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

More on F versus t tests for unit roots when there is no trend

John Elder
North Dakota State University

Peter E. Kennedy
Simon Fraser University

Abstract

Rodrigues and Tremayne (2004) interpret a problematic size result in a Monte Carlo study reported in Elder and Kennedy (2001) as arising from Elder and Kennedy's use of an inappropriate testing equation. In expositing their result, Rodrigues and Tremayne inadvertently lead readers to believe that the Elder and Kennedy conclusion is in error. We clarify the Rodrigues and Tremayne contribution, putting the validity of the Elder and Kennedy result in proper perspective and underlining the important role played by the starting value in Monte Carlo analyses.

Our thanks, without implication, to Paulo Rodrigues and Andrew Tremayne for comments on an earlier version.
Citation: Elder, John and Peter E. Kennedy, (2004) "More on F versus t tests for unit roots when there is no trend." *Economics Bulletin*, Vol. 3, No. 37 pp. 1−6
Submitted: September 16, 2004. **Accepted:** October 8, 2004.
URL: http://www.economicsbulletin.com/2004/volume3/EB−04C10023A.pdf

1 Introduction

In Elder and Kennedy (2001), henceforth EK, we show that when it is known that there is no trend in the data, the classic Φ_1 F test for a unit root, introduced in Dickey and Fuller (1981), is dominated by the Dickey-Fuller one-sided t test (called τ_μ. in the literature). This happens because, remarkably, adding an extra constraint to the null for this application has no influence on the test statistic, and so the one-sided nature of the t test causes it to be more powerful.

The EK result applies to a context in which it is known that there is no trend in the data, so that the null is a unit root with no drift. Unfortunately, when we simulated data for a true null of a unit root, we allowed the intercept to be non-zero, resulting in trending data. A consequence of this is that in the final two columns of our tables, the type I error for the t test is not equal to its chosen value of 0.05. The context is no trend in the data, so although for the alternative hypothesis an intercept is allowed when generating Monte Carlo data, for the unit root null no intercept is allowed. The test equation, however, must have an intercept, to eliminate the potential influence of this nuisance parameter on test statistics, but should not have a trend term because the trend is known to be irrelevant. We included an intercept in the test regression, and did not include a trend, so our power results, the focus of our paper, are unaffected by our failure to generate appropriate null data; fortuitously, then, this oversight is of no consequence for our results.

2 Anomalous Size

Rodriques and Tremayne (2004), henceforth RT, have explained that the anomalous size results in EK's final two columns result from EK not employing a test equation with a trend term. In one respect, this is correct – if the null is a unit root with drift, a trend is necessary in the testing equation, implying that EK have erred in not using an appropriate testing equation. In another respect this is not correct – if the null is a unit root with no drift, no trend is necessary in the testing equation, implying only that EK have erred by generating inappropriate Monte Carlo null data. RT have postulated that the former is the case, and have presented for readers an exposition of why using an inappropriate testing equation can be disastrous in that it creates a test with incorrect size, possibly dramatically so. They have presented Monte Carlo results illustrating how using an appropriate testing equation, in this case a testing equation including a trend, creates a test with proper size.

RT's purpose was to warn readers that care must be taken in choosing the testing equation for unit root testing. Their intention was to focus entirely on the issue of size, ignoring the issue of power addressed by EK. In this respect they did not intend to suggest that the EK results were incorrect, and a very close reading of their paper shows that they have come close to succeeding. Unfortunately, they have included power results in their Monte Carlo study, and have commented on power in their conclusion. For example, they have phrased their final sentence in such a way ("no blanket recommendation that this form of the test should always be used obtains") as to lead readers to believe that the EK result (that when it is known that there is no trend, use the t test, not the F test) is invalid. The purpose of the remainder of this paper is to ensure that readers do not draw this conclusion from the RT paper, and to underline the fact that starting values for Monte Carlo studies can markedly influence power results.

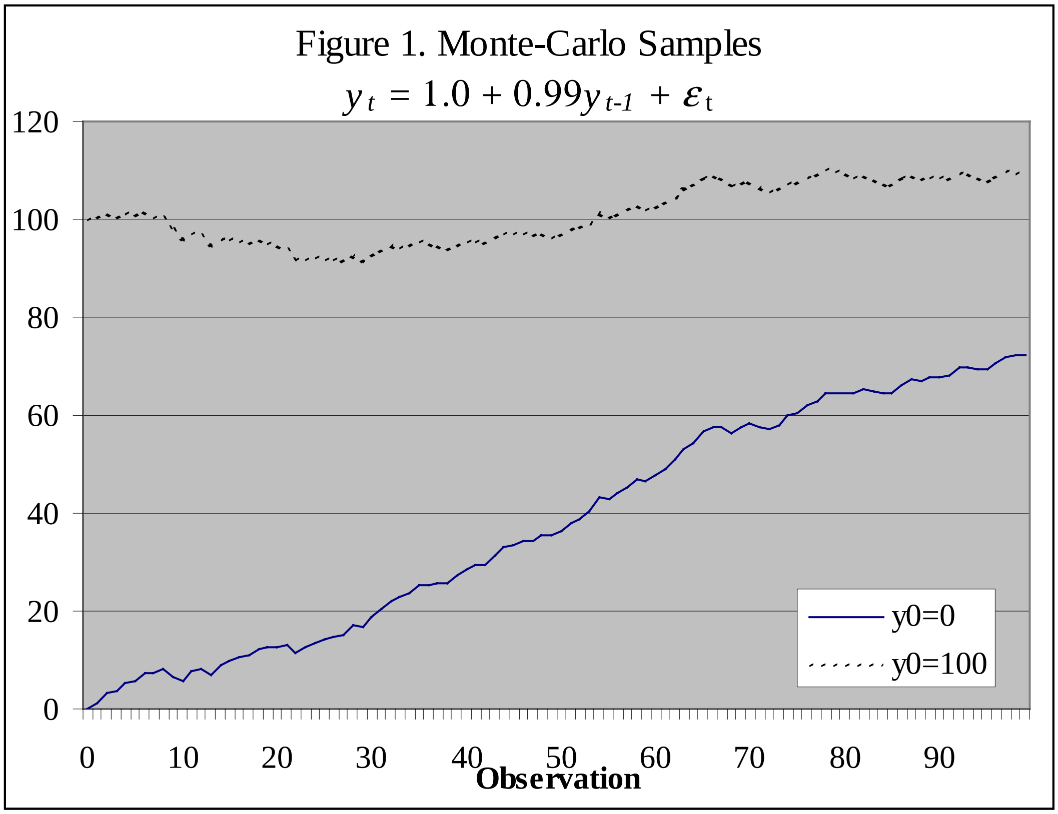
EK's anomalous size results led RT to perform a Monte Carlo study that investigates a different question than that examined by EK; they look at the size (and, despite their intentions, power) of the Dickey-Fuller Φ_2 and Φ_3 F tests rather than at the Φ_1 F test which is relevant when we know that there is no trend. (The Φ_1 test tests the joint hypothesis of a unit root and a zero intercept; the Φ_2 test tests the joint hypothesis of a unit root, a zero intercept, and a zero trend; the Φ_3 test tests the joint hypothesis of a unit root and a zero trend.) The null of the Φ_1 F test is that the intercept is zero and there is a unit root; the testing equation does not have a trend term because we know that there is no trend. The associated t test when we know there is no trend is denoted τ_μ. In contrast, the Φ_2 and Φ_3 F tests are used whenever we do not know that there is no trend; because of this the testing equation must include a trend. The associated t test in this case is denoted τ_τ, which has a different non-standard distribution than its counterpart τ_μ.

Except for their choice of starting value, RT generate their Monte Carlo data exactly as did EK, but then used a test equation including a trend term. This caused the EK anomalous size results to disappear, but produced power results that appear to contradict the EK results. It is this set of power results, along with ambiguous RT commentary, that will cause readers of the RT paper to conclude that the EK results are invalid. It turns out, however, that these differing power results were produced by RT's choice of starting value for generation of their Monte Carlo data.

3 Influence of the Starting Value

EK discuss the starting value problem at some length, noting that in samples of the size found in most empirical work (EK used a sample size of 100 in their Monte Carlo work, as did Dickey and Fuller), the starting value used for generation of Monte Carlo data can be of consequence. Indeed, a considerable literature has examined this issue recently, concluding that power results can be markedly affected by the initial condition; see, for example, Cook (2004) and Muller and Elliott (2003). One result from this literature is that choice of unit root test should bear in mind the distance of the starting value from the unconditional mean, should something be known about that distance. In general, of course, nothing is known in this regard, so this advice is of questionable value. But there is an important implication for Monte Carlo studies: The starting value should be chosen to reflect the context of the problem under investigation.

To illustrate the role of the starting value, we plot in Figure 1 two artificial samples generated by the stationary first-order autoregressive process $y_t = 1.0 + 0.99y_{t-1} + \varepsilon_t$ with $\varepsilon_t \sim N(0,1)$, where the starting value for one sample is $y_0=0$, and the starting value for the other sample is $y_0=100$, the unconditional mean of the process. Note that the sample with a starting value of zero displays strong evidence of a trend as the process increases steadily in search of its mean. Even after 100 iterations, the process does not reach its mean. The effect of such starting values can be made arbitrarily small by increasing the



sample size, but clearly, the effect is substantial in sample sizes often associated with Monte-Carlo studies. Because the context of the EK paper was no trend, we recommended using a starting value equal to the unconditional mean, or, preferably, its stochastic version. This ensured that there would be no trend in the data, allowing those data to represent the situation under investigation more accurately than would an arbitrary choice of starting value.

RT chose instead to use starting value $y_0=0$. In the context of their primary interest, namely the size of the unit root test, the starting value does not matter even in small samples, so their choice cannot be criticized. Unfortunately, RT report power results in addition to size results, evidently not recognizing that these power results are sensitive to the starting value. The artificial trend created by using starting value $y_0=0$ causes the trend term in the testing equation to play an active role when computing power. In contrast, if the starting value had been the unconditional mean, this artificial trend would not have been present, so the trend term in the testing equation would have played little role in computing power. We illustrate this in the Monte Carlo results reported below.

4 Monte Carlo Replication

We have replicated RT's Monte Carlo study using their starting value of $y_0=0$. The results, similar to those in RT, are reported in the bottom half of Table I. In the bottom half of Table II we report comparable results obtained using the stochastic version of a starting value equal to the unconditional mean of the process, what we consider to be the most appropriate way of selecting a starting value for the context of the issue EK were addressing. When the null of a unit root is true (i.e., when ρ equals one) the unconditional mean is undefined, so any starting value can be employed; we used $y_0=0$. The data have been generated, as in both EK and RT, using the equation $y_t = \alpha + \rho y_{t-1} + \varepsilon_t$. In the bottom half of both tables, for the tests Φ_2, Φ_3, and τ_τ, the testing equation is equivalent to RT's testing equation 8 (which adds a trend to the EK testing equation): $y_t = \alpha + \beta t + \rho y_{t-1} + u_t$. In the top half of both tables, for the tests Φ_1 and τ_μ., the testing equation does not have a trend: $y_t = \alpha + \rho y_{t-1} + u_t$. Several results are of note.

First, the powers are markedly different in Table I than in Table II, because of the artificial trend created by using the $y_0=0$ starting value. Although RT claim (p.4) that they "experimented with other starting values and obtained directly comparable results," we suspect that by this comment they are referring to their size results. Had they looked at their power figures, they should have seen a marked difference between those resulting from using starting values far from the unconditional mean and those resulting from using starting values close to the unconditional mean. One conclusion that could be drawn from these power differences is that the EK result depends on the starting value being close to the unconditional mean so that there is unequivocally no trend in the data, as emphasized in the EK paper. Because the Table I results do not reflect the context of the EK study, henceforth we disregard the Table I results and consider only the Table II results.

Second, in the top half of Table II are reported results that replicate the EK results, with na in place of the irrelevant results reported in EK. The main point of EK is evident from these results: The one-sided t test τ_μ is everywhere more powerful than the Φ_1 F test. These results also show, as EK claimed, that the magnitude of the intercept α does not affect power for any of these tests; in each row, for each ρ value the powers for a test are identical for different α values.

Third, in the bottom half of Table II are reported results from RT's Monte Carlo setup but using the stochastic version of the unconditional mean as the starting value. The EK results are maintained, but with all tests exhibiting power lower than those reported in the upper half of this table. This makes sense: Because under the alternative there is no trend, including an irrelevant trend in the testing equation should reduce power.

Finally, we note that the results reported in the bottom half of Table II should not be interpreted as showing that the EK conclusion applies in the presence of trending data. The Monte Carlo data used in the bottom half of Table II contain no trend except when both ρ equals one and α is nonzero; it is only the testing equation that contains a trend term, in accordance with the RT study.

Table I – Empirical Size and Power of 0.05 Tests for Sample Size 100 with $y_0 = 0$

Power and size are computed from 10,000 samples.

	$\rho = 0.8$			$\rho = 0.9$			$\rho = 0.95$			$\rho = 0.99$			$\rho = 1.0$		
Estimated Equation/	α			α			α			α			α		
Test Statistic	**0.00**	**0.50**	**1.00**	**0.00**	**0.50**	**1.00**	**0.00**	**0.50**	**1.00**	**0.00**	**0.50**	**1.00**	**0.00**	**0.50**	**1.00**
$y_t = \alpha + \rho y_{t-1} + \varepsilon_t$															
Φ_1	0.77	0.82	0.93	0.23	0.34	0.75	0.08	0.22	0.91	0.05	0.78	1.00	0.05	na	na
τ_μ one-sided	0.86	0.90	0.96	0.31	0.42	0.75	0.12	0.23	0.67	0.06	0.07	0.20	0.06	na	na
$y_t = \alpha + \beta t + \rho y_{t-1} + \varepsilon_t$															
Φ_2	0.40	0.47	0.63	0.09	0.13	0.39	0.04	0.10	0.62	0.04	0.67	1.00	0.05	0.97	1.00
Φ_3	0.55	0.60	0.75	0.14	0.19	0.43	0.06	0.11	0.40	0.05	0.07	0.15	0.05	0.05	0.05
τ_τ one-sided	0.64	0.68	0.78	0.18	0.22	0.33	0.08	0.09	0.14	0.05	0.05	0.04	0.05	0.05	0.05

Table II. Empirical Size and Power of 0.05 Tests for Sample Size 100 with $y_0 \sim N(\alpha/(1-\rho), \sigma^2/(1-\rho^2))$

Power and size are computed from 10,000 samples. The starting value when $\rho = 1.0$ is $y_0=0$.

	$\rho = 0.8$			$\rho = 0.9$			$\rho = 0.95$			$\rho = 0.99$			$\rho = 1.0$		
Estimated Equation/	α			α			α			α			α		
Test Statistic	**0.00**	**0.50**	**1.00**	**0.00**	**0.50**	**1.00**	**0.00**	**0.50**	**1.00**	**0.00**	**0.50**	**1.00**	**0.00**	**0.50**	**1.00**
$y_t = \alpha + \rho y_{t-1} + \varepsilon_t$															
Φ_1	0.79	0.79	0.79	0.25	0.25	0.25	0.09	0.09	0.09	0.05	0.05	0.05	0.05	na	na
τ_μ one-sided	0.87	0.87	0.87	0.33	0.33	0.33	0.13	0.13	0.13	0.06	0.06	0.06	0.05	na	na
$y_t = \alpha + \beta t + \rho y_{t-1} + \varepsilon_t$															
Φ_2	0.43	0.43	0.43	0.09	0.09	0.09	0.04	0.04	0.04	0.04	0.04	0.04	0.05	0.97	1.00
Φ_3	0.58	0.58	0.58	0.15	0.15	0.15	0.07	0.07	0.07	0.05	0.05	0.05	0.05	0.05	0.05
τ_τ one-sided	0.66	0.66	0.66	0.19	0.19	0.19	0.08	0.08	0.08	0.05	0.05	0.05	0.05	0.05	0.05

5 Conclusion

An oversight produced irrelevant columns in our Monte Carlo results reported in EK. This was an unfortunate lapse on our part, but, fortuitously, it did not affect our main conclusion: When it is known that a variable has no trend, the one-sided Dickey-Fuller t test τ_μ dominates the Dickey-Fuller Φ_1 F test.

Two positive results have come from RT's comment on our original EK paper. First, it has allowed us to stress that our conclusion favoring the one-sided t test applies only in a context in which it is known that there is no trend in the data. And second, it has forcefully illustrated the important role played by the starting value used to generate Monte Carlo data for time series analysis.

References

Cook, Steven (2004) “On the finite sample power of modified Dickey-Fuller tests: The role of the initial condition” *Economics Bulletin* **3(11)**, 1-9.

Dickey, David. A., and Wayne. A. Fuller (1981) “Likelihood ratio statistics for autoregressive time series with a unit root” *Econometrica* **49**, 1057-1072.

Elder, John, and Peter E. Kennedy (2001) “F versus t tests for unit roots.” *Economics Bulletin* **3(3)**, 1-6.

Muller, Ulrich K., and Graham Elliott (2003) Tests for unit roots and the initial condition *Econometrica* **71**, 1269-1286.

Rodrigues, Paulo, and Andrew Tremayne (2004) “F versus t tests for unit roots: a comment” *Economics Bulletin* **3(12)**, 1-7.