Is Spurious Regression an Issue for two independent stationary AR(1) processes?

Universidad Complutense de Madrid
November 15, 2017

Christos Agiakloglou and Charalampos Agiropoulos
Department of Economics
University of Piraeus

Spurious Regression

- The pioneer work of Granger and Newbold (1974) was just the introduction of a new phenomenon, known as spurious regression.
- Actually, this phenomenon was initially presented by Yule (1926) as a spurious correlation phenomenon.
- Granger and Newbold (1974) showed, using a Monte Carlo analysis, that the regression of two independent drift-free random walk processes will produce strong evidence of a linear relationship.
- Phillips (1986) proved mathematically this behavior showing that the usual t-statistic does not have a limiting distribution.
- Entorf (1997) consider random walk with drifts.
- Granger, Hyung and Jeon (2001) found spurious results even for two stationary independent AR(1) processes.

Serially Correlated Errors

- Spurious regression is linked to serially correlated errors.
- Granger and Newbold (1974) pointed out that along with the large t-values strong evidence of serially correlated errors will appear in regression analysis, stating that when a low value of the Durbin-Watson statistic is combined with a high value of the t-statistic the relationship is not true.
- Marmol (1995), who generalized the work of Phillips (1986) for high-order integrated processes, showed that the Durbin Watson statistic will converge in probability to zero and therefore low values of this statistic are expected to appear in the presence of spurious regressions.
- Newbold and Davies (1978) found autocorrelated errors for non-stationary moving average processes.
- Agiakloglou (2009) reported serially correlated errors for the first and second moment, indicating the presence ARCH(1) error structure, for stationary AR(1) processes.
- Agiakloglou, Tsimbos and Tsimpanos (2015) found evidence of spatially correlated errors for stationary spatial autoregressive SAR(1) processes.

Two Examples

- Harvey (1980) “Econometrics – Alchemy or science?” studied the relationship between rainfall and inflation rate in U.K..
- Ferson, Sarkissian and Simin (2003) “Spurious regression in financial Economics” found that many predictive stock return regressions in the literature, based on individual predicting variables, may be spurious.

Spurious Regression vs. Correlation

- Two similar if not identical terms referring to the same phenomenon of obtaining false evidence about the existence of a linear relationship between two variables.
- Spurious Regression:

  Large t-value along with low Durbin-Watson statistic.
- Spurious Correlation:

  No hint.
- Agiakloglou and Tsimpanos (2012) examined the spurious correlation phenomenon for two independent stationary AR(1) processes and they found no evidence of spurious behavior using the true variance of their sample correlation coefficient.
- Also, Agiakloglou and Agiropoulos (2016) examined the balance between size and power in testing for linear association between two independent stationary AR(1) processes.
- More difficult to detect spurious correlations.

The Issue

- Consider the following regression model:

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

- where the error term is normally distributed with mean zero and constant variance and the variables X_t and Y_t are generated by the following DGP:

$$Y_t = \varphi Y_{t-1} + \varepsilon_{yt} \qquad \text{and} \qquad X_t = \varphi X_{t-1} + \varepsilon_{xt}$$

- where the errors ε_{xt} and ε_{yt} are each white noise N(0, 1) processes independent of each other and the autoregressive parameters are allowed to take values of 0.0, 0.2, 0.5, 0.8 and 0.9.
- Note that if φ = 1, both processes are non-stationary random walk processes without drift, whereas if φ = 0, both processes are white noise processes.

Testing for a Linear Relationship

- The test of no linear relationship is based on the following hypotheses:

  H_0: β = 0 against H_1: β ≠ 0

  and it is implemented by the usual t statistic, i.e.,

$$t = \frac{\hat{\beta}}{se(\hat{\beta})}$$

  where the t statistic follows a *t* distribution with (*T* - 2).
- The null hypothesis will be accepted if its absolute value is less than the critical value.
- Note that t must have zero mean and variance one, asymptotically.
- Unfortunately, as it is known, false conclusions will be drawn for time series data obtained from independent non-stationary and/or stationary processes.

Two independent non-stationary processes

- Consider two independent random walk processes without drift.
- In this case, the null hypothesis is rejected too often, reaching to 100% for large sample sizes.
- The percentage of rejections increases as sample size increases.
- The t distribution does not convert to a standard normal distribution.
- The variance of the estimated coefficient is not properly defined.

Percentage of rejections of the null hypothesis that β = 0 at the nominal 5% level ($|t| > 1.96$) along with the standard deviation of the t statistic for two independent Random Walk without drift processes for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

	Sample Size			
	50	**100**	**500**	**1000**
% of Rejections	68	76	89	93
St. dev.	5.201	7.294	16.431	23.831

Two stationary Processes

- Consider two independent stationary AR(1) processes.
- The null hypothesis is also rejected very often.
- The percentage of rejections depends only on the magnitude of the autoregressive parameter.
- The percentage of rejections is not affected by the sample size.
- The t distribution does not convert to a standard normal distribution.
- The variance of the estimated coefficient is not properly defined.

Percentage of rejections of the null hypothesis that $\beta = 0$ at the nominal 5% level ($|t| > 1.96$) along with the standard deviation of the t statistic for two independent stationary AR(1) processes regardless of sample size, based on 10,000 replications

	φ				
	0.0	**0.2**	**0.5**	**0.8**	**0.9**
% of Rejections	5	6	13	35	52
St. dev.	1.003	1.046	1.288	2.138	3.049

Evidence of serially correlated errors

- Spurious Regression is related to autocorrelated errors.
- The Durbin-Watson (DW) statistic is affected by the magnitude of the autoregressive parameter and by the sample size.
- The value of the DW statistic <u>decreases</u> as the <u>value</u> of the autoregressive parameter <u>increases</u>.
- The value of the DW statistic <u>decreases</u> as <u>sample size increases</u> reaching to some predetermined value as Marmol (1995) has indicated for two independent non-stationary processes.
- The DW statistic converges to its population value, i.e.,

  $$DW = 2(1 - \rho).$$
- So, for $\rho = 1$, $DW = 0$, for $\rho = 0.9$, $DW = 0.2$ etc.

Mean values of the Durbin-Watson statistic for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

	φ					
Sample Size	**0.0**	**0.2**	**0.5**	**0.8**	**0.9**	**1.0**
50	2.0014	1.6479	1.1237	0.6203	0.4693	0.3332
100	1.9981	1.6222	1.0613	0.5081	0.3216	0.1735
500	2.0003	1.6049	1.0123	0.4265	0.2240	0.0360
1000	1.9995	1.6021	1.0061	0.4102	0.2119	0.0183
10000	2.0009	1.6010	1.0009	0.4010	0.2010	0.0018

Dealing with Autocorrelation

- Granger, Hyung and Jeon (2001) proposed a new method called BART that improved the behavior of this phenomenon only for very large sample sizes. However, for moderate and even for large sample sizes the spuriocity was not totally removed.
- For example, for large value of the autoregressive parameter equal to 0.9, Granger *et al.* (2001) have reported 23.9 empirical percentage level of rejecting the null hypothesis at the 5% nominal level, for sample size of 500 observations using the BART method.
- Agiakloglou (2013) showed significant improvement on the performance of the test as well as on the Durbin-Watson statistic for small and moderate sample sizes for two independent stationary AR(1) processes and for two non-stationary I(1) processes by estimating the original simple regression model either with a lagged dependent variable or in first differences respectively.
- Agiakloglou, Tsimbos and Tsimpanos (2015) showed that in spatial analysis at the global level the employment of the classical LM spatial dependence specification tests will suggest other forms of spatial model estimation that will not deliver spurious results and spatially autocorrelated errors.

Autocorrelated Errors

- Consider the same DGP where now the error term is generated by the following AR(1) process, i.e.,

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

- where the absolute value of the autoregressive coefficient ρ is less than one, i.e., $|\rho| < 1$, and the error term u_t is considered to be *iid* normally distributed with mean zero and constant variance σ_u^2, i.e., $u_t \sim iid\mathrm{N}(0, \sigma_u^2)$.
- Estimation of the original simple regression model with OLS, under the assumption of serially correlated errors, will produce the same estimates, which will still be unbiased, but with different variance.

Variance under Autocorrelated Errors

- The variance of the new OLS-AUTO estimator $\tilde{\beta}$ will be obtained as:

$$Var(\tilde{\beta}) = \frac{\sigma_\varepsilon^2}{\sum(X_t - \bar{X})^2}\left[1 + 2\rho\frac{\sum(X_t - \bar{X})(X_{t-1} - \bar{X})}{\sum(X_t - \bar{X})^2} + 2\rho^2\frac{\sum(X_t - \bar{X})(X_{t-2} - \bar{X})}{\sum(X_t - \bar{X})^2} + \cdots\right]$$

- or equivalently

$$Var(\tilde{\beta}) = Var(\hat{\beta})[1 + 2\rho r + 2\rho^2 r^2 + \cdots]$$

- where Var($\hat{\beta}$) is the OLS variance of $\hat{\beta}$ without the presence of serially correlated errors, *r* is the sample autocorrelation of the independent variable, knowing that X_t follows an AR(1) process, and ρ is the correlation coefficient of the error term. Hence,

$$Var(\tilde{\beta}) = Var(\hat{\beta})A$$

- where A > 0 such as $A = [1 + 2\rho r + 2\rho^2 r^2 + \cdots]$

The factor A

- The correct variance under AUTO is greater than the incorrect OLS variance times a factor A, i.e.,

$$Var(\tilde{\beta}) = Var(\hat{\beta})A$$

$$A = [1 + 2\rho r + 2\rho^2 r^2 + \cdots]$$

- The value of A can be obtained either as an approximation of the first two terms, i.e.,

$$A_1 = 1 + 2\rho r$$

- or as an infinite sum of a geometric process, i.e., as:

$$A_2 = 1 + 2\rho r(1 + \rho r + \rho^2 r^2 + \cdots) = 1 + 2\rho r \frac{1}{1 - \rho r} = \frac{1 + \rho r}{1 - \rho r}$$

- A defines the ratio of the two variances.

Comments on A

- For small values of ρ and r both approximations of A will take the same value, i.e., for $\rho = r = 0.2$, $A_1 = A_2 = 1.08$.
- For large values of ρ and r the two approximations of A will take different values, i.e., for $\rho = r = 0.8$, $A_1 = 2.28$ and $A_2 = 4.56$.
- Hence, if the presence of serially correlated errors is ignored and the incorrect variance, $Var(\hat{\beta})$, is used to calculate the relevant t statistic, instead of the correct variance, $Var(\tilde{\beta})$, the values of the t statistic will be larger, resulting to more rejections of the null hypothesis, since the incorrect variance is smaller than the correct variance.
- In this case, the variance of the estimator is underestimated, the magnitude of which depends on the approximation of A that is used.
- For small values of ρ and r the variance will be underestimated equally at a small level, i.e., for $\rho = r = 0.8$ is 7.4%, while for large values of ρ and r the underestimation will differ significantly. For example, for $\rho = r = 0.8$ the variance is underestimated 56% using the A_1 approximation and 78% (1/4.56=0.22) using the A_2 approximation.
- Thus, the use of the incorrect variance alters the distribution of the t statistic which does not convert to a standard normal distribution, producing misleading statistical results.

Testing the null under AUTO

- The objective is to correct the behavior of the t statistic by importing the correct variance so that the test will have better performance, knowing that the variance that was used was smaller than the correct variance.
- Keeping the same simulation process alive, the relevant t statistic, for testing the null hypothesis that $\beta = 0$, is now calculated by replacing the OLS incorrect variance with the correct variance, obtained under autocorrelated errors and the new two t statistics, using the two approximations of A, are obtained as follows:

$$t_1 = \frac{\hat{\beta}}{se(\hat{\beta})\sqrt{A_1}} \quad \text{and} \quad t_2 = \frac{\hat{\beta}}{se(\hat{\beta})\sqrt{A_2}}$$

- where the values of A_1 and A_2 are calculated individually in every trial using the sample correlation coefficient r of X_t and the estimate of ρ obtained by estimating the AR(1) for the errors using the residuals resulting from the OLS estimation of the original simple regression model.
- The whole simulation process is conducted in *R*.

Percentage of rejections of the null hypothesis that β = 0 at the nominal 5% level (|t| > 1.96) using the t_1 and t_2 statistics along with their standard deviations for two independent stationary AR(1) processes for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

			φ				
			0.0	**0.2**	**0.5**	**0.8**	**0.9**
% of rej.	**50**	t_1	5.8	5.9	7.9	18.1	28.4
		t_2	5.8	5.8	7.0	10.4	13.8
	100	t_1	5.6	6.0	7.5	17.7	30.0
		t_2	5.6	5.9	6.4	8.2	10.8
	500	t_1	5.1	5.4	6.9	17.5	31.4
		t_2	5.1	5.4	5.7	6.2	6.8
	1000	t_1	5.0	4.9	6.3	17.3	30.4
		t_2	5.0	4.8	5.0	5.2	5.9
st. dev.	**50**	t_1	1.027	1.038	1.115	1.504	1.907
		t_2	1.027	1.036	1.075	1.234	1.397
	100	t_1	1.025	1.031	1.093	1.465	1.940
		t_2	1.025	1.029	1.047	1.127	1.243
	500	t_1	1.007	1.013	1.073	1.441	1.944
		t_2	1.007	1.011	1.020	1.039	1.068
	1000	t_1	0.998	0.999	1.051	1.421	1.920
		t_2	0.998	0.998	0.998	1.015	1.032

Comments on the t_1 test results

- The performance of the test has significantly improved simply by correcting the variance of the estimator, regardless of the formula that is used for getting values of A.
- Using the A_1 approximation of A, the relevant t_1 statistic has the same behavior as that of the classical t statistic under spurious regression, meaning that the number of rejections of the null hypothesis is only affected by the magnitude of the autoregressive parameter and not by the sample size.
- The difference is that the null hypothesis is rejected less frequently than using the classical t statistic for every value of the autoregressive parameter and sample size.
- For example, using the classical t statistic the null hypothesis is rejected 52% and 35% for $\varphi = 0.9$ and 0.8 respectively at the 5% nominal level, whereas using the t_1 statistic the null hypothesis is rejected 30% and 17.5% respectively.
- Therefore, the problem of getting spurious results has decreased, but not removed, simply because the values that the A_1 approximation of A was taking <u>were not large enough</u> to significantly increase the variance of the estimator and thus the distribution of the t_1 statistic did not convert to a standard normal distribution.

Comments on the t_2 test results

- Using the A_2 approximation for A, the relevant t_2 statistic produced even better results, showing evidence of convergence to the right nominal levels.
- Thus, the issue of getting spurious results for two independent stationary AR(1) processes can totally be removed asymptotically, since for large sample sizes the empirical levels are very close to the nominal levels regardless of the value of the autoregressive parameter and the standard deviations of the t_2 statistic converge to 1.
- Actually, the performance of the t_2 statistic depends not only on the value of the autoregressive parameter but also on the sample size.
- For small and moderate sample sizes the percentage of rejections of the null hypothesis increases as the value of the autoregressive parameter increases.
- For large sample sizes though the number of rejections of the null hypothesis decreases for all values of the autoregressive parameter, reaching happily the nominal level.
- It seems, therefore, that in this case the value of A_2 was large enough to alter the magnitude of the variance and produce smaller number of rejections.

Test performance

- Clearly, the use of the correct variance, when serially correlated errors are observed, improves the test results, regardless of the approximation that is used, but it did not remove the spurious regression phenomenon for small sample sizes.
- The test behaves better using the A_2 approximation of A, instead of A_1, especially, for large sample sizes and large values of the autoregressive parameter.
- For small values of the autoregressive parameter both approximations work equally well.
- However, for small and moderate sample sizes and for large values of the autoregressive parameter the empirical levels are not close to the nominal ones using the t_2 statistic.
- A more thoroughly examination of the simulation process suggests that the source of this problem, along with the asymptotic behavior of the t_2 statistic, comes from the estimated values of ρ and r used to calculate the value of A_2, since these values are not close enough to their theoretical ones for small and moderate sample sizes.

Mean values of *r* and *ρ* for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

	T	φ				
		0.0	**0.2**	**0.5**	**0.8**	**0.9**
r	**50**	0.001	0.193	0.479	0.759	0.852
	100	0.002	0.198	0.490	0.780	0876
	500	-0.002	0.197	0.496	0.796	0.895
	1000	-0.001	0.198	0.498	0.797	0.897
ρ	**50**	-0.017	0.163	0.429	0.685	0.763
	100	-0.007	0.183	0.467	0.746	0.836
	500	-0.000	0.198	0.494	0.790	0.888
	1000	-0.003	0.196	0.495	0.794	0.893

The estimated values of *ρ* and *r*

- The estimated mean values of ρ and r, used to calculate the value of A_2, are smaller than the theoretical ones, but as sample size increases, these values become identical to the theoretical generated ones.
- For this see also Agiakloglou and Agiropoulos (2016).
- This finding suggests <u>that the value of</u> A_2, unlike the value of A_1, will strongly be affected by these estimated values.
- For example, for $\rho = r = 0.9$ ➔ $A_2 = 9.526$ and $A_1 = 2.62$.
- However, if these values of ρ and r are replaced by their mean estimated values obtained through the simulation process we have:
- A) For $T = 50$, $\rho = 0.763$ and $r = 0.852$ ➔ $A_2 = 4.716$ and $A_1 = 2.300$.
- B) For $T = 100$, $\rho = 0.836$ and $r = 0.876$ ➔ $A_2 = 6.472$ and $A_1 = 2.465$.
- C) For $T = 500$, $\rho = 0.888$ and $r = 0.895$ ➔ $A_2 = 8.745$ and $A_1 = 2.590$.
- D) For $T = 1000$, $\rho = 0.893$ and $r = 0.897$ ➔ $A_2 = 9.097$ and $A_1 = 2.604$.
- Hence, for small sample sizes, the value of A_2 is significantly smaller of what it should be.
- Therefore, it is underestimated.
- More concerns on the estimated values of ρ.

Two propositions

- To investigate further the performance of the test we tried using :
- I) The theoretical generated values of ρ and r.
- II) The unbiased estimate of the variance of the error term, when serially correlated errors are observed.

The generated values of *ρ* and *r*

- The correct variance is calculated using <u>constant values</u> for A_1 and A_2 throughout the simulation process, based on the generated values of *ρ* and *r,* to calculate respectively the new two relevant t_1' and t_2' statistics.
- The performance of test using the t_1' statistic, was not affected at all, simply because, as previously discussed, the value of A_1 is small enough to significantly change the magnitude of the variance.
- On the other hand, the empirical size of the test using the t_2' statistic, based on the A_2 approximation of A, has really produced very astonishing results.
- For all values of the autoregressive parameter and for all sample sizes the empirical levels of this test are very close to the nominal levels, indicating that the value of A_2 was large enough to change significantly the magnitude of the variance of the estimator and the relevant statistic did convert to a standard normal distribution.
- It will be almost impossible to visualize in practice a scenario like this, in which case the analyst will a priori know the true generating values of an observed series.

Percentage of rejections of the null hypothesis that β = 0 at the nominal 5% level (|t| > 1.96) using the t'_1 and t'_2 statistics along with their standard deviations for two independent stationary AR(1) processes for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

			φ				
			0.0	**0.2**	**0.5**	**0.8**	**0.9**
% of rej.	**50**	t'_1	5.7	5.6	6.8	15.1	25.2
		t'_2	5.7	5.6	5.6	4.8	3.3
	100	t'_1	5.6	5.8	6.7	16.3	28.3
		t'_2	5.6	5.8	5.4	5.1	4.4
	500	t'_1	5.1	5.3	6.8	17.2	31.1
		t'_2	5.1	5.3	5.4	5.4	5.2
	1000	t'_1	5.0	5.0	6.4	17.2	30.3
		t'_2	5.0	4.9	4.8	4.9	5.0
st. dev.	**50**	t'_1	1.026	1.026	1.070	1.398	1.751
		t'_2	1.026	1.024	1.015	0.989	0.919
	100	t'_1	1.024	1.024	1.070	1.414	1.864
		t'_2	1.024	1.023	1.015	1.000	0.977
	500	t'_1	1.007	1.011	1.068	1.431	1.929
		t'_2	1.007	1.010	1.013	1.012	1.012
	1000	t'_1	0.998	0.999	1.049	1.416	1.912
		t'_2	0.998	0.997	0.995	1.002	1.003

The unbiased estimate of the variance

- If $\rho = 0$, OLS uses: $s_O^2 = \frac{1}{T-2}\sum \hat{\varepsilon}_t^2$ since $E(\sum \hat{\varepsilon}_t^2) = (T-2)\sigma_\varepsilon^2$.
- If $\rho \neq 0$, then: $E(\sum \tilde{\varepsilon}_t^2) = [(T-1) - A]\sigma_\varepsilon^2$ and therefore the unbiased estimate of the variance will be obtained as:

$$s_A^2 = \frac{1}{[(T-1)-A]}\sum \hat{\varepsilon}_t^2$$

- Note that the values of the residuals will not be affected. Auto does not affect the estimates.
- Notice also that if $\rho = 0$, then $A = 1$ and both expressions are identical.
- Hence: $s_A^2 = K\, s_O^2$ where $K = \frac{T-2}{[(T-1)-A]}$
- K is a positive number greater than one defining the ration of the two estimated variances of the error term.
- If $\rho = 0$, then $K = 1$ and, therefore, both variances are identical.
- K will also take the value of one asymptotically, as sample size increases.

Comments on the value of *K*

- Hence: $s_A^2 = K\, s_O^2$ where $K = \frac{T-2}{[(T-1)-A]}$
- *K* is a positive number greater than one defining the ration of the two estimated variances of the error term.
- If $\rho = 0$, then $K = 1$ and, therefore, both variances are identical.
- Thus, if there is any significant contribution to the magnitude of the variance, using the unbiased estimate of the variance of the error term, that is expected only to happen for <u>small sample sizes</u> and <u>large values</u> of the autoregressive parameter.
- For example, for $\rho = r = 0.8$ and $T = 50$, K = 1.08, while for $T = 500$, K = 1.0072, using the A_2 approximation of A, declaring an underestimation of the variance at the level of 7.4% and 0.71% respectively.
- The value of *K* will be even smaller using the estimated values of ρ and r and/or using the A_1 approximation of A.
- Basically, significantly changes are not expected.

The new Variance

- The new variance will be obtained as: $Var(\tilde{\beta}_A) = Var(\hat{\beta})KA$
- And the test will be implemented based on the new two test statistics:

$$t_1'' = \frac{\widehat{\beta}}{se(\widehat{\beta})\sqrt{K_1 A_1}} \quad \text{and} \quad t_2'' = \frac{\widehat{\beta}}{se(\widehat{\beta})\sqrt{K_2 A_2}}$$

- using both approximations of *A*.
- As expected, the performance of the test has been improved, but by little and only for small sample size and for large value of the autoregressive parameter.
- For example, for $T = 50$ and for $\varphi = 0.9$ the percentage of rejections of the null hypothesis at the 5% nominal level is now 12.7% instead of 13.8% using the A_2 approximations of A.
- In general, the use of the unbiased estimator of the variance of the error term did not remove the spurious behavior for small sample sizes and for large values of the autoregressive parameter.

Percentage of rejections of the null hypothesis that β = 0 at the nominal 5% level (|t| > 1.96) using the t_1'' and t_2'' statistics along with their standard deviations for two independent stationary AR(1) processes for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

			φ				
			0.0	**0.2**	**0.5**	**0.8**	**0.9**
% of rej.	**50**	t_1''	5.8	5.9	7.8	17.6	27.8
		t_2''	5.7	5.8	6.9	9.9	12.7
	100	t_1''	5.6	6.0	7.5	17.5	29.7
		t_2''	5.6	5.9	6.3	7.8	10.0
	500	t_1''	5.1	5.4	6.8	17.4	31.3
		t_2''	5.1	5.4	5.7	6.1	6.6
	1000	t_1''	5.0	4.9	6.3	17.3	30.4
		t_2''	5.0	4.8	5.0	5.2	5.8
st. dev.	**50**	t_1''	1.027	1.038	1.110	1.488	1.883
		t_2''	1.027	1.036	1.070	1.209	1.351
	100	t_1''	1.025	1.030	1.090	1.456	1.926
		t_2''	1.025	1.029	1.044	1.112	1.212
	500	t_1''	1.007	1.013	1.072	1.439	1.941
		t_2''	1.007	1.011	1.019	1.035	1.060
	1000	t_1''	0.998	0.999	1.051	1.420	1.918
		t_2''	0.998	0.998	0.998	1.013	1.028

The Cochrane-Orcutt procedure

- The natural way to deal with autocorrelated errors in time series econometrics is to apply the Cochrane –Orcutt (CO) procedure, although:

✓ I) Granger and Newbold (1974) have stated that the CO procedure will fail to correct this problem.

✓ II) Granger *et al.* (2001) have also stated "Patch-work procedures, such as the CO correction, will be inefficient compared to using a wider specification".

- As shown, the problem of not getting the right size for the test, using the correct variance under serially correlated AR(1) errors, was he estimated values of ρ.
- One way of getting better estimated values of ρ Is to apply the iterative Cochrane-Orcutt procedure.

The Cochrane-Orcutt procedure Cont'd

- The CO procedure is based on the generalized difference equation:

$$Y_t - \rho Y_{t-1} = \alpha(1-\rho) + \beta(X_t - \rho X_{t-1}) + \varepsilon_t - \rho\varepsilon_{t-1}$$

- where $Y_0 = \sqrt{1-\rho^2}Y_1$ and $X_0 = \sqrt{1-\rho^2}X_1$.
- The variance of the estimated coefficient β will asymptotically equal to:

$$Var(\hat{\beta}_{CO}) = \frac{(1-\rho^2)\sigma_\varepsilon^2}{T\sigma_X^2[1+\rho^2-2\rho r]}$$

- And the relevant t statistic for testing the null hypothesis that β = 0 is calculated as:

$$t_{CO} = \frac{\widehat{\beta}_{CO}}{se(\widehat{\beta}_{CO})}$$

- The results are astonishing.
- The application of the CO procedure has removed the spurious regression phenomenon for two independent stationary AR(1) processes.

Percentage of rejections of the null hypothesis that β = 0 at the nominal 5% level ($|t| > 1.96$) using the t_{CO} statistic along with its standard deviation for two independent stationary AR(1) processes for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

		φ				
		0.0	**0.2**	**0.5**	**0.8**	**0.9**
% of rej.	**50**	6.8	6.8	7.0	7.6	7.9
	100	6.2	6.2	6.2	6.4	6.6
	500	5.3	5.3	5.3	5.3	5.3
	1000	5.0	5.0	5.0	5.0	5.0
st. dev.	**50**	1.069	1.070	1.076	1.100	1.133
	100	1.049	1.044	1.046	1.051	1.058
	500	1.012	1.012	1.012	1.011	1.011
	1000	0.999	0.999	0.999	1.000	1.000

Three Additional Comments

- The results obtained by CO procedure:

✓ The estimated values of ρ.

Indeed they are closer to the true values.

✓ Compared with GLS method.

Similar results, slightly better for small sample sizes.

✓ Applied to two independent random walk process without drift.

It works even for those processes.

Mean values of ρ using OLS and Cochrane-Orcutt for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

	T	φ				
		0.0	**0.2**	**0.5**	**0.8**	**0.9**
ρ_{OLS}	**50**	-0.017	0.163	0.429	0.685	0.763
	100	-0.007	0.183	0.467	0.746	0.836
	500	-0.000	0.198	0.494	0.790	0.888
	1000	-0.003	0.196	0.495	0.794	0.893
ρ_{CO}	**50**	-0.022	0.166	0.446	0.722	0.807
	100	-0.009	0.184	0.475	0.763	0.857
	500	-0.002	0.197	0.495	0.793	0.892
	1000	-0.001	0.199	0.498	0.797	0.896

Percentage of rejections of the null hypothesis that β = 0 at the nominal 5% level (|t| > 1.96) using the t_{GLS} statistic along with its standard deviation for two independent stationary AR(1) processes for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

		φ				
		0.0	**0.2**	**0.5**	**0.8**	**0.9**
% of rej.	**50**	6.8	6.9	6.8	7.4	7.4
	100	6.1	6.2	6.3	6.5	6.4
	500	5.2	5.3	5.3	5.3	5.3
	1000	5.0	5.0	4.9	5.0	5.0
st. dev.	**50**	1.069	1.069	1.072	1.084	1.093
	100	1.044	1.044	1.045	1.049	1.053
	500	1.012	1.012	1.012	1.012	1.012
	1000	0.999	0.999	0.999	1.000	1.000

Percentage of rejections of the null hypothesis that β = 0 at the nominal 5% level (|t| > 1.96) using the t_{CO} statistic along with its standard deviation for two independent random walk processes without drift for sample sizes of 50, 100, 500 and 1000 observations, based on 10,000 replications

	Sample Size			
	50	**100**	**500**	**1000**
% of Rejections	9.5	7.5	5.4	5.1
st. dev.	1.268	1.112	1.019	1.004

General Comments

- It is very difficult to analyze time series data.
- Spuriocity is an issue.
- The Cochrane-Orcutt procedure works for two independent stationary and non-stationary processes.
- GLS also works.
- Agiakloglou (2013) showed that regressing the original simple regression model in first difference for two random walk without drift processes one will get 5% empirical levels for sample sizes of 50 and 100 observations.