

SPURIOUS NONLINEAR REGRESSIONS IN ECONOMETRICS

TAE-HWAN KIM, YOUNG-SOOK LEE and PAUL NEWBOLD

School of Economics, University of Nottingham, Nottingham NG7 2RD, UK

January 2004

In this paper we consider the situation where two independent random walks are used in various frequently-employed nonlinear test and estimation procedures. We show analytically and by simulation that all nonlinear test and estimation procedures wrongly indicate that (i) the two independent random walks have a significant nonlinear relationship, and (ii) the spurious nonlinear relationship becomes stronger as the sample size approaches infinity.

1 Introduction

The seminal study of Granger and Newbold (1974) showed that when two independent random walks are used in a linear regression, one tends to find a significant relationship between the variables. This phenomenon is termed ‘spurious regression.’ Granger and Newbold discovered this phenomenon by undertaking a Monte Carlo experiment and later Phillips (1986) provided a rigorous and elegant proof.

The results in the two papers served as a springboard to a subsequent long series of investigations of the phenomenon for different types of regression and different types of data generating process. While Granger and Newbold (1974) and Phillips (1986) used driftless random walks, Entorf (1997) analyzed two independent random walks with non-zero drifts. It is now well understood that the spurious regression phenomenon also occurs for a wider class of time series; I(2) processes (Haldrup, 1994), non-stationary fractionally integrated processes (Marmol, 1998), stochastic unit root processes (Granger and Swanson, 1997) and some particular types of stationary processes (Tsay and Chung, 2000, Granger et al, 2001, Kim et al, 2003).

All these studies, however, rely on the use of linear regression and, therefore, demonstrate the existence of ‘spurious linear regression.’ To the best of our knowledge, no previous papers investigate the possibility of spurious nonlinear relationship between two independent variables. The use of linear regression is mainly motivated not by underlying economic insights or theories, but by its computational convenience and tractable interpretation of the results. It is more plausible and less restrictive to assume that the relationship between many economic variables is better characterized by a nonlinear specification. Many nonlinear models can provide better economic insights and have been increasingly used by applied economists.

This paper investigates the finite sample and large sample behavior of some popular nonlinear tests when they are employed to test if there exists a nonlinear relationship between two independent random walks. The paper is organized as follows. Section 2 briefly explains the nonlinear tests used in our study. We present in Section 3 some Monte Carlo evidence for spurious nonlinearity and provide theoretical explanations for the findings. Section 4 concludes the paper.

2 Nonlinear Tests

We consider two independent random walks y_t and x_t generated from the data generating process (DGP):

$$\begin{aligned}y_t &= y_{t-1} + \varepsilon_{yt}, \\x_t &= x_{t-1} + \varepsilon_{xt}.\end{aligned}\tag{1}$$

where $t = 1, 2, \dots, T$ and the error terms ε_{yt} and ε_{xt} are assumed to satisfy the following conditions.

Assumption 1. (i) ε_{yt} is i.i.d($0, \sigma_y^2$), (ii) ε_{xt} is i.i.d($0, \sigma_x^2$) and (iii) ε_{yt} and ε_{xt} are independent.

To investigate a possible nonlinear relationship between y_t and x_t , a researcher might conduct various nonlinear tests. Most nonlinear tests are usually based on the model:

$$y_t = \tilde{x}_t' \theta + g(x_t, \alpha, \beta) + \varepsilon_t$$

where $\tilde{x}_t = (1, x_t)'$ and $\theta = (\theta_0, \theta_1)'$. The function $g(x_t, \alpha, \beta)$ captures a possible nonlinear contribution of x_t , and α, β are some parameters governing the nonlinear behavior of that function. The nonlinear function g is assumed to satisfy $g(x_t, \alpha, \beta = 0) = 0$ so that the natural null hypothesis is then $H_0 : \beta = 0$. We include in our study five popular nonlinear tests and one newly developed nonlinear test by Hamilton (2000). All six tests are based as a preliminary step on the OLS regression of y_t on \tilde{x}_t . For our subsequent discussion, we define $\hat{\theta}$ to be the OLS estimator from that regression, $\hat{\varepsilon}_t$ to be the residuals ($\hat{\varepsilon}_t = y_t - \tilde{x}_t' \hat{\theta}$) and f_t to be the fitted value ($f_t = \tilde{x}_t' \hat{\theta}$).

1. The Ramsey RESET test: The Ramsey RESET test (regression specification error test) proposed by Ramsey (1969) exploits the idea that if there is no nonlinearity, then any nonlinear transformation of f_t should not be useful in explaining y_t . The test is based on the following auxiliary regression:

$$\hat{\varepsilon}_t = \tilde{x}_t' \tilde{\theta} + \sum_{k=1}^r \hat{a}_k f_t^{k+1} + \tilde{v}_t \quad (2)$$

for some $r \geq 1$. The null hypothesis is $H_0 : a_1 = \dots = a_r = 0$. The Ramsey RESET test statistic is

$$RESET = TR^2 \xrightarrow{d} \chi^2(r)$$

where R^2 is the uncentered R^2 from (2). When r is large, f_t^{k+1} can be highly correlated. To avoid the multicollinearity problem, the largest $r^* (< r)$ principal components are used.

2. The McLeod and Li test: The McLeod and Li test developed in McLeod and Li (1983) is obtained by applying the standard Ljung and Box Q-statistic to the squared residuals $\hat{\varepsilon}_t^2$. This test captures departures from linearity in mean. The test statistic is

$$McLeod = T(T+2) \sum_{k=1}^m \frac{\hat{\rho}(k)^2}{T-k} \xrightarrow{d} \chi^2(m)$$

where $\hat{\rho}(k)$ is the sample k^{th} order autocorrelation function of \hat{e}_t^2 given by

$$\hat{\rho}(k) = \frac{\sum_{t=k+1}^T (\hat{e}_t^2 - \hat{\sigma}^2)(\hat{e}_{t-k}^2 - \hat{\sigma}^2)}{\sum_{t=1}^T (\hat{e}_t^2 - \hat{\sigma}^2)^2}$$

with $\hat{\sigma}^2 = T^{-1} \sum_{t=1}^T \hat{e}_t^2$.

3. The Keenan test: Keenan (1985) suggests a test based on Tukey's non-additivity test using f_t^2 as in the RESET test. The procedure to obtain the statistic is as follows: (i) regress y_t on \tilde{x}_t to obtain f_t and \hat{e}_t (ii) regress f_t^2 on \tilde{x}_t and calculate the residuals denoted by $\hat{\xi}_t$, and (iii) regress \hat{e}_t on $\hat{\xi}_t$ and obtain the residuals \hat{v}_t . The Keenan test asks whether the squared fitted value f_t^2 has any additional forecasting ability for y_t . The test statistic is

$$Keenan = (T - 4) \frac{\hat{e}' \hat{\xi} (\hat{\xi}' \hat{\xi})^{-1} \hat{\xi}' \hat{e}}{\hat{v}' \hat{v}} \xrightarrow{d} \chi^2(1) \quad (3)$$

where \hat{e} , $\hat{\xi}$, \hat{v} are $T \times 1$ vectors of \hat{e}_t , $\hat{\xi}_t$, \hat{v}_t respectively.

4. The neural network test: The neural network test proposed by White (1989) is based on the following nonlinear specification

$$y_t = \tilde{x}_t' \theta + \sum_{k=1}^q \beta_k \psi(\tilde{x}_t' \gamma_k) + \varepsilon_t$$

where $\psi(z) = (1 + e^{-z})^{-1}$. The null hypothesis is $H_0 : \beta_1 = \dots = \beta_q = 0$ for particular choice of q and γ_k . Usually γ_k are randomly drawn from the uniform distribution within $[0, 1]$. The test statistic is

$$Neural = TR^2 \xrightarrow{d} \chi^2(q)$$

where R^2 is the uncentered R^2 from the OLS regression of \hat{e}_t on \tilde{x}_t and $\psi(\tilde{x}_t' \gamma_1), \dots, \psi(\tilde{x}_t' \gamma_q)$. As in the RESET test, the largest $q^* (< q)$ principal components are used to avoid the multicollinearity problem when q is large.

5. The White dynamic (first order) information matrix test: The dynamic information matrix test proposed by White (1992) is based on the following dynamic information matrix indicators:

$$m_t(x_t, \theta, \sigma) = S \text{vec}[s_t s_{t-1}']$$

where $s_t = s_t(x_t, \theta, \sigma)$ is the conditional score function of the likelihood of a linear model $y_t = \tilde{x}_t' \theta + \varepsilon_t$ with $\varepsilon_t \sim N(0, \sigma^2)$ and S is a nonstochastic selection matrix that can be used to focus attention on particular elements of s_t . The score function is

$$s_t(x_t, \theta, \sigma) = \sigma^{-1}(u_t, u_t x_t, u_t^2 - 1)'$$

where $u_t = \varepsilon_t/\sigma$. Let $\hat{s}_t = s_t(x_t, \tilde{\theta}, \tilde{\sigma})$ and $\hat{m}_t = S \text{vec}[\hat{s}_t \hat{s}'_{t-1}]$ where $\tilde{\theta}$ and $\tilde{\sigma}$ are the quasi-maximum likelihood estimators. Three test statistics can be obtained

$$White1 = T^{-1} \hat{M}'_T \hat{J}_T^{-1} \hat{M}_T \xrightarrow{d} \chi^2(q)$$

where $\hat{M}_T = T^{-1} \sum_{t=1}^T \hat{m}_t$, $\hat{J}_T = T^{-1} \sum_{t=1}^T \hat{m}_t \hat{m}'_t - (T^{-1} \sum_{t=1}^T \hat{m}_t \hat{s}'_t)(T^{-1} \sum_{t=1}^T \hat{s}_t \hat{s}'_t)^{-1} (T^{-1} \sum_{t=1}^T \hat{s}_t \hat{m}'_t)$ and q is the dimension of m_t ;

$$White2 = TR^2 \xrightarrow{d} \chi^2(q)$$

where R^2 is the uncentered R^2 from the OLS regression of the generalized residual $\hat{u}_t = (y_t - \tilde{x}'_t \tilde{\theta})/\tilde{\sigma}$ on \tilde{x}_t and \hat{m}_t/\hat{u}_t ;

$$White3 = TR^2 \xrightarrow{d} \chi^2(q)$$

where R^2 is the uncentered R^2 from the OLS regression of one on \hat{m}_t and \hat{s}_t .

6. The flexible nonlinear test: The flexible nonlinear test in Hamilton (2000) is based on

$$y_t = \tilde{x}'_t \theta + \lambda m(gx_t) + \varepsilon_t$$

where $m(\cdot)$ is a unknown nonlinear function. Naturally, the null hypothesis is $H_0 : \lambda = 0$ and its test statistic is given by

$$\begin{aligned} Hamilton &= \frac{[\hat{e}' H \hat{e} - \hat{\sigma}^2 \text{tr}(M H M)]^2}{\hat{\sigma}^4 \left(2 \text{tr} \{ [M H M - (T - k - 1)^{-1} M \text{tr}(M H M)]^2 \} \right)} \\ &\xrightarrow{d} \chi^2(1) \end{aligned}$$

where $M = I_T - X(X'X)^{-1}X'$, X is the $T \times 2$ matrix made of \tilde{x}_t , and H is a $T \times T$ matrix whose $(i, j)^{th}$ element H_{ij} is given by

$$\begin{aligned} H_{ij} &= 1 - \frac{h}{2} |x_i - x_j| & \text{if } \frac{h}{2} |x_i - x_j| \leq 1 \\ &= 0 & \text{otherwise} \end{aligned}$$

with $h = 2[T^{-1} \sum_{t=1}^T (x_t - \bar{x})^2]^{-1/2}$ where $\bar{x} = T^{-1} \sum_{t=1}^T x_t$. The Hamilton test is different from the previous five tests in that the user does not need to specify the nonlinear function $m(\cdot)$ and this is why the method is called 'flexible'.

3 Spurious Nonlinearity

We first present some simulation evidence for spurious nonlinearity. Two independent random walks y_t and x_t are generated through (1) and then the

six tests discussed in Section 2 are employed to test whether y_t is nonlinearly related to x_t for various sample sizes $T = 50, 100, 500, 1000$ and 10000 . The simulation results are reported in Table 1. First of all, it is striking that even with T as low as 50, the rejection rates of all six tests are considerably larger than the nominal level 5%. All White tests reject the null of linearity in favor of nonlinearity almost 100% of the time, displaying most severe spurious nonlinearity. They are followed by the McLeod (67%) and Hamilton (51%) tests. The RESET (40%), Keenan (38%) and Neural (38%) tests are the most resistant among the six tests considered. As the sample size increases, the rejection rates for the White, McLeod and Hamilton tests quickly converge to one while the rate of convergence is much slower for the other tests. For example, for the Keenan and Neural tests, rejection rates are 95% and 96% even when T is 10000. We have conducted an additional simulation for these tests and confirmed that the rejection rate does converge to one eventually.

The simulation results suggest that none of the six test statistics converge in distribution to their intended asymptotic distributions shown in Section 2, and in fact they tend to diverge to infinity. Given this strong simulation evidence, we now provide a theoretical explanation for this spurious nonlinear regression phenomenon. In particular, we investigate the asymptotic behavior of the RESET and Keenan tests since these two are not only among the most frequently used nonlinear tests by applied economists, but they also render the theoretical analysis tractable. Theorem 1 below shows the limiting distributions of the RESET and Keenan tests when scaled by T^{-1} .

Theorem 1 *Suppose y_t and x_t are generated by (1) and Assumption 1 holds. When $r = 1$ for the RESET test, we have*

$$\begin{aligned} T^{-1}RESET &\Rightarrow \Psi_{e\zeta}^2 / \Psi_{\zeta\zeta} \Psi_{ee}, \\ T^{-1}Keenan &\Rightarrow \Psi_{e\zeta}^2 / (\Psi_{\zeta\zeta} \Psi_{ee} - \Psi_{e\zeta}^2), \end{aligned}$$

where

$$\begin{aligned} \Psi_{e\zeta} &= \sigma_y^3 \zeta (2\omega - \pi) \left\{ \int_0^1 W(r) V(r) dr - \int_0^1 W(r) dr \int_0^1 V(r) dr \right\} \\ &\quad + \sigma_y^3 \zeta^2 \left\{ \int_0^1 W(r)^2 V(r) dr - \int_0^1 W(r)^2 dr \int_0^1 V(r) dr \right\} \\ &\quad - \sigma_y^3 \zeta^3 \left\{ \int_0^1 W(r)^3 dr - \int_0^1 W(r)^2 dr \int_0^1 W(r) dr \right\} \\ &\quad - \sigma_y^3 \zeta^2 (2\omega - \pi) \left[\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2 \right], \\ \Psi_{\zeta\zeta} &= \sigma_y^4 \zeta^2 (2\omega - \pi)^2 \left[\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2 \right] \end{aligned}$$

$$\begin{aligned}
& +2\sigma_y^4\zeta^3(2\omega - \pi) \left\{ \int_0^1 W(r)^3 dr - \int_0^1 W(r) dr \int_0^1 W(r)^2 dr \right\} \\
& +\sigma_y^4\zeta^4 \left[\int_0^1 W(r)^4 dr - \left\{ \int_0^1 W(r)^2 dr \right\}^2 \right], \\
\Psi_{ee} & = \sigma_y^2 \int_0^1 V(r)^2 dr - \left\{ \int_0^1 V(r) dr \right\}^2 - \sigma_y^2\zeta^2 \left[\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2 \right], \\
\zeta & = \frac{\int_0^1 W(r) V(r) dr - \int_0^1 W(r) dr \int_0^1 V(r) dr}{\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2}, \\
\omega & = \int_0^1 V(r) dr - \zeta \int_0^1 W(r) dr, \\
\pi & = 2\omega + \zeta \frac{\int_0^1 W(r)^3 dr - \int_0^1 W(r)^2 dr \int_0^1 W(r) dr}{\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2},
\end{aligned}$$

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

The limiting distributions are not Chi-square distributions and the results confirm our observation from the simulation that the statistics diverge to infinity, rejecting the null of linearity in favor of spurious nonlinearity.

4 Summary

While most papers in the spurious regression literature exclusively focus on the linear regression framework, in this paper we present both small sample and large sample evidence showing that ‘spurious nonlinear regressions’ can occur when two independent random walks are used in nonlinear estimation and testing procedures. Thus, this paper shows that the so-called spurious regression phenomenon can occur in practice more frequently and more widely than previously perceived. We conclude the paper by emphasizing the point that when interpreting nonlinear test results in favor of nonlinearity, applied economists should weigh the evidence against the persistence structure of the time series under investigation.

References

- [1] Entorf, H. (1997) Random walks with drifts: Nonsense regression and spurious fixed-effect estimation. *Journal of Econometrics* **80**, 287-296.
- [2] Granger, C.W.J., Hyung, N. and Jeon, Y. (2001) Spurious regressions with stationary series. *Applied Economics* **33**, 899-904.

- [3] Granger, C.W.J. and Newbold, P. (1974) Spurious regressions in econometrics. *Journal of Econometrics* **2**, 111-120.
- [4] Granger, C.W.J. and Swanson, N. (1997) An introduction to stochastic unit-root processes. *Journal of Econometrics* **80**, 35-62.
- [5] Haldrup, N. (1994) The asymptotics of single-equation cointegration regressions with I(1) and I(2) variables. *Journal of Econometrics* **63**, 153-181.
- [6] Hamilton, J.D. (2000) A parametric approach to flexible nonlinear inference. *Econometrica* **69**, 801-812.
- [7] Keenan, K.L. (1985) A Tukey nonadditivity type test for time series nonlinearity. *Biometrika* **72**, 39-44.
- [8] Kim, T., Lee, Y. and Newbold, P. (2003) Spurious regressions with processes around linear trends or drifts. Mimeo, University of Nottingham, School of Economics.
- [9] Marmol, F. (1998) Spurious regression theory with nonstationary fractionally integrated processes. *Journal of Econometrics* **84**, 233-250.
- [10] McLeod, A.I. and Li, W.I. (1983) Diagnostic checking ARMA time series models using squared residual autocorrelation. *Journal of Time Series Analysis* **4**, 169-176.
- [11] Phillips, P.C.B. (1986) Understanding spurious regressions in econometrics. *Journal of Econometrics* **33**, 311-340.
- [12] Ramsey, J.B. (1969) Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society B* **31**, 350-371.
- [13] Tsay, W.J. and Chung, C.F. (2000) The spurious regression of fractionally integrated processes. *Journal of Econometrics* **96**, 155-182.
- [14] White, H. (1989) An additional hidden unit test for neglected nonlinearity in multilayer feedforward networks. in: Proceedings of the international joint conference on neural networks, Washington, DC (IEEE Press, New York), vol. 2: 451-455.
- [15] White, H. (1992) *Estimation, Inference and Specification Analysis*. Cambridge University Press, New York.

APPENDIX

Proof of Theorem 1. We first consider the Keenan test. The test statistic in (3) can be written as

$$\begin{aligned} T^{-1}Keenan &= \frac{(\sum_{t=1}^T \hat{e}_t \hat{\xi}_t)^2}{\sum_{t=1}^T \hat{\xi}_t^2 \sum_{t=1}^T \hat{v}_t^2} + o_p(1) \\ &= \frac{(T^{-5/2} \sum_{t=1}^T \hat{e}_t \hat{\xi}_t)^2}{(T^{-3} \sum_{t=1}^T \hat{\xi}_t^2)(T^{-2} \sum_{t=1}^T \hat{e}_t^2) - (T^{-5/2} \sum_{t=1}^T \hat{e}_t \hat{\xi}_t)^2} + o_p(1) \end{aligned} \quad (4)$$

where $\hat{e}_t = y_t - \hat{\theta}_0 - \hat{\theta}_1 x_t$, $\hat{\xi}_t = f_t^2 - \hat{a} - \hat{b}x_t$, $\hat{v}_t = \hat{e}_t - \hat{\delta}\hat{\xi}_t$, and \hat{a} , \hat{b} , $\hat{\delta}$ are the corresponding OLS estimators. There are three terms in (4) and we examine them separately. The first term is

$$\begin{aligned} T^{-5/2} \sum_{t=1}^T \hat{e}_t \hat{\xi}_t &= \{2(T^{-1/2}\hat{\theta}_0)\hat{\theta}_1 - (T^{-1/2}\hat{b})\} \{T^{-2} \sum_{t=1}^T (y_t - \bar{y})(x_t - \bar{x})\} \\ &\quad + \hat{\theta}_1 \{T^{-5/2} \sum_{t=1}^T (y_t - \bar{y})(x_t^2 - \bar{x}^2)\} \\ &\quad - \hat{\theta}_1^3 \{T^{-5/2} \sum_{t=1}^T (x_t - \bar{x})(x_t^2 - \bar{x}^2)\} \\ &\quad - \{2(T^{-1/2}\hat{\theta}_0)\hat{\theta}_1^2 - \hat{\theta}_1(T^{-1/2}\hat{b})\} \{T^{-2} \sum_{t=1}^T (x_t - \bar{x})^2\} \end{aligned}$$

where \bar{y} , \bar{x} , \bar{x}^2 are the sample means of y_t , x_t , x_t^2 respectively. The limit of each term of $T^{-5/2} \sum_{t=1}^T \hat{e}_t \hat{\xi}_t$ is as follows:

$$\hat{\theta}_1 \Rightarrow \frac{\sigma_y \left\{ \int_0^1 W(r) V(r) dr - \int_0^1 W(r) dr \int_0^1 V(r) dr \right\}}{\sigma_x \left[\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2 \right]} \equiv \frac{\sigma_y}{\sigma_x} \zeta.$$

Note that $T^{-3/2} \sum_{t=1}^T x_t \Rightarrow \sigma_x \int_0^1 W(r) dr$, $T^{-3/2} \sum_{t=1}^T y_t \Rightarrow \sigma_y \int_0^1 V(r) dr$ and $T^{-2} \sum_{t=1}^T x_t y_t \Rightarrow \sigma_x \sigma_y \int_0^1 W(r) V(r) dr$ from which we have:

$$T^{-1/2}\hat{\theta}_0 \Rightarrow \sigma_y \left\{ \int_0^1 V(r) dr - \zeta \int_0^1 W(r) dr \right\} \equiv \sigma_y \omega,$$

$$\begin{aligned} T^{-2} \sum_{t=1}^T (y_t - \bar{y})(x_t - \bar{x}) &= T^{-2} \sum_{t=1}^T x_t y_t - (T^{-3/2} \sum_{t=1}^T x_t)(T^{-3/2} \sum_{t=1}^T y_t) \\ &\Rightarrow \sigma_x \sigma_y \int_0^1 W(r) V(r) dr - \sigma_x \sigma_y \int_0^1 W(r) dr \int_0^1 V(r) dr, \end{aligned}$$

$$\begin{aligned}
T^{-5/2} \sum_{t=1}^T (y_t - \bar{y})(x_t^2 - \bar{x}^2) &= T^{-5/2} \sum_{t=1}^T x_t^2 y_t - (T^{-2} \sum_{t=1}^T x_t^2)(T^{-3/2} \sum_{t=1}^T y_t) \\
&\Rightarrow \sigma_x^2 \sigma_y \int_0^1 W(r)^2 V(r) dr - \sigma_x^2 \sigma_y \int_0^1 W(r)^2 dr \int_0^1 V(r) dr,
\end{aligned}$$

$$\begin{aligned}
T^{-5/2} \sum_{t=1}^T (x_t - \bar{x})(x_t^2 - \bar{x}^2) &= T^{-5/2} \sum_{t=1}^T x_t^3 - (T^{-2} \sum_{t=1}^T x_t^2)(T^{-3/2} \sum_{t=1}^T x_t) \\
&\Rightarrow \sigma_x^3 \int_0^1 W(r)^3 dr - \sigma_x^3 \int_0^1 W(r)^2 dr \int_0^1 W(r) dr,
\end{aligned}$$

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

$$\begin{aligned}
T^{-1/2} \hat{b} &= \frac{T^{-5/2} \sum_{t=1}^T (x_t - \bar{x})(f_t^2 - \bar{f}^2)^2}{T^{-2} \sum_{t=1}^T (x_t - \bar{x})^2} \\
&= 2(T^{-1/2} \hat{\theta}_0) \hat{\theta}_1 + \frac{\hat{\theta}_1^2 \{T^{-5/2} \sum_{t=1}^T (x_t - \bar{x})^2 (x_t^2 - \bar{x}^2)\}}{T^{-2} \sum_{t=1}^T (x_t - \bar{x})^2} \\
&\Rightarrow \zeta \frac{\sigma_y^2}{\sigma_x} \left[2\omega + \zeta \frac{\int_0^1 W(r)^3 dr - \int_0^1 W(r)^2 dr \int_0^1 W(r) dr}{\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2} \right] \equiv \zeta \frac{\sigma_y^2}{\sigma_x} \pi.
\end{aligned}$$

Combining all the limit results, we obtain

$$\begin{aligned}
T^{-5/2} \sum_{t=1}^T \hat{\epsilon}_t \hat{\xi}_t &\Rightarrow \sigma_y^3 \zeta (2\omega - \pi) \left\{ \int_0^1 W(r) V(r) dr - \int_0^1 W(r) dr \int_0^1 V(r) dr \right\} \\
&\quad + \sigma_y^3 \zeta^2 \left\{ \int_0^1 W(r)^2 V(r) dr - \int_0^1 W(r)^2 dr \int_0^1 V(r) dr \right\} \\
&\quad - \sigma_y^3 \zeta^3 \left\{ \int_0^1 W(r)^3 dr - \int_0^1 W(r)^2 dr \int_0^1 W(r) dr \right\} \\
&\quad - \sigma_y^3 \zeta^2 (2\omega - \pi) \left[\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2 \right] \\
&\equiv \Psi_{e\zeta}.
\end{aligned}$$

Next, we turn to the second term $T^{-3} \sum_{t=1}^T \hat{\xi}_t^2$ which is given by

$$T^{-3} \sum_{t=1}^T \hat{\xi}_t^2 = \{2(T^{-1/2} \hat{\theta}_0) \hat{\theta}_1 - (T^{-1/2} \hat{b})\}^2 \{T^{-2} \sum_{t=1}^T (x_t - \bar{x})^2\}$$

$$\begin{aligned}
& +2\hat{\theta}_1^2\{2(T^{-1/2}\hat{\theta}_0)\hat{\theta}_1 - (T^{-1/2}\hat{b})\}\{T^{-5/2}\sum_{t=1}^T(x_t - \bar{x})(x_t^2 - \bar{x}^2)\} \\
& +\hat{\theta}_1^4\{T^{-3}\sum_{t=1}^T(x_t^2 - \bar{x}^2)^2\}.
\end{aligned}$$

The only new term is the last term whose limit is given by

$$\begin{aligned}
T^{-3}\sum_{t=1}^T(x_t^2 - \bar{x}^2)^2 &= T^{-3}\sum_{t=1}^T x_t^4 - (T^{-2}\sum_{t=1}^T x_t^2)^2 \\
&\xrightarrow{d} \sigma_x^4 \int_0^1 W(r)^4 dr - \sigma_x^4 \left\{ \int_0^1 W(r)^2 dr \right\}^2.
\end{aligned}$$

Hence, it follows that

$$\begin{aligned}
T^{-3}\sum_{t=1}^T \hat{\xi}_t^2 &\Rightarrow \sigma_y^4 \zeta^2 (2\omega - \pi)^2 \left[\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2 \right] \\
&\quad + 2\sigma_y^4 \zeta^3 (2\omega - \pi) \left\{ \int_0^1 W(r)^3 dr - \int_0^1 W(r) dr \int_0^1 W(r)^2 dr \right\} \\
&\quad + \sigma_y^4 \zeta^4 \left[\int_0^1 W(r)^4 dr - \left\{ \int_0^1 W(r)^2 dr \right\}^2 \right] \\
&\equiv \Psi_{\zeta\zeta}.
\end{aligned}$$

The last term in (4) can be written as

$$\begin{aligned}
T^{-2}\sum_{t=1}^T \hat{e}_t^2 &= T^{-2}\sum_{t=1}^T \{(y_t - \bar{y}) - \hat{\theta}_1(x_t - \bar{x})\}^2 \\
&\Rightarrow \sigma_y^2 \int_0^1 V(r)^2 dr - \left\{ \int_0^1 V(r) dr \right\}^2 \\
&\quad - \sigma_y^2 \zeta^2 \left[\int_0^1 W(r)^2 dr - \left\{ \int_0^1 W(r) dr \right\}^2 \right] \\
&\equiv \Psi_{ee}.
\end{aligned}$$

Therefore, we have the desired result:

$$T^{-1}Keenan \Rightarrow \frac{\Psi_{e\zeta}^2}{\Psi_{\zeta\zeta}\Psi_{ee} - \Psi_{e\zeta}^2}.$$

Next, we turn to the RESET test. When $r = 1$, the statistic is TR^2 and R^2 is the uncentered R^2 from the following auxiliary regression:

$$\hat{e}_t = \tilde{\theta}_0 + x_t \tilde{\theta}_1 + \hat{a}_1 f_t^2 + \tilde{v}_t. \tag{5}$$

Now let us compare this regression with the regression in the third step for the Keenan test, which is

$$\begin{aligned}\hat{e}_t &= \hat{\delta}\hat{\xi}_t + \hat{v}_t \\ &= \hat{\delta}(f_t^2 - \hat{a} - \hat{b}x_t) + \hat{v}_t.\end{aligned}\tag{6}$$

Hence, it can be easily seen that the two auxiliary regressions in (5) and (6) are identical, which implies that the uncentered R^2 can be calculated using (6) and is given by

$$R^2 = \frac{(\sum_{t=1}^T \hat{e}_t \hat{\xi}_t)^2}{\sum_{t=1}^T \hat{\xi}_t^2 \sum_{t=1}^T \hat{e}_t^2}.$$

Therefore, we have

$$\begin{aligned}T^{-1}RESET &= \frac{(T^{-5/2} \sum_{t=1}^T \hat{e}_t \hat{\xi}_t)^2}{(T^{-3} \sum_{t=1}^T \hat{\xi}_t^2)(T^{-2} \sum_{t=1}^T \hat{e}_t^2)} \\ &\Rightarrow \frac{\Psi_{e\zeta}^2}{\Psi_{\zeta\zeta}\Psi_{ee}}\end{aligned}$$

which completes the proof. ■

Table 1. Proportion of rejections of nonlinear tests at nominal 0.05 level

T	50	100	500	1000	10000
RESET	0.40	0.63	0.89	0.96	1.00
McLeod	0.67	0.99	1.00	1.00	1.00
Keenan	0.38	0.55	0.80	0.85	0.95
Neural	0.38	0.56	0.80	0.85	0.96
White1	1.00	1.00	1.00	1.00	1.00
White2	1.00	1.00	1.00	1.00	1.00
White3	1.00	1.00	1.00	1.00	1.00
Hamilton	0.51	0.73	0.98	1.00	1.00