also, his Monte Carlo evidence is built upon the specification of the same AR parameter to generate both, x_t and y_t in each simulation. Given the potential usefulness of Agiakloglou's proposal, we aim to cope with such limitations and thus clarify its reach. Our strategy is threefold:

We first provide a new set of DGPs, close to the one employed by Agiakloglou (all the processes are stationary), but allowing for constant terms and different parameter values for x_t and y_t (this is, the constants, the autogressive terms and the variances are different for each variable). The results appear in Table (1).

The rejection rates presented in Table (1) confirm Agiakloglou's findings: the size distortions are greatly reduced when the regression in estimated in first differences or with a lag of the dependent variable. Model (2) is particularly useful when the AR parameters are large, whilst model (3) when the AR parameters are rather small. More importantly, these simulations also provide some evidence that both proposals are useful when the DGP parameters are not the same for each variable, something not considered in Agliakloglou's work. Moreover, the presence of constant terms does not seem to affect the size-distortion reduction capabilities of the proposed specifications. It is noteworthy to emphasize that the simulated processes in table (1) are stationary.

T	Model	AR(1)					AR(2)
		$\phi_y = 0.10$	$\phi_y = 0.30$	$\phi_y = 0.50$	$\phi_y = 0.70$	$\phi_y = 0.90$	$\phi_{1,y} = -0.30, \phi_{2,y} = 0.20$
		$\phi_x = 0.25$	$\phi_x = 0.45$	$\phi_x = 0.65$	$\phi_x = 0.85$	$\phi_x = 0.95$	$\phi_{1,x} = 0.10, \phi_{2,x} = -0.50$
100	(1)	0.065	0.082	0.156	0.298	0.558	0.018
	(2)	0.101	0.079	0.063	0.067	0.052	0.046
	(3)	0.056	0.058	0.059	0.071	0.089	0.035
200	(1)	0.055	0.070	0.164	0.327	0.594	0.015
	(2)	0.086	0.076	0.078	0.060	0.061	0.041
	(3)	0.047	0.042	0.047	0.062	0.089	0.029
500	(1)	0.063	0.088	0.166	0.329	0.587	0.023
	(2)	0.116	0.076	0.068	0.047	0.047	0.035
	(3)	0.057	0.056	0.045	0.049	0.060	0.042
1,000	(1)	0.068	0.079	0.178	0.322	0.587	0.018
	(2)	0.094	0.073	0.068	0.059	0.045	0.040
	(3)	0.065	0.047	0.053	0.053	0.058	0.036

Table 1: *Proportion of rejection of the null hypothesis for the t-test with a 5% nominal value. All the simulated processes are stationary. The constant term and the standard deviation of y_t and x_t,* are 0.3 and 1.8, and 0.5 and 1.4, respectively.

In the second part of our strategy, we provide the theoretical underpinning for Agiakloglou' procedure for the unit root case, and show that his proposal is correct except that a nonstandard asymptotically pivotal distribution should be employed instead of the classical Normal distribution under the null hypothesis; this result is summarized in the following Theorem.

Theorem. *Let $\{y_t\}_{t=1}^{\infty}$ and $\{x_t\}_{t=1}^{\infty}$ be two independent $I(1)$ processes, $z_t = z_{t-1} + u_{zt}$, for $z = x, y$, $u_{zt} \sim iid(0, \sigma_z^2)$. Then, as $T \to \infty$:*
***1.** Model (2):*

$$T^{1/2}\hat{\beta} \overset{D}{\to} \frac{\omega_{xy}(1)\sigma_y}{\sigma_x},$$
$$t_{\hat{\beta}} \overset{D}{\to} \omega_{xy}(1),$$

***2.** Model (3):*

$$T\hat{\beta} \overset{D}{\to} \frac{\sigma_y}{\sigma_x}\gamma_y \frac{\int \omega_x d\omega_y + [\frac{1}{2}(\omega_y(1)^2 - 1)]\gamma_{xy} + \omega_y(1)[\int \omega_x \int \omega_y^2 - \int \omega_x\omega_y \int \omega_y]}{\gamma_y \int \omega_x^2 + (\int \omega_x)^2 \int \omega_y^2 - 2\int \omega_x \int \omega_y \int \omega_x\omega_y + (\int \omega_x\omega_y)^2},$$
$$t_{\hat{\beta}} \overset{D}{\to} \frac{\gamma_y \int \omega_x d\omega_y + [\frac{1}{2}(\omega_y(1)^2 - 1)]\gamma_{xy} + \omega_y(1)[\int \omega_x \int \omega_y^2 - \int \omega_x\omega_y \int \omega_y]}{\sqrt{\gamma_y[\int \omega_x^2\gamma_y + (\int \omega_x)^2 \int \omega_y^2 - 2\int \omega_x \int \omega_y \int \omega_x\omega_y + (\int \omega_x\omega_y)^2]}},$$

with

$$\gamma_y := \left(\int \omega_y\right)^2 - \int \omega_y^2 \qquad \textit{and,} \qquad \gamma_{xy} := \int \omega_y\omega_x - \int \omega_x \int \omega_y,$$

where $\overset{D}{\to}$ denotes convergence in distribution, $\omega_z(r)$ represents a standard Wiener process, and $\int \omega_z(r)$ represents the stochastic integral $\int_0^1 \omega_z(r)dr$.

Proof: see appendix A.
Under model (2) the estimated β collapses to zero at a rate of $T^{-1/2}$ while t_β converges to a standard normal. This last result should come as no surprise given that we are under a process I(1), which, when differentiated, becomes an I(0) process. The Theorem also shows that under model (3) the statistic t_β does not converge to a standard normal but to a nonstandard distribution (a function of Brownian motions) without unknown nuisance parameters. Given the above, a new set of critical values must be computed. We did so using a Monte Carlo simulation with $100,000$ replications for different sample sizes. The critical values are shown in Table 2.

T	1%	5%	10%
50	± 3.4677	± 2.7057	± 2.3088
100	± 3.4123	± 2.6912	± 2.3079
250	± 3.3847	± 2.6852	± 2.3055
500	± 3.3733	± 2.6703	± 2.3068
∞	± 3.3685	± 2.6752	± 2.3031

Table 2: *Critical values of the t-ratio associated to $\hat{\beta}$ in eq. (3).*

Note, however, that these critical values are adequate only under the assumption of driftless unit roots with iid noise at the base. The inclusion, for example, of a drift in the DGP (not considered in Agiakloglou) invalidates them. To clarify and extend the later point, we present the third part of our strategy by providing Monte Carlo evidence using different DGPs. It would be a strong assumption in practice to consider that economic time series can always be represented as AR(1) processes. By allowing a wider variety of DGPs, we show that the size-reduction capacities of Agiakloglou's procedure can be considerably lower in some cases. In other words, we obtain evidence that the spurious regression phenomenon is not systematically eliminated when applying Models (2) and (3). On Table 3 we present the proportion of rejections (out of $1,000$ replications for samples sizes of $T = 100$, 200, 500, and $1,000$) of the null hypothesis for the t-ratio with a nominal value of 5%. We considered six different DGPs: ARFIMA(0,d,0), AR(2), I(1), 3-Regime SETAR, and I(1)+drift+break. All the DGPs are nonstationary, except the (nonlinear) 3-Regime SETAR, which is globally stationary although the corridor regime is not (for y_t). Note also that all the considered DGPs (except the I(1) DGP) include deterministic components. The parameters of the simulations appear in Table 4 in the appendix. For reasons of space we only present a subset of our simulation study. The complete experiment is available as supplementary material.
Table (3) allow us to make two main observations: First, note that the proportion of rejections of the null hypothesis in all cases (except the I(1)* case, where correct critical values have been used) is greater than the nominal value of 5%. This indicates that if the practitioner were to work with variables generated by some of these DGPs, which she probably would not know *ex-ante*, by following Agiakloglou's procedure, the risk of erroneously conclude that there is a relation between the variables, when in fact there is none, is greater. That said, note that, for all DGPs considered in Table (3), Agiakloglou's alternative specifications still reduce considerably the size distortions; for the nonstationary ARFIMA case, for instance, size distortions are reduced from around 0.8 to close to

T	Model	ARFIMA(0,d,0)	AR(2)	ARIMA(1,1,0)	I(1)*	3-R SETAR	I(1)+drift +breaks
100	(1)	0.695	0.954	1.000	—	0.965	0.999
	(2)	0.058	0.147	0.258	—	0.113	0.304
	(3)	0.175	0.656	0.359	0.052 (0.165)	0.839	1.000
200	(1)	0.786	0.988	1.000	—	0.850	1.000
	(2)	0.078	0.126	0.260	—	0.118	0.318
	(3)	0.250	0.805	0.371	0.048 (0.174)	0.693	1.000
500	(1)	0.819	1.000	1.000	—	0.553	1.000
	(2)	0.082	0.149	0.250	—	0.126	0.341
	(3)	0.166	0.894	0.376	0.052 (0.176)	0.383	1.000
1,000	(1)	0.808	1.000	1.000	—	0.271	1.000
	(2)	0.095	0.140	0.272	—	0.136	0.347
	(3)	0.119	0.931	0.400	0.049 (0.173)	0.227	1.000

Table 3: *Proportion of rejection of the null hypothesis for the t-test with a 5% nominal value. All the simulated processes are nonstationary except the nonlinear 3-Regime SETAR (with a corridor regime being nonstationary). Column I(1)* allows to compare the results using the critical values of Table 2, and standard critical values (in parentheses).*

0.10 when the sample size is $1,000$. The same can be said when the underlying DGP is a nonlinear 3-R SETAR.[*] When the variables are generated as unit root processes with drift and a break or as nonstationary AR(2) processes, model (2) also reduces considerably the size distortions, although the rejection rate remains close to 0.30 in the former case, and around 0.14 in the latter.

Second, the inclusion of breaks greatly exacerbates the problem. This can be seen on the last column of Table (3), the I(1)+drift+break case, having the greater rejection proportions for either model. Recall that the reasoning behind Agiakloglou' proposal was that spurious regression come from an uncontrolled autocorrelation problem, so a word of caution must be made if structural breaks are suspected.

To summarize, our comment on Agiakloglou's procedure provides a threefold extension: First, it extends the finite sample experiment by allowing constant terms and different parameter values for each variable; the size-distortion reduction of Agiakloglou's proposal is preserved for autoregressive stationary data-generating processes. Second, it also provides the underlying asymptotic theory for the driftless unit root case. In short, limit expressions for the slope estimate and its associated t-ratio under stochastic nonstationarity were presented. The t_β for model (2) converges to a normal standard, whilst, for model (3), it converges to a nuisance-parameter-free function of Brownian motions (critical values for the latter case were calculated). Finally, it also limits its reach: practitioners have to be cautious, and aware of the limitations of Agiakloglou's spurious regression solution. When applied to a wider variety of data-generating processes other than simple AR(1) processes, the spurious regression phenomenon, in the form of greater rejection rate of the null than the nominal value of the test, although reduced, is still present.

[*]The size-distortion reductions also occur for different specifications of the 3-R SETAR, and for other persistence parameter values in ARFIMA models, as can be seen in Tables 14 and 21, respectively, in the supplementary material.

A Proof of Theorem

1.- **Model 2**: Given that x_t and y_t are unit-root processes and the regression is estimated using the variables in first differences, the results in part (1) of the Theorem are straightforward. First differences of both unit-root processes behave as simple iid I(0) independent processes. Therefore, classic econometric theory applies, so we know that, under the null, the t-ratios converge to a Standard Normal random variable (*rv*). Note that the standard Brownian motion, $w_{x,y}(\cdot)$, stated in the first part of the Theorem, is precisely a standard Normal *rv*. Note too that, when the differenced variables are not iid but rather stationary ARMA processes (for instance), typical autocorrelation ensues.

2.- **Model 3**: We present a guide as to how to obtain the asymptotic expression of the estimates and their associated t-ratios using OLS where the variables y_t and x_t are generated as independent unit root processes, and specification 3 is estimated. Let $\hat{\Theta} = (\hat{\alpha}, \hat{\beta}, \hat{\delta})'$:

$$\hat{\Theta} = \left(X'X\right)^{-1} X'Y, \quad Var(\hat{\Theta}) = \hat{\sigma}^2 \left(X'X\right)^{-1},$$

$$t_{\hat{\alpha}} = \frac{\hat{\alpha}}{\sqrt{\hat{\sigma}^2_{\hat{\alpha}}}}, \quad t_{\hat{\beta}} = \frac{\hat{\beta}}{\sqrt{\hat{\sigma}^2_{\hat{\beta}}}}, \quad t_{\hat{\delta}} = \frac{\hat{\delta}}{\sqrt{\hat{\sigma}^2_{\hat{\delta}}}},$$

where,

$$X'X = \begin{bmatrix} T & \sum x_t & \sum y_{t-1} \\ \sum x_t & \sum x_t^2 & \sum x_t y_{t-1} \\ \sum y_{t-1} & \sum x_t y_{t-1} & \sum y_{t-1}^2 \end{bmatrix}; \quad X'Y = \begin{bmatrix} \sum y_t \\ \sum x_t y_t \\ \sum y_t y_{t-1} \end{bmatrix};$$

and,

$$\begin{aligned} \hat{\sigma}^2 = \quad & T^{-1} \Big[\sum y_t^2 + \hat{\alpha}^2 T + \hat{\beta}^2 \sum x_t^2 + \hat{\delta}^2 \sum y_{t-1}^2 - 2\hat{\alpha} \sum y_t - 2\hat{\beta} \sum x_t y_t \\ & -2\hat{\delta} \sum y_t y_{t-1} + 2\hat{\alpha}\hat{\beta} \sum x_t + 2\hat{\alpha}\hat{\delta} \sum t + 2\hat{\beta}\hat{\delta} \sum x_t y_{t-1} \Big]. \end{aligned}$$

The estimated parameters $\hat{\alpha}$, $\hat{\beta}$, and $\hat{\delta}$ and their corresponding t-ratios are functions of the following expressions (initial conditions are set as $x_0 = y_0 = 0$). We use well-known asymptotic results (available, *inter alia*, in Hamilton, 1994, Proposition 17.1, p.486). Let $z = y, x$ and $\xi_{zt} = \sum_{i=1}^{t} u_{zi}$:

$$T^{-3/2} \sum \xi_{zt-1} \xrightarrow{D} \sigma_z \int_0^1 \omega_z(r) dr,$$

$$T^{-2} \sum \xi_{zt-1}^2 \xrightarrow{D} \sigma_z^2 \int_0^1 [\omega_z(r)]^2 dr,$$

$$T^{-1} \sum \xi_{zt-1} u_{zt} \xrightarrow{D} \sigma_z^2 \frac{1}{2} ([\omega_z(1)]^2 - 1),$$

$$T^{-2} \sum \xi_{xt-1} \xi_{yt-1} \xrightarrow{D} \sigma_x \sigma_y \int_0^1 \omega_x \omega y(r) dr,$$

$$T^{-1} \sum \xi_{yt-1} u_{xt} \xrightarrow{D} \sigma_x \sigma y \int_0^1 \omega_x(r) d\omega_y(r),$$

where $\omega_z(1)$ is a standard brownian motion and $\stackrel{D}{\rightarrow}$ denotes convergence in law. The later results allow us to fill the OLS matrices and then compute the parameter estimates and the t-statistic associated with each. Note that we also require sums involving the lagged dependent variable, which can be easily obtained:

$$
\begin{aligned}
\sum z_t &= \sum \xi_{zt-1} + \sum u_{zt}, \\
\sum y_{t-1} &= \sum \xi_{yt-1}, \\
\sum z_t^2 &= \sum \xi_{zt-1}^2 + \sum u_{zt}^2 + 2\sum \xi_{zt-1} u_{zt}, \\
\sum y_{t-1}^2 &= \sum \xi_{yt-1}^2, \\
\sum x_t y_t &= \sum \xi_{xt-1}\xi_{yt-1} + \sum \xi_{yt-1} u_{xt} + \sum \xi_{xt-1} u_{yt} + \sum u_{xt} u_{yt}, \\
\sum x_t y_{t-1} &= \sum \xi_{xt-1}\xi_{yt-1} + \sum \xi_{yt-1} u_{xt}, \\
\sum y_{t-1} y_t &= \sum \xi_{yt-1}^2 + \sum \xi_{yt-1} u_{yt}.
\end{aligned}
$$

Replacing the sums that appear in the OLS formulae with the above asymptotic expressions and letting $T \to \infty$ yields the results stated in the Theorem. The algebra involved is cumbersome but does not present any particular complication. Codes in algebraic manipulation programs with the calculations are available upon request.

B Data Generating Processes

ARFIMA(0,d,0)	$(1-L)^{0.9}x_t = u_{x,t} \quad u_{x,t} \sim iidN(0,1)$ $(1-L)^{0.8}y_t = u_{y,t} \quad u_{y,t} \sim iidN(0,1)$
AR(2)	$x_t = 2.00 + 0.45x_{t-1} + 0.55x_{t-2} + u_{x,t}, \quad u_{x,t} \sim iidN(0,5)$ $y_t = 0.80 + 0.50y_{t-1} + 0.50y_{t-2} + u_{y,t}, \quad u_{y,t} \sim iidN(0,5)$
ARIMA(1,1,0)	$(1-L)x_t = 5.00 + 0.70x_{t-1} + u_{x,t}, \quad u_{x,t} \sim iidN(0,5)$ $(1-L)y_t = 5.00 + 0.80y_{t-1} + u_{y,t}, \quad u_{y,t} \sim iidN(0,5)$
I(1)	$x_t = x_{t-1} + u_{x,t}, \quad u_{x,t} \sim iidN(0,5)$ $y_t = y_{t-1} + u_{y,t}, \quad u_{y,t} \sim iidN(0,5)$
3-R SETAR	$x_t = \begin{cases} 1 + 0.50x_{t-1} + u_{x1,t} & \text{if } x_{t-1} > 18 \\ 1 + 0.30x_{t-1} + 0.20x_{t-2} + u_{x2,t} & \text{if } 15 < x_{t-1} \leq 18 \\ 1 + 0.40x_{t-1} + u_{x3,t} & \text{if } x_{t-1} \leq 15 \end{cases}$ $u_{x1,t} \sim iidN(0,0.4),\ u_{x2,t} \sim iidN(0,\sqrt{0.6}),\ u_{x3,t} \sim iidN(0,\sqrt{0.4})$ $y_t = \begin{cases} 1 + 0.10y_{t-1} + u_{y1,t} & \text{if } y_{t-1} > 16 \\ 1 + 0.65y_{t-1} + 0.35y_{t-2} + u_{y2,t} & \text{if } 12 < y_{t-1} \leq 16 \\ 1 + 0.60y_{t-1} + u_{y3,t} & \text{if } y_{t-1} \leq 12 \end{cases}$ $u_{y,t,1} \sim iidN(0,5),\ u_{y,t,2} \sim iidN(0,\sqrt{2.2}),\ u_{y,t,3} \sim iidN(0,\sqrt{1.5})$
I(1)+drift+break	$x_t = 5.00 - 0.50t + 2.00DT_{xt} + \xi_{x,t}$ $\xi_{x,t} = \xi_{x,t-1} + u_{x,t}, \quad u_{x,t} \sim iidN(0,5^2)$ $y_t = 0.50 - 3.00t + 7.00DT_{yt} + \xi_{y,t}$ $\xi_{y,t} = \xi_{y,t-1} + u_{y,t}, \quad u_{y,t} \sim iidN(0,3^2)$ $DT_{xt} = DT_{yt} = (0,0,\ldots,0,1,2,\ldots,T/2)'$

Table 4: *Data-generating processes. L is the lag operator.*

References

Agiakloglou, C. (2013). Resolving spurious regressions and serially correlated errors. *Empirical Economics*, 45(3):1361–1366.

Granger, C. W. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of econometrics*, 2(2):111–120.

Hamilton, J. D. (1994). *Time series analysis*. Princeton university press Princeton.

Kolev, G. I. (2011). The 'spurious regression problem' the classical regression model framework. *Economics Bulletin*, 31(1):925–937.

Martínez-Rivera, B. and Ventosa-Santaulària, D. (2012). A comment on 'is the spurious regression problem spurious?'. *Economics Letters*, 115(2):229–231.

McCallum, B. T. (2010). Is the spurious regression problem spurious? *Economics Letters*, 107(3):321–323.

Phillips, P. C. (1986). Understanding spurious regressions in econometrics. *Journal of econometrics*, 33(3):311–340.

Ventosa-Santaulària, D. (2009). Spurious regression. *Journal of Probability and Statistics*, 2009.

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*, pages 1–63.