University of Victoria
Econometrics Working Paper EWP 0203 **ISSN 1485-6441**

On the Futility of Testing the Error Term Assumptions in Spurious Regressions

David E. A. Giles

Department of Economics, University of Victoria, B.C., Canada

May, 2002

Abstract

A spurious regression model is one in which the dependent and independent variables are non-stationary, but not cointegrated, and the data are *not* filtered (*e.g.*, by differencing) before the model is estimated. It is well known that in this case the asymptotic behaviour of the least squares parameter estimates, their "t-ratios", the Durbin-Watson statistic and the R^2, are all non-standard. In particular, the parameter estimates and R^2 converge weakly to functionals of standard Brownian motions; the "t-ratios" diverge in distribution; and the Durbin-Watson statistic converges in probability to zero. In this paper we show that similar results apply to other common tests of a spurious regression model's specification. In particular, standard tests of the Normality and homoskedasticity of the error term are doomed to always reject the null hypotheses, asymptotically. These results further reinforce the need to avoid the estimation of spurious regressions.

Key Words: Spurious regression; normality; homoskedasticity; asymptotic theory; unit roots

JEL Classifications: C12; C22; C52;

Author Contact: David Giles, Department of Economics, University of Victoria,
P.O. Box 1700 STN CSC, Victoria, B.C., Canada, V8W 2Y2
e-mail: dgiles@uvic.ca; FAX: (250) 721 6214; Phone: (250) 721 8540

1. Introduction

Testing and allowing for non-stationary time-series data has been one of the major themes in econometrics over the past quarter-century or so. In their influential and relatively early contribution, Granger and Newbold (1974) drew our attention to some of the likely consequences of estimating a "spurious regression" model. They argued that the "levels" of many economic time-series are integrated or nearly so, and that if such data are used in a regression model a high R^2 value is likely to be found even when the series are independent of each other. Moreover, they illustrated that the regression residuals are likely to be autocorrelated, as evidenced by a very low value for the Durbin-Watson (DW) statistic. Students of econometrics soon, rather simplistically, equated a "spurious regression" with one in which $R^2 > DW$. Granger and Newbold (1977) and Plosser and Schwert (1978) added to our awareness and understanding of spurious regressions, but it was Phillips (1986) who provided a formal analytical explanation for the behaviour of the Ordinary Least Squares (OLS) coefficient estimator, the associated "t-statistics" and "F-statistic", and the R^2 and DW statistics in the context of such models.

Phillips (1986) developed a sophisticated asymptotic theory that he used to prove that in a spurious regression, *inter alia*, the DW statistic converges in probability to zero, the OLS parameter estimators and R^2 converge weakly to non-standard limiting distributions, and the "t-ratios" and "F-statistic" diverge in distribution as $T \uparrow \infty$. In short, Phillips "solved" the spurious regression problem, and in the process he proved that the unfortunate consequences of modelling with integrated data cannot be eliminated by increasing the sample size. This paper uses Phillips' non-standard asymptotic theory to demonstrate that the pitfalls of estimating a spurious regression extend to the application of standard diagnostic tests for the normality or homoskedasticity of the model's error term. We prove that the associated test statistics diverge in distribution as the sample size grows, so that one is led inevitably to the false conclusion that there is a "problem" with the usual assumptions about the error term. In fact, the real "problem" is a failure to take account of the non-stationarity of the data when specifying the model.

The next section establishes some of the basic asymptotic results that we use in the later analysis. Section 3 establishes and illustrates the asymptotic behaviour of the Jarque and Bera (1980) normality test; and some simple variants of the homoskedasticity tests proposed by Breusch and Pagan (1980) and Godfrey (1988) are examined in a similar way in section 4. Some concluding remarks are given in section 5.

2. Some Basic Asymptotic Results

For our purposes, it is sufficient to consider the simple univariate regression model, estimated by Ordinary Least Squares (OLS):

$$\hat{y}_t = \hat{a} + \hat{b}x_t + \hat{u}_t \quad . \tag{1}$$

The regression is "spurious" because both the dependent variable and the regressor follow independent I(1) processes:

$$y_t = y_{t-1} + v_t \ ; \quad v_t \sim iid(0, s_v^2) \tag{2}$$

$$x_t = x_{t-1} + w_t \ ; \quad w_t \sim iid(0, s_w^2) \tag{3}$$

with v_t and w_t independent for all t, and (without loss of generality) $v_0 = w_0 = 0$. In fact [Phillips (1986, p.313)] v_t and w_t may be heterogeneous, a point that is relevant in section 4 below. So, the true values of the parameters are $a = b = 0$.

From Phillips (1986, pp.315 and 326) we know that, by the strong law of McLeish (1975, Theorem 2.10) for weakly dependent sequences, and the Functional Central Limit Theorem [*e.g.*, Hamilton (1994, pp. 479-480)]:

$$T^{-3/2}\sum_t x_t \Rightarrow s_w \int_0^1 W(r)dr = s_w x_1, \quad \text{say} \tag{4}$$

and

$$T^{-3/2}\sum_t y_t \Rightarrow s_v \int_0^1 V(r)dr = s_v h_1, \quad \text{say} \tag{5}$$

where $\Rightarrow$ denotes weak convergence of the associated probability measures as $T \uparrow \infty$, and $W(r)$ and $V(r)$ are independent Wiener processes on C[0,1], the space of all real-valued functions on [0,1]. Using the same approach as Phillips it is also readily shown that

$$T^{-(k+2)/2}\sum_t x_t^k \Rightarrow s_w^k \int_0^1 \left(W(r)\right)^k dr = s_w^k x_k \ , \quad \text{say;} \quad k = 1, 2, 3, 4, ... \tag{6}$$

and

$$T^{-(k+2)/2}\sum_t y_t^k \Rightarrow s_v^k \int_0^1 \left(V(r)\right)^k dr = s_v^k h_k \ , \quad \text{say ;} \quad k = 1, 2, 3, 4, ... \tag{7}$$

From Phillips (1986, p.315) we also know that

$$T^{-2}\sum_t (x_t - \bar{x})^2 \Rightarrow s_w^2\left(x_2 - x_1^2\right) \tag{8}$$

$$T^{-2}\sum_t (y_t - \bar{y})^2 \Rightarrow s_v^2\left(h_2 - h_1^2\right) \tag{9}$$

and

$$T^{-2}\sum_t x_t y_t \Rightarrow s_w s_v y_{11}, \tag{10}$$

where

$$y_{ij} = \int_0^1 \left(W(r)\right)^i \left(V(r)\right)^j dr \ ; \quad \text{i, j} = 1, 2, 3, 4, ... \tag{11}$$

3. Asymptotic Behaviour of the Jarque-Bera Test

The test of Jarque and Bera (1980) is perhaps the one that is used most widely by econometricians to test for the Normality of OLS regression errors, and the test statistic is

$$JB = (T/6)\left[m_3 + (m_4 - 3)^2 / 4\right] \tag{12}$$

where, in general,

$$m_3 = \left[T^{-1} \sum_t (\hat{u}_t - \overline{u})^3 / s^3 \right]^2 \tag{13}$$

$$m_4 = T^{-1} \sum_t (\hat{u}_t - \overline{u})^4 \; / s^4 \tag{14}$$

and

$$s^2 = T^{-1} \sum_t (\hat{u}_t - \overline{u})^2 . \tag{15}$$

If the model includes an intercept, then of course $\overline{u} = 0$, and for a regression model with stationary data, the limiting null distribution of JB is c_2^2. However, in the case of a spurious regression the situation is fundamentally different.

Theorem 1

When applied to the spurious regression model (1), $(T^{-1} JB)$ *converges weakly as* $T \uparrow \infty$ *, and so* JB *itself diverges at the rate "T".*

Proof

From Phillips (1986, pp. 330-331):

$$\hat{b} \Rightarrow (s_v / s_w)\left[(y_{11} - x_1 h_1) / (x_2 - x_1^2) \right] = (s_v / s_w) q \quad , \quad \text{say} \tag{16}$$

and

$$T^{-1} s^2 \Rightarrow s_v^2 \left[h_2 - h_1^2 - q^2 (x_2 - x_1^2) \right] \quad . \tag{17}$$

So, by the Continuous Mapping Theorem [*e.g.*, Billingsley (1968, pp. 30-31)],

$$\hat{b}^k \Rightarrow (s_v / s_w)^k q^k \quad ; \qquad k = 1, 2, 3, ... \tag{18}$$

and

$$T^{-k} s^{2k} \Rightarrow s_v^{2k} \left[h_2 - h_1^2 - q^2 (x_2 - x_1^2) \right]^k ; \qquad k = 1, 2, 3, ... \tag{19}$$

First, consider m_3 in (13). Defining $y_t^* = (y_t - \bar{y})$ and $x_t^* = (x_t - \bar{x})$, note that

$$\sum_t \hat{u}_t^3 = \sum_t y_t^{*3} - 3\hat{b}\sum_t y_t^{*2} x_t^* + 3\hat{b}^2 \sum_t y_t^* x_t^{*2} - \hat{b}^3 \sum_t x_t^{*3} . \qquad (20)$$

From results A.1 - A.4 in the Appendix,

$$T^{-5/2}\sum_t x_t^{*3} \Rightarrow s_w^3\left[x_3 - 3x_1 x_2 + 2x_1^3\right] \qquad (21)$$

$$T^{-5/2}\sum_t y_t^{*3} \Rightarrow s_v^3\left[h_3 - 3h_1 h_2 + 2h_1^3\right] \qquad (22)$$

$$T^{-5/2}\sum_t y_t^{*2} x_t^* \Rightarrow (s_v^2 s_w)\left[y_{12} - 2h_1 y_{11} - x_1(h_2 + 2h_1^2)\right] \qquad (23)$$

$$T^{-5/2}\sum_t x_t^{*2} y_t^* \Rightarrow (s_w^2 s_v)\left[y_{21} - 2x_1 y_{11} - h_1(x_2 + 2x_1^2)\right] \quad . \qquad (24)$$

So, applying the Continuous Mapping Theorem to (20), and using results (18) and (21) - (24), the quantity $T^{-5/2}\sum_t \hat{u}_t^3$ converges weakly as $T \uparrow \infty$. Finally, using this result and (19) (with $k = 3$), and applying the Continuous Mapping Theorem to the terms in (13), we see that m_3 converges weakly with increasing "T".

Second, consider m_4 in (14), and note that

$$\sum_t \hat{u}_t^4 = \sum_t y_t^{*4} + 6\hat{b}^2\sum_t y_t^{*2} x_t^{*2} - 4\hat{b} \sum_t y_t^{*3} x_t^* - 4\hat{b}^3 \sum_t y_t^* x_t^{*3} + \hat{b}^4 \sum_t x_t^{*4} . \qquad (25)$$

From results A.5 - A.9 in the Appendix,

$$T^{-3}\sum_t y_t^{*4} \Rightarrow s_v^4\left[h_4 - 4h_1 h_3 + 6h_1^2 h_2 - 3h_1^4\right] \qquad (26)$$

$$T^{-3}\sum_t x_t^{*4} \Rightarrow s_w^4\left[x_4 - 4x_1x_3 + 6x_1^2x_2 - 3x_1^4\right] \tag{27}$$

$$T^{-3}\sum_t y_t^{*2}x_t^{*2} \Rightarrow \left(s_v^2 s_w^2\right)\left[y_{22} - 2x_1y_{12} + x_1^2h_1 + h_1^2x_2 - 2h_1y_{21} + 4x_1h_1y_{11} - 3h_1^2x_1^2\right] \tag{28}$$

$$T^{-3}\sum_t y_t^{*3}x_t^{*} \Rightarrow (s_v^3 s_w)\left[y_{13} - 3h_1y_{12} - x_1h_3 + 3h_1^2y_{11} - 3x_1h_1^3\right] \tag{29}$$

$$T^{-3}\sum_t x_t^{*3}y_t^{*} \Rightarrow (s_v s_w^3)\left[y_{31} - 3x_1y_{21} - h_1x_3 + 3x_1^2y_{11} - 3h_1x_1^3\right] \tag{30}$$

Again, applying the Continuous Mapping Theorem to (25), and using results (18) and (26) - (30), the quantity $T^{-3}\sum_t \hat{u}_t^4$ converges weakly as $T \uparrow \infty$. Finally, using this result and (19) (with $k = 4$), and applying the Continuous Mapping Theorem to the terms in (13), we see that m_4 converges weakly with increasing "T". Finally, it follows immediately from (12) that $(T^{-1}JB)$ converges weakly, so JB diverges at the rate "T" as $T \uparrow \infty$. (

The implication of this result is analogous to that associated with Phillips' (1986, pp. 333-334) result that $(T \times DW)$ converges weakly in the case of a spurious regression, and hence DW itself has a zero probability limit as $T \uparrow \infty$. That is to say, testing for serial independence or for Normality in the errors of a spurious regression will always lead to a rejection of the associated null hypotheses, for large enough T , whether these hypotheses are false or true. The application of these diagnostic tests is pointless when the data are non-stationary (and not cointegrated). It should also be noted that these results are quite independent of the initial values and distributions of v_t and w_t in (2) and (3). In particular, these random errors need not be Normally distributed. Of course, the latter point is of particular interest in the case of the Jarque-Bera test. Table 1 presents some Monte Carlo evidence to illustrate this point, and the results there also demonstrate the rate of divergence of the JB statistic as $T \uparrow \infty$. The Monte Carlo experiment involved 5,000 replications with the values of v_t and w_t generated as Standard Normal, Uniform(0,1), the inverse of Standard Normal, and log-Normal independent random variables. The simulations were conducted using the SHAZAM (2001) econometrics package.

4. Asymptotic Behaviour of the Breusch-Pagan-Godfrey Test for Homoskedasticity

As is well known, many of the familiar tests for the homoskedasticity of regression errors can be formulated as Lagrange Multiplier (LM) tests. For example, see Breusch and Pagan (1980) and Godfrey (1988). One simple example of the Breusch-Pagan-Godfrey (BPG) test involves an alternative hypothesis in which the regression error's variance is proportional to a linear combination of the regressors. For the simple regression model, the implementation of the test involves obtaining the OLS residuals, $\hat{u}_t$, from (1), and then fitting the following auxiliary regression:

$$\hat{u}_t^2 = a + bx_t + e_t \quad . \tag{31}$$

Let $\hat{a}$ and $\hat{b}$ be the OLS estimators of a and b, and as before let $s^2 = T^{-1}\sum_t \hat{u}_t^2$. Then the coefficient of determination associated with the estimation of (31) can be expressed as:

$$R^2 = \left[\hat{b}^2 \sum_t x_t^{*2}\right] / \left[\sum_t (\hat{u}_t - s^2)^2\right] \tag{32}$$

and an LM test of the homoskedasticity of the errors in (1) can be constructed using the statistic (TR^2) . For this model, if the variables in (1) were stationary then the test statistic would converge in distribution to c_1^2 if the null hypothesis were true. As is discussed by Godfrey (1988, Chap. 4) and Greene (2000, pp.509-510), an asymptotically equivalent LM test can be based on the statistic $(SSR/2)$, where "SSR" denotes the "regression" ("explained") sum of the squares from OLS estimation of the model

$$(\hat{u}_t^2 / s^2) = a' + b'x_t + e'_t \ . \tag{33}$$

In the case of a spurious regression, these two test statistics no longer converge in distribution to c_1^2 under the null of homoskedasticity. As was the case for the JB test for normality of the errors, the statistics for both the (TR^2) and $(SSR/2)$ variants of the LM test diverge in distribution as $T \uparrow \infty$, as we now show.

Theorem 2

When applied to the spurious regression model (1), R^2 defined in (32) converges weakly as $T \uparrow \infty$, and so (TR^2) diverges at the rate "T".

Proof

We can write (32) as

$$R^2 = \left[\left(T^{-1/2}\hat{b}\right)^2\left(T^{-2}\sum_t x_t^{*2}\right)\right] / \left[T^{-3}\sum_t (\hat{u}_t - s^2)^2\right]. \tag{34}$$

First, note that

$$T^{-1/2}\hat{b} = \left[T^{-5/2}\left(\sum_t x_t^* \hat{u}_t^2 - s^2 \sum_t x_t^*\right)\right] / \left[T^{-2}\sum_t x_t^{*2}\right] \tag{35}$$

$$= \left[T^{-5/2}\sum_t x_t^* \hat{u}_t^2\right] / \left[T^{-2}\sum_t x_t^{*2}\right], \tag{37}$$

so

$$R^2 = \left[\left(T^{-5/2}\sum_t x_t^* \hat{u}_t^2\right)^2 / \left(T^{-2}\sum_t x_t^{*2}\right)\right] / \left[T^{-3}\sum_t (\hat{u}_t - s^2)^2\right]. \tag{38}$$

Now, note that

$$T^{-3}\sum_t (\hat{u}_t^2 - s^2)^2 = T^{-3}\sum_t \hat{u}_t^4 + T^{-2}s^4 - 2T^{-1}s^2(T^{-2}\sum_t \hat{u}_t^2) \tag{39}$$

and so using (17) - (19), and (25) - (30) above, the expression in (39) converges weakly as $T \uparrow \infty$, by the Continuous Mapping Theorem.

Further, we can write

$$T^{-5/2}\sum_t x_t^* \hat{u}_t^2 = T^{-5/2}\left[\sum_t x_t^* y_t^{*2} + \hat{b}^2 \sum_t x_t^{*3} - 2\hat{b}\sum_t x_t^{*2} y_t^*\right] , \qquad (40)$$

and by using (18), (21), (23) and (24), the expression in (40) also converges weakly as $T \uparrow \infty$. Finally, using (8), (39) and (40), the Continuous Mapping Theorem ensures the weak convergence of R^2 in (38). Accordingly, (TR^2) diverges at the rate "T" as $T \uparrow \infty$. (

Theorem 3

When applied to the spurious regression model (1),the statistic $(T^{-1}SSR)$ *converges weakly as* $T \uparrow \infty$ *, and so the LM test statistic* $(SSR/2)$ *diverges at the rate "T".*

Proof

We can write

$$SSR = (\hat{b}')^2 \sum_t x_t^{*2} \quad , \qquad (41)$$

where $\hat{b}'$ is the OLS estimator of b' in (33). Noting that the sample mean of the dependent variable in (33) is unity, we have

$$\hat{b}' = \left[\sum_t x_t^* (\hat{u}_t^2 / s^2)\right] / \left[\sum_t x_t^{*2}\right] \qquad (42)$$

and so

$$(\hat{b}')^2 = T^{-1}\left[T^{-5/2}\sum_t x_t^* \hat{u}_t^2\right]^2 / \left[(T^{-1}s^2)\sum_t T^{-2}x_t^{*2}\right]^2 \tag{43}$$

and

$$T^{-1}SSR = \left(T^{-2}\sum_t x_t^{*2}\right)\left[T^{-5/2}\sum_t x_t^* \hat{u}_t^2\right]^2 / \left[(T^{-1}s^2)\sum_t T^{-2}x_t^{*2}\right]^2 \tag{44}$$

so, using results (8), (17) and (40), it follows by the Continuous Mapping Theorem that $(SSR/2)$ itself diverges at the rate "T" as $T \uparrow \infty$. (

From Theorems 2 and 3, we see that however it is formulated, there is no point in using this simple variant of the BPG test for homoskedasticity in the context of a spurious regression. Recall from section 2 that v_t and w_t need not be homoskedastic for our various asymptotic results to hold. So, regardless of whether the null hypothesis under test here is true or false, it will be rejected with increasing probability as the sample size grows. This is illustrated in Tables 2 and 3 for the case where the null hypothesis is true, and a nominal 10% significance level (based on the asymptotic c_1^2 distribution that would apply for stationary data) is used. The experimental design is the same as in section 3 above. The rates of divergence of the two test statistics, and the commensurate size distortions for the LM tests, can be seen in these tables for various distributions for the errors. Except for the $(SSR/2)$ version of the test, with T = 10 and either normal or uniform errors, there is positive size distortion. As the sample size grows, applying the BPG test in the context of a spurious regression leads one to increasingly come to the wrong conclusion that the errors are heteroskedastic. Although this point has been illustrated here with a very simple alternative hypothesis (namely that the variance of the regression errors is proportional to the sole regressor), it is clear that the same basic result also applies to more general variants of the BPG test in which the error variance is proportional to some linear combination of variables under the alternative hypothesis. These results also apply to either the $(SSR/2)$ or

(TR^2) versions of White's (1980) test for homoskedasticity against an arbitrary heteroskedastic alternative, and to other similar tests.

5. Conclusions

Many of the basic pitfalls associated with the use of non-stationary data in regression analysis have been well documented. In particular, Phillips (1986) exposed the underlying reasons for several observed empirical features of "spurious regressions". Among other things, he showed that the standard "t-test" and "F-test" statistics diverge as $T \uparrow \infty$, and the Durbin-Watson statistic converges to zero in probability. Thus, each of the associated null hypotheses will be rejected with increasing probability as the sample size grows, even though in fact they are actually true. This paper follows this theme and extends these results by considering what will be encountered by an applied researcher who (wisely) undertakes some other common types of regression diagnostic testing, but (unwisely) does so in the context of a spurious regression model.

We have shown that just as it is pointless to test for the independence of the model's errors via the Durbin-Watson test, it is equally futile to test for the normality or homoskedasticity of these errors in the spurious regression context. As the sample size grows, these standard diagnostic tests will increasingly reject these hypotheses, even when they are true. To then conclude that the model needs to be reformulated in order to deal with discovered "problems" associated with the error term would be as spurious as the estimation of the original model itself. Although our formal proofs are set in the context of a simple regression model, it is clear from Phillips (1986, pp.319-322) that they extend directly to the multiple regression model.

All of this underscores the importance of testing appropriately for unit roots (and cointegration) prior to the formulation and estimation of a regression model based on time series data, a point that was made very clearly by Granger and Newbold (1974, p.117):

> "*In our opinion the econometrician can no longer ignore the time series properties of the variables with which he is concerned - except at his peril. The fact that many economic 'levels' are near random walks or integrated processes means that considerable care has to be taken*

in specifying one's equations."

Acknowledgment

I am most grateful to Lauren Dong for her helpful comments on an earlier version of this paper.

Table 1

Mean Values and Rejection Rates for JB Test Statistic

(Monte Carlo Simulation, 5000 Repetitions, Nominal 10% Significance Level)[1]

	N(0,1)		U(0,1)		[1/N(0,1)]		exp[N(0,1)]	
T	**RR (%)**	**Mean**	**RR (%)**	**Mean (%)**	**RR**	**Mean (%)**	**RR**	**Mean**
10	1.34	0.91	0.54	0.78	10.02	2.03	4.24	1.28
20	3.36	1.45	1.34	1.08	23.06	8.22	11.26	2.90
50	13.84	2.97	8.06	2.30	55.64	49.38	26.94	7.09
100	45.78	5.95	35.10	4.48	81.08	158.42	55.28	11.37
250	79.68	14.29	67.84	8.77	93.54	1036.56	77.72	20.90
500	91.14	29.57	89.36	22.25	97.20	6594.69	91.28	31.61
1000	95.90	57.94	95.34	45.02	98.74	19997.91	95.86	53.83
2000	98.08	121.83	97.96	90.52	99.24	92016.16	97.74	100.63
5000	99.30	290.83	99.18	231.64	99.76	39813.92	99.04	233.01
10000	99.50	600.22	99.50	451.83	99.82	60762.71	99.64	479.90

1. The 10% critical value for the Chi Square distribution with 2 degrees of freedom is 4.60517. "RR" denotes the "rejection rate" - *i.e.*, the percentage of the 5,000 simulated values for JB that exceeded this critical value.

Table 2

Mean Values and Rejection Rates for BPG Test Statistic:

SSR Version

(Monte Carlo Simulation, 5000 Repetitions, Nominal 10% Significance Level)[1]

	N(0,1)		U(0,1)		[1/N(0,1)]		exp[N(0,1)]	
T	RR (%)	Mean	RR (%)	Mean (%)	RR	Mean (%)	RR	Mean
10	4.88	0.73	6.14	0.79	15.50	1.40	15.08	1.27
20	12.16	1.13	14.40	1.22	34.00	2.91	27.84	2.23
50	30.32	2.38	34.88	2.87	56.28	7.53	45.56	4.74
100	45.74	4.73	50.34	5.71	68.24	14.61	57.68	8.35
250	64.92	11.74	67.60	14.38	79.78	39.61	70.60	17.96
500	74.72	23.84	75.94	28.30	86.68	78.11	77.34	32.34
1000	82.76	48.17	82.90	56.61	90.84	153.11	84.42	63.12
2000	86.46	94.37	88.16	114.92	92.80	303.22	89.56	122.28
5000	92.12	228.55	91.56	284.06	95.36	774.02	92.50	296.50
10000	93.98	460.56	94.78	570.44	97.10	1605.40	94.02	556.36

1. The 10% critical value for the Chi Square distribution with 1 degrees of freedom is 2.70554. "RR" denotes the "rejection rate" - *i.e.*, the percentage of the 5,000 simulated values for BPG that exceeded this critical value.

Table 3

Mean Values and Rejection Rates for BPG Test Statistic:

TR^2 Version

(Monte Carlo Simulation, 5000 Repetitions, Nominal 10% Significance Level)[1]

	N(0,1)		U(0,1)		[1/N(0,1)]		exp[N(0,1)]	
T	**RR (%)**	**Mean**	**RR (%)**	**Mean (%)**	**RR**	**Mean (%)**	**RR**	**Mean**
10	12.68	1.14	14.30	1.20	25.32	1.86	25.36	1.74
20	19.54	1.55	21.58	1.59	39.58	3.17	35.54	2.60
50	38.88	3.24	42.86	3.56	58.76	7.52	52.28	5.30
100	53.14	6.26	57.94	7.00	70.80	14.98	63.82	9.49
250	69.82	15.44	72.94	17.57	80.80	70.70	74.96	20.76
500	78.00	31.54	79.94	34.37	87.54	75.01	81.06	38.27
1000	84.84	64.62	85.34	70.01	90.80	149.41	86.84	75.54
2000	88.54	123.67	90.14	140.97	93.22	298.48	91.08	147.95
5000	93.34	304.76	93.06	339.70	95.80	736.72	93.50	358.96
10000	94.78	614.62	95.86	698.52	97.44	1575.00	94.94	675.14

1. The 10% critical value for the Chi Square distribution with 1 degrees of freedom is 2.70554. "RR" denotes the "rejection rate" - *i.e.*, the percentage of the 5,000 simulated values for BPG that exceeded this critical value.

Appendix

Results Used for the Proof of Theorem 1

A.1: Derivation of Equation (21)

$$T^{-5/2}\sum_t x_t^{*3} = T^{-5/2}\sum_t \left(x_t^3 - 3x_t^2\bar{x} + 3x_t\bar{x}^2 - \bar{x}^3\right)$$

$$= T^{-5/2}\sum_t x_t^3 - 3T^{-5/2}(T^{-1}\sum_t x_t)(\sum_t x_t^2) + 3(T^{-1}\sum_t x_t)^2(T^{-5/2}\sum_t x_t) - T^{-3/2}(T^{-1}\sum_t x_t)^3$$

$$= (T^{-5/2}\sum_t x_t^3) - 3(T^{-3/2}\sum_t x_t)(T^{-2}\sum_t x_t^2) + 3(T^{-3/2}\sum_t x_t)^2(T^{-3/2}\sum_t x_t) - (T^{-3/2}\sum_t x_t)^3$$

Then, repeatedly applying result (6) in section 2 (with k = 1, 2, 3), and appealing to the Continuous Mapping Theorem, we obtain:

$$T^{-5/2}\sum_t x_t^{*3} \Rightarrow s_w^3\left[x_3 - 3x_1x_2 + 2x_1^3\right] \quad .$$

A.2: Derivation of Equation (22)

This follows as for equation (21) above. Then, repeated application of result (7) in section 2 yields:

$$T^{-5/2}\sum_t y_t^{*3} \Rightarrow s_v^3\left[h_3 - 3h_1h_2 + 2h_1^3\right] \quad .$$

A.3: Derivation of Equation (23)

$$T^{-5/2}\sum_t y_t^{*2}x_t^* = T^{-5/2}\left(\sum_t x_t y_t^2 - 2\bar{y}\sum_t x_t y_t + \bar{y}^2\sum_t x_t - \bar{x}\sum_t y_t^2 + 2\overline{xy}\sum_t y_t - \sum_t \overline{xy}^2\right)$$

$$= (T^{-5/2}\sum_t x_t y_t^2) - 2(T^{-3/2}\sum_t y_t)(T^{-2}\sum_t x_t y_t) - (T^{-3/2}\sum_t x_t)(T^{-2}\sum_t y_t^2)$$

$$+ 2(T^{-3/2}\sum_t x_t)(T^{-3/2}\sum_t y_t)^2$$

Then, by the Continuous Mapping Theorem, using definition (11) and results (4), (5), (10) in section 2, and generalizing the last of these three results in a natural manner, we get:

$$T^{-5/2}\sum_t y_t^{*2} x_t^* \Rightarrow (s_v^2 s_w)\left[y_{12} - 2h_1 y_{11} - x_1(h_2 + 2h_1^2)\right] \quad .$$

A.4: Derivation of Equation (24)

This follows as for equation (23) in A.3 above. Interchanging variables, we get:

$$T^{-5/2}\sum_t x_t^{*2} y_t^* \Rightarrow (s_w^2 s_v)\left[y_{21} - 2x_1 y_{11} - h_1(x_2 + 2x_1^2)\right] \quad .$$

A.5: Derivation of Equation (26)

$$T^{-3}\sum_t y_t^{*4} = T^{-3}\left(\sum_t y_t^4 - 4\bar{y}\sum_t y_t^3 + 6\bar{y}^2\sum_t y_t^2 - 4\bar{y}^3\sum_t y_t + \sum_t \bar{y}^4\right)$$

$$= (T^{-3}\sum_t y_t^4) - 4(T^{-3/2}\sum_t y_t)(T^{-5/2}\sum_t y_t^3) + 6(T^{-5/2}\sum_t y_t)^2(T^{-2}\sum_t y_t^2)$$

$$-4(T^{-3/2}\sum_t y_t)^4 + (T^{-3/2}\sum_t y_t)^4 \quad .$$

Then, using the Continuous Mapping Theorem, result (7) from section 2 repeatedly, and gathering terms, we get:

$$T^{-3}\sum_t y_t^{*4} \Rightarrow s_v^4\left[h_4 - 4h_1 h_3 + 6h_1^2 h_2 - 3h_1^4\right] \quad .$$

A.6: Derivation of Equation (27)

This follows as for equation (26) in A.5 above. Interchanging variables, we get:

$$T^{-3}\sum_t x_t^{*4} \Rightarrow s_w^4\left[x_4 - 4x_1 x_3 + 6x_1^2 x_2 - 3x_1^4\right] \quad .$$

A.7: Derivation of Equation (28)

$$T^{-3}\sum_t y_t^{*2}x_t^{*2} = T^{-3/2}(y_t^2x_t^2 - 2\bar{x}x_ty_t + \bar{x}^2y_t^2 - 2\bar{y}y_tx_t^2 + 4\overline{xy}x_ty_t$$

$$-2\overline{yx}^2y_t + \bar{y}^2x_t^2 - 2\overline{xy}^2x_t + \bar{x}^2\bar{y}^2)$$

$$= (T^{-3}\sum_t x_t^2y_t^2) - 2(T^{-3/2}\sum_t x_t)(T^{-5/2}\sum_t x_ty_t^2) + (T^{-3/2}\sum_t x_t)^2(T^{-2}\sum_t y_t^2)$$

$$-2(T^{-3/2}\sum_t y_t)(T^{-5/2}\sum_t y_tx_t^2) + 4(T^{-3/2}\sum_t x_t)(T^{-3/2}\sum_t y_t)(T^{-2}\sum_t x_ty_t)$$

$$-2(T^{-3/2}\sum_t y_t)^2(T^{-3/2}\sum_t x_t)^2 + (T^{-3/2}y_t)^2(T^{-2}\sum_t x_t^2)$$

$$-2(T^{-3/2}\sum_t x_t)^2(T^{-3/2}\sum_t y_t)^2 + (T^{-3/2}\sum_t x_t)^2(T^{-3/2}\sum_t y_t)^2 \quad .$$

By the Continuous Mapping Theorem, using definition (11) and results (6), (7) and (10) in section 2, and generalizing the last of these three results in a natural manner, we get:

$$T^{-3}\sum_t y_t^{*2}x_t^{*2} \Rightarrow \left(s_v^2 s_w^2\right)\left[y_{22} - 2x_1y_{12} + x_1^2h_2 + h_1^2x_2 - 2h_1y_{21} + 4x_1h_1y_{11} - 3h_1^2x_1^2\right].$$

A.8: Derivation of Equation (29)

$$T^{-3}\sum_t y_t^{*3}x_t^{*} = T^{-3}(\sum_t x_ty_t^3 - \bar{y}\sum_t x_ty_t^2 + \bar{x}y_t^3 + \overline{xy}\sum_t y_t^2 - 2\bar{y}\sum_t x_ty_t^2 + 2\bar{y}\sum_t x_ty_t^2$$

$$+2\bar{y}^2\sum_t x_ty_t + 2\overline{xy}\sum_t y_t^2 - 2\overline{xy}^2\sum_t y_t + \bar{y}^2\sum_t x_ty_t - \bar{y}^2\bar{x}\sum_t y_t + T\overline{xy}^3)$$

$$= (T^{-3}\sum_t x_ty_t^3) - 3(T^{-3/2}\sum_t y_t)(T^{-5/2}\sum_t x_ty_t^2) - (T^{-3/2}\sum_t x_t)(T^{-5/2}\sum_t y_t^3)$$

$$+3(T^{-3/2}\sum_t x_t)(T^{-3/2}\sum_t y_t)(T^{-2}\sum_t y_t^2) + 3(T^{-3/2}\sum_t y_t)^2(T^{-2}\sum_t x_ty_t)$$

$$-3(T^{-3/2}\sum_t x_t)(T^{-3/2}\sum_t y_t)^3 \quad .$$

Again, by the Continuous Mapping Theorem, using definition (11) and results (7) and (10) in section 2, and generalizing the last of these three results in a natural manner, we get:

$$T^{-3}\sum_t y_t^{*3} x_t^* \Rightarrow (s_v^{\,3} s_w)\left[y_{13} - 3h_1 y_{12} - x_1 h_3 + 3h_1^2 y_{11} - 3x_1 h_1^3 + 3x_1 h_1 h_2\right] \quad .$$

A.9: Derivation of Equation (30)

This follows as for equation (29) in A.8 above, with the use of result (6). Interchanging variables, we get:

$$T^{-3}\sum_t x_t^{*3} y_t^* \Rightarrow (s_v s_w^{\,3})\left[y_{31} - 3x_1 y_{21} - h_1 x_3 + 3x_1^2 y_{11} - 3h_1 x_1^3 + 3h_1 x_1 x_2\right]$$

References

Billingsley, D. P. (1968), *Convergence of probability measures* (Wiley, New York).

Breusch, T. S. and A. R. Pagan (1980), The Lagrange multipler test and its applications to model specification in econometrics, *Review of Economic Studies*, 47, 239-253.

Godfrey, L. G. (1988), *Misspecification tests in econometrics: The Lagrange multiplier principle and other methods* (Cambridge University Press, Cambridge).

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

Granger, C. W. J. and P. Newbold (1977), *Forecasting economic time series* (academic Press, New York).

Greene, W. H. (2000), *Econometric analysis*, 4th. ed. (Prentice Hall, Upper Saddle River NJ)

Hamilton, J. B. (1994), *Time series analysis* (Princeton University Press, Princeton NJ).

Jarque, C. and A. Bera (1980), Efficient tests for normality, homoskedasticity and serial independence of regression residuals, *Economics Letters*, 6, 255-259.

McLeish, D. L. (1975), A Maximal inequality and dependent strong laws, *Annals of Probability*, 3, 829-839.

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

Plosser, C. I. And G. W. Schwert (1978), Money, income and sunspots: Measuring economic relationships and the effects of differencing, *Journal of Monetary Economics*, 4, 637-660.

SHAZAM (2001), *SHAZAM econometrics package: Users' reference manual, version 9* (McGraw-Hill, New York).

White, H. (1980), A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica*, 48, 817-838.