E C O N O M I C S B U L L E T I N

# F versus t tests for unit roots: a comment

Paulo M. M. Rodrigues
*Faculty of Economics, University of Algarve*

Andrew Tremayne
*Department of Economics, University of York*

## *Abstract*

In this note we provide justification for some Monte Carlo results presented by Elder and Kennedy (2001). In particular we show that the severe size distortions observed by Elder and Kennedy are due to the presence of nuisance parameters in the data generation process, but ignored in the test regression. As is shown in a small Monte Carlo exercise, correct size for the statistics is obtained when an adequate test regression is considered.

**Citation:** M. M. Rodrigues, Paulo and Andrew Tremayne, (2004) "F versus t tests for unit roots: a comment." *Economics Bulletin*, Vol. 3, No. 12 pp. 1−7
**Submitted:** February 20, 2004. **Accepted:** April 9, 2004.
**URL:** http://www.economicsbulletin.com/2004/volume3/EB−04C40002A.pdf

# F versus t tests for unit roots: a comment


Paulo M. M. Rodrigues
Faculty of Economics
University of Algarve
Campus de Gambelas
8005-139 Faro, Portugal
*prodrig@ualg.pt*

Andrew Tremayne
University of York
Department of Economics
Heslington,
York YO10 5DD, UK
*art7@york.ac.uk*


February, 2004


**Abstract**

In this note we provide justification for some Monte Carlo results presented by Elder and Kennedy (2001). In particular we show that the severe size distortions observed by Elder and Kennedy are due to the presence of nuisance parameters in the data generation process, but ignored in the test regression. As is shown in a small Monte Carlo exercise, correct size for the statistics is obtained when an adequate test regression is considered.


## 1 Introduction

In a recent paper, Elder and Kennedy (2001), henceforth EK, present some interesting results concerning F and t tests for unit roots, though their Monte Carlo results deserve some discussion. In particular the last two columns of Tables I - III require further consideration.

The data generation process (DGP) considered by Elder and Kennedy is,

$$(1-\rho L)(y_t-\mu)=\varepsilon_t, \tag{1}$$

where $\varepsilon_t$ is an error term with mean zero and variance $\sigma^2$ and $L$ is the usual lag operator. The appropriate test regression to consider in this context is,

$$\begin{aligned} y_t &= \mu(1-\rho)+\rho y_{t-1}+\varepsilon_t \\ &= \alpha+\rho y_{t-1}+\varepsilon_t \end{aligned} \tag{2}$$

The test regression is obtained under the assumption that the underlying (DGP) is a first order autoregression with a nonzero mean, $\mu \neq 0$, when $|\rho| < 1$. Typically then, to achieve similarity, i.e. to eliminate the potential influence of nuisance parameters on the test statistics, deterministic variables (in this particular case a constant which has the effect of demeaning the data) are included in the test regression. Following, Dickey and Fuller (1979, 1981), the unit root ($\rho = 1$) can either be tested by means of a one-sided t-test on the ordinary least-squares estimate, $\widehat{\rho}$, or a joint test can be adopted to investigate the validity of the joint null hypothesis $(\alpha, \rho) = (0, 1)$.

The main conclusion of EK's analysis suggests that under stationarity the value of $\alpha$ has no impact on the F-statistics, but we shall see that this is not the crucial feature of these tests because the null hypothesis of a unit root must be handled more carefully than in EK.

# 2 Starting Values

A further important consideration relates to the role of initial values. This can be observed from the following, assuming that the initial value $y_0$ is independent of $\varepsilon_t$. Then, in accordance with (2), the variance of $y_t$ is represented as,

$$var(y_t) = \rho^{2t} var(y_0) + \left( \sum_{j=0}^{t-1} \rho^{2j} \right) \sigma^2. \tag{3}$$

Thus, when $|\rho| \geq 1$, $y_t$ cannot be stationary since $var(y_t)$ will increase without bound as $t \to \infty$. On the other hand, for $y_t$ to be stationary when $|\rho| < 1$, it is necessary that $var(y_t) = ... = var(y_0) = \gamma_0$. It is straightforward to show that, if

$$var(y_0) = \frac{\sigma^2}{(1 - \rho^2)}, \tag{4}$$

then

$$var(y_t) = \frac{\sigma^2}{(1 - \rho^2)} = \gamma_0 \tag{5}$$

and the autocovariaces are

$$\gamma_k = \rho^k \gamma_0 \tag{6}$$

for all $t$. If, however, $var(y_0) \neq \frac{\sigma^2}{(1-\rho^2)}$, then $var(y_t)$ will not be a constant independent of $t$ (as long as $\sigma^2 > 0$), and $y_t$ will not be stationary. Hence, for $y_t$ to be stationary, it is not enough to have $|\rho| < 1$ and $y_0$ independent of $\varepsilon_t$; $var(y_0) = \frac{\sigma^2}{(1-\rho^2)}$ must also hold. Clearly, the assumptions on the initial values are relevant for the stationarity condition; see, for example, the discussion in Dickey, Bell and Miller (1986, Appendix A).

In recent unit root literature, Monte Carlo investigations commonly contemplate two scenarios: *i*) the conditional case, where the starting values are assumed to be fixed or random $O_p(1)$ variables; and *ii*) the unconditional case, where the starting values are assumed to be random draws from a distribution with zero mean and variance $\frac{\sigma^2}{(1-\rho^2)}$.

The results found in Tables I and II of EK's paper are obtained under the conditional case ($y_0 = 0$ and $y_0 = \frac{\sigma^2}{(1-\rho^2)}$, respectively), while, the results in Table III are computed based on the unconditional case ($y_0 \backsim N[\frac{\alpha}{(1-\rho)}, \frac{\sigma^2}{(1-\rho^2)}]$); see Elliott (1999) for a more detailed discussion of this issue.

It is pertinent to ask why the values in the last two columns of Table I are so far away from 0.05. It turns out that the answer to this question stems from the fact that an inappropriate test regression is being employed.

# 3 Monte Carlo Contradiction

The findings generated from the last three columns of Tables I-III in EK merit some further consideration. One would expect the size of the tests to be close to the nominal 5% level considered, however, we observe in all three tables that for $\rho = 1$ and $\alpha \neq 0$, $\Phi_1$ and their $\widehat{\tau}_{\mu,two-sided}$ are oversized and that $\widehat{\tau}_{\mu,one-sided}$ is considerably undersized. For instance, although using the procedure applied by Godfrey and Orme (2000, page 75) to assess whether observed rejection proportions under the null are acceptable, all three leftmost columns of the rightmost panels of Tables I-III are clearly satisfactory, we see that none of the other entries corresponding to $\alpha \neq 0$ in those (rightmost) panels are. For example, if an allowance of 0.5% either side of a nominal level of 5% is deemed satisfactory, using Godfrey and Orme's tests based on least favourable null hypotheses with 10,000 replications, the estimated size must lie in the interval $4.16 - 5.88\%$; all of the 18 entries to which we refer are well outside these bounds.

This is consistent with the notion that a non-similar regression was used. More specifically, if (2) was used as DGP as well as test regression, one should note that, for $|\rho| < 1$ the test regression remains similar because it accounts for the presence of a non zero intercept under the alternative hypothesis.

However, when $\rho = 1$, this is not the case. It is straightforward to see in this case that (2) can be rewritten as,

$$y_t = \underset{O(T)}{\alpha t} + \underset{O_p(1)}{y_0} + \underset{O_p(\sqrt{T})}{\sum_{k=1}^{t} \varepsilon_k} \tag{7}$$

(the orders of magnitude of each term on the right hand side are given below them). As a result, when $\alpha \neq 0$, a drift term is introduced into the DGP when $\rho = 1$, resulting in non-trivial implications for the test statistics, as noted by Haldrup and Hylleberg (1995), and as can be observed from Tables I-III in EK.

From (7) the main problem resides in the deterministic drift term, which is $O(T)$. It dominates the stochastic trend (which is $O_p(\sqrt{T})$, and consequently the use of a non-similar regression will, without careful consideration, imply the use of inappropriateate distributions. As was shown by Haldrup and Hylleberg (1995) and Rodrigues (2001), when deterministic trends present in the DGP are not accounted for in the test regression, then the t-statistic on $\widehat{\rho}$ will converge to a mixture of normal distributions and Brownian motions, with the normal component displaying greater dominance as the magnitude of $\alpha$ increases. Nankervis and Savin (1987) and West (1988) also have interesting discussions of this problem.

# 4 Simulation Study

In this Section we perform a small simulation study to show that test statistics computed from test regressions which account for the presence of potential nuisance parameters present in the DGP will display correct empirical size. The tests considered are: $t_{\tau-one-sided}$, the left sided t-test that $\phi < 0$; $t_{\tau-two-sided}$ representing the two-sided t-test that $\phi \neq 0$; $\Phi_2$ the joint test of the null hypothesis $H_0 : (\gamma_0, \gamma_1, \phi) = (0, 0, 0)$; and $\Phi_3$ the joint test of the null hypothesis $H_0 : (\gamma_0, \gamma_1, \phi) = (\gamma_0, 0, 0)$. The notation for these last two tests follows that of Dickey and Fuller (1981); notice that the set of test statistics we consider does not include $\Phi_1$, unlike EK.

The DGP considered in this simulation is (2) with $\rho = \{1, 0.99, 0.95, 0.90, 0.80, 0.70\}$ and $\alpha \in \{0, 0.5, 1.0\}$; compare EK's tables and also Table VII of Dickey and Fuller (1981). Furthermore, $\varepsilon_t \backsim N(0, 1)$ and the RNDN function of GAUSS for Windows NT/95 Version 3.2.38 is used. Thus, the DGP considered is the same as in EK's paper (with sample size reported 100, as there), however the test statistics are computed from the test regression,

$$\Delta y_t = \gamma_0 + \gamma_1 t + \phi y_{t-1} + u_t. \tag{8}$$

Note that EK in all three tables consider stochatically bounded ($O_p(1)$) starting values. Hence, under the test regression (8) considered these will have no impact on the test statistics. Hence, without loss of generality (see Phillips (1987)), we consider $y_0 = 0$. We experimented with other starting values and obtained directly comparable results; details are available from the authors on request. The results of our small experiment are presented in Table I, but they convincingly demonstrate the thrust of our argument. All entries are based on 20000 replications.

**Table I - Empirical Size and Power of 0.05 Tests for Sample Size 100** ($y_0 = 0$)

| | $\alpha$ | $\rho$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | 1 | 0.99 | 0.95 | 0.90 | 0.80 | 0.70 |
| $t_{\tau-one-sided}$ | 0 | .046 | .049 | .083 | .186 | .642 | .960 |
| $t_{\tau-two-sided}$ | | .048 | .046 | .045 | .101 | .457 | .889 |
| $\Phi_2$ | | .048 | .033 | .038 | .085 | .415 | .859 |
| $\Phi_3$ | | .047 | .046 | .063 | .138 | .549 | .930 |
| $t_{\tau-one-sided}$ | 0.5 | .046 | .045 | .091 | .214 | .670 | .963 |
| $t_{\tau-two-sided}$ | | .048 | .064 | .053 | .121 | .486 | .894 |
| $\Phi_2$ | | .975 | .659 | .091 | .128 | .450 | .870 |
| $\Phi_3$ | | .047 | .064 | .104 | .183 | .580 | .935 |
| $t_{\tau-one-sided}$ | 1.0 | .046 | .034 | .130 | .310 | .736 | .969 |
| $t_{\tau-two-sided}$ | | .048 | .107 | .080 | .186 | .561 | .913 |
| $\Phi_2$ | | 1.00 | 1.00 | .538 | .315 | .561 | .898 |
| $\Phi_3$ | | .047 | .139 | .353 | .362 | .676 | .948 |

As can be observed from this Table, the results for $\Phi_2$, $\Phi_3$ and the left sided t-ratio, $t_{\tau-one-sided}$, are in accordance with the results in Dickey and Fuller (1981, p.1067). This confirms the crucial importance of considering adequate test regressions in order to obtain asymptotically invariant tests, which is our basic message. In more detail, the large rejection frequencies for $\Phi_2$ when $\gamma_0 \neq 0$ are detecting this failure of the null; this is mirrored in EK's work by the seemingly poor 'size' properties of $\Phi_1$ seen in the last two columns of the rightmost blocks of their tables. When the null fails because $\rho < 1$ (with $\gamma_0 = 0$), the $t_{\tau-one-sided}$ test is preferable to all the others because it is one-sided; when this zero restriction fails, the tests based on the $F-distribution$ can sometimes outperform the $t-tests$ (particularly when $\rho = 0.95$ or 0.99), whether or not they are implemented in their one-sided guise.

# 5 Conclusion

This note reconsiders some of the points made in the paper by Elder and Kennedy (2001) for some unit root test statistics and provides an explanation for some of their Monte Carlo results. We show that the severe size distortions observed by Elder and Kennedy in their work are due to the presence of nuisance parameters in the data generation process, but ignored in the test regression. This renders the size of their tests unacceptable when implemented as they suggest. In a small Monte Carlo exercise correct size for the statistics is obtained when an adequate test regression is considered and indicative results for power are presented showing that there is an advantage from using a t-test for a unit root rather than the Dickey-Fuller (1981) statistics based on the $F-$distribution when $\gamma_0$ is zero in truth. This is understandable given the one-sided nature of the former and the two-sided nature of the latter, but no blanket recommendation that this form of the test should always be used obtains.

# References


[1] Dickey, D.A., W.R. Bell and R.B. Miller (1986) "Unit Roots in Time Series Models: Tests and Implications" *The American Statistician* 40, 12-26.

[2] Dickey, D.A., and W.A. Fuller (1979) "Distribution of the Estimators for Autoregressive Time Series with a Unit Root" *Journal of the American Statistical Association **74**, 427-431.*

[3] Dickey, D.A., and W.A. Fuller (1981) "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root" *Ecoometrica **49**, 1057-1072.*

[4] Elder, J., and P. Kennedy (2001) "F versus t Tests for a Unit Root" *Economics Bulletin **3**, 1-6.*

[5] Elliott, G. (1999) "Efficient Tests for a Unit Root when the Initial Observation is Drawn From its Unconditional Distribution" *International Economic Review **40**, 767-783.*

[6] Godfrey, L.G., and C.D. Orme (2000) "Controlling the Significance Levels of Prediction Error Tests for Linear Regression Models" *Econometrics Journal **3**, 66-83.*

[7] Haldrup, N., and S. Hylleberg (1995) "A Note on the Distribution of the Least Squares Estimator of a Random Walk with Drift: Some Analytical Evidence" *Economics Letters **48**, 221-228.*

[8] Nankervis, J.C., and N.E. Savin (1987) "Finite Sample Distributions of $t$ and $F$ Statistics in an AR(1) Model with an Exogenous variable" *Econometric Theory* **3**, *387-408.*

[9] Phillips, P.C.B. (1987) "Towards a Unified Asymptotic Theory for Autoregression" *Biometrika **74**, 535-547.*

[10] Rodrigues, P.M.M. (2001) " The Behaviour of Seasonal Unit Root Tests under neglected Local Drifts" *Portuguese Economic Journal **1**,* 27-46.

[11] West, K.D. (1988) " Asymptotic Normality when Regressors have a Unit Root" *Econometrica* ***56****, 1397-1417.*