F versus t tests for unit roots

John Elder
North Dakota State University

Peter E. Kennedy
Simon Fraser University

Abstract

F tests which test jointly for a unit root and a zero intercept, and so compete against Dickey−Fuller t tests, are shown not to enhance power because they are invariant to the intercept value in the absence of a unit root. Monte Carlo results in the literature that indicate otherwise are shown to have resulted from the use of special starting values. Testing procedures that employ these F tests to enhance power should be revised.

1. Introduction

In their simplest form, tests for unit roots are introduced in the context of

$$y_t = \rho \, y_{t-1} + \alpha + \varepsilon_t \tag{1}$$

where ε_t is an error term with mean zero and variance σ^2, and the time series y_t is said to have a unit root if $\rho = 1$. Judging by textbook expositions, and comments such as that of Ayat and Burridge (2000, p.74), the augmented Dickey-Fuller (ADF) test has become the most popular of many competing tests in the literature. In its original, unaugmented form, known as the DF test, this test is undertaken by estimating equation (1) and conducting a traditional one-sided t test of $\rho = 1$, using special critical values.[1]

The seminal paper of Dickey and Fuller (1981) introduced a competing F test, Φ_1, based on testing the joint null $(\rho, \alpha) = (1, 0)$ with the usual F statistic but using special critical values. Equation (1) is appropriate whenever y_t is known not to have a trend, in which case the null of a unit root implies that the intercept or "drift" term α must be zero. This suggests that Φ_1, which looks at both ρ and α, may have more power than the competing t test. Dickey and Fuller (1981, p.1069) state unequivocally that their t test is generally less powerful than Φ_1 when ρ is less than one. Hatanaka (1996, pp.49-50) also has an unequivocal statement: "The t-statistic for $\rho = 1$ can also be used……. but the F-statistic is more desirable." Hamilton's (1994, p.494) endorsement of Φ_1 is less direct: "Thus, it might seem more natural to test for a unit root in this specification by testing the joint hypothesis that $\alpha = 0$ and $\rho = 1$." Several other references in the literature, such as Perron (1988, p.317), Holden and Perman (1994, pp.64-65), Enders (1995, p.257), and Ayat and Burridge (2000, p.80) can be interpreted as implicitly endorsing Φ_1 as being more powerful than the t test. To our knowledge there are no statements to the contrary in the literature.

The purpose of this note is to point out that this view in the literature is incorrect, and that Monte Carlo results in Dickey and Fuller (1981) showing otherwise are misleading because they result from the use of special starting values.[2] A consequence of this is that unit root testing strategies that make use of this test, such as those of Perron (1988), Holden and Perman (1994), and Enders (1995) should be revised, and recommendations favoring this F test should be ignored.

2. F Test Redundancy

The result just noted stems from a peculiar feature of unit root F tests in the absence of a unit root, apparently unnoticed in the literature. Since this peculiarity is common to both DF and ADF tests, for simplicity we confine our discussion to the DF test. Φ_1 tests the joint null $(\rho, \alpha) = (1, 0)$ in equation (1) when we know that y_t has no trend. The alternative is $\rho < 1$ with any value for α, although in most realistic cases α should be

---

[1] It is more common to find this test undertaken by regressing the first difference of y_t on y_{t-1} and testing the coefficient on y_{t-1} against zero. Since this transformation is of no consequence for what follows, for simplicity we ignore it.

[2] A similar result holds for Dickey and Fuller's Φ_2 test as a supplement to their Φ_3 test. We do not exposit this result here because the Φ_2 test is not in common use.

positive. The t test is one-sided, whereas the F test is not, so that for small values of α we would expect the t test to have greater power. For large values of α, however, we might expect the F test to have larger power because it attends to the value of α. Unfortunately, this is incorrect because it happens that when there is no unit root (i.e., when we are examining power), the value of α has no influence on the F statistic.

This result can most easily be demonstrated by showing that the restricted and unrestricted sums of squared residuals used to calculate the F statistic are unaffected by the intercept, so long as the starting value y_0 is equal to its unconditional mean $\alpha_0/(1-\rho_0)$, where α_0 and ρ_0 are the true values of α and ρ, respectively. (The role of the starting value is discussed further below.) The AR(1) process can be written as

$$y_t = \alpha \sum_{\tau=1}^{t} \rho^{\tau-1} + \sum_{\tau=1}^{t} \rho^{t-\tau} \varepsilon_\tau + \rho^t y_0 \qquad \text{for } t = 1, 2, 3, \dots \tag{2}$$

Imposing the constraints that $y_0 = \alpha_0/(1-\rho_0)$, and $\rho_0 \neq 1$, this can be rewritten as

$$y_t = \frac{\alpha_0}{1-\rho_0} + \sum_{\tau=0}^{t} \rho_0^{t-\tau} \varepsilon_\tau \qquad \text{for } t = 0, 1, 2, \dots \tag{3}$$

where $\varepsilon_0 = 0$ by definition because of the way in which y_0 is chosen. (Note that equation (3) describes the full sequence, including y_0.) From this equation it is clear that $y_t - y_{t-1}$, and thus the restricted sum of squared residuals used in calculating the F test of the hypothesis that $(\rho, \alpha) = (1, 0)$, is not affected by the intercept value α_0.

Consider now the unrestricted sum of squared residuals for this F test, with the residuals denoted $y_t - \hat{\alpha} - \hat{\rho} y_{t-1}$. If we change the true value of α from α_0 to α^*, and y_0 to its unconditional mean $\alpha^*/(1-\rho_0)$, then the residuals become $y_t^* - \hat{\alpha}^* - \hat{\rho}^* y_{t-1}^*$. From (3) and the properties of the OLS estimator

$$y_t^* = y_t + \frac{\alpha^* - \alpha_0}{1-\rho_0} \quad \text{for} \quad t = 0, 1, 2, \dots \tag{4}$$

$$\hat{\alpha}^* = \hat{\alpha} + \frac{\alpha^* - \alpha_0}{1-\rho_0}(1-\hat{\rho}) \quad \text{and} \tag{5}$$

$$\hat{\rho}^* = \hat{\rho}. \tag{6}$$

Substituting (4) through (6) into $y_t^* - \hat{\alpha}^* - \hat{\rho}^* y_{t-1}^*$ produces $y_t - \hat{\alpha} - \hat{\rho} y_{t-1}$, so the unrestricted sum of squared errors is unaffected by the intercept value.

All this implies that, in the absence of a unit root, the true value of α affects neither the restricted nor the unrestricted sum of squared errors, and thus does not affect the F statistic – with unconditional mean starting values, the calculated F value is numerically identical for different α values. This implies that there is no gain in power[3] from using Φ_1: the t test dominates.

---

[3] One reason this result has gone unnoticed in the literature is that under the null of a unit root, where the profession has directed its attention, the intercept does affect the F value. It is only when examining power

3. Starting Value Influence

To examine the influence of the starting value, substitute $y_0 = \delta + \alpha_0/(1-\rho_0)$ in equation (2), where δ is the extent to which the starting value differs from its unconditional mean, and calculate the comparable equation (4) result, obtaining

$$y_t^* = y_t + \frac{\alpha^* - \alpha_0}{1 - \rho_0} + \rho_0^{\ t}\delta \quad \text{for} \quad t = 0, 1, 2, \ldots \tag{7}$$

This result illustrates the general nature of the impact of the starting value – if it differs from its unconditional mean the results we derived earlier will not follow – y_t^* and the other results are slightly different, implying that the sums of squared residuals are no longer unchanged when the intercept changes. Equation (7) suggests that this difference will disappear asymptotically (because $\rho_0^{\ t}\delta$ disappears asymptotically) and indeed straightforward but tedious algebra shows that this is the case. Intuitively, the process starts at a value considerably different from its unconditional mean, but being stable, works its way into a neighborhood of its unconditional mean. For a time, however, it produces unrepresentative observations, affecting the F statistic. If δ is related to the intercept, the intercept could in this way affect the power of the F statistic. This phenomenon helps explain seemingly anomalous Monte Carlo results in the literature, as explained in the next section.

4. Monte Carlo Contradiction

One reason why the profession has failed to notice this peculiarity of F tests in the unit root context is that Dickey and Fuller (1981, p.1067) in their Table VII present Monte Carlo results indicating that the Φ_1 test gains power as the intercept α grows. As explained below, these results have come about because they held constant the initiating y value at zero as they moved from $\alpha = 0$ to $\alpha = 0.5$ to $\alpha = 1.0$. This causes the extent to which the process begins away from its unconditional mean to vary as the intercept value varies. The Dickey-Fuller Monte Carlo power results must be interpreted as showing that the power of the F test increases as the distance of the process starting value from its unconditional mean increases, rather than, as is implicitly suggested in their paper, increasing as the intercept increases.

Consider their case in which $\rho = 0.9$. When $\alpha = 0$ the equilibrium (unconditional mean) value of y is zero, so the starting y value should be zero if fixed and have mean zero if random; when $\alpha = 0.5$, the equilibrium value of y is 5, so the starting y value should be 5 if fixed and have mean 5 if random; and when $\alpha = 1.0$ the equilibrium value of y is 10, so the starting y value should be 10 if fixed and have mean 10 if random. But in the Dickey-Fuller Monte Carlo study, when $\alpha = 0.5$ the process always began 5 units below its conditional mean, and when $\alpha = 1.0$ it always began 10 units below its conditional mean. This resulted in a major string of outliers, always in the same direction in repeated samples, biasing the Monte Carlo results. This bias disappears asymptotically because the y values eventually move to a range consistent with their equilibrium

---

(i.e., when $\rho < 1$) that this result comes into play. It should also be clear that this result does not hold when testing a null other than a unit root.

(unconditional mean) value, but in finite samples (of size 100, say, as employed by Dickey and Fuller) the bias is enough to produce markedly misleading results.

Our observation that the power of the Φ_1 test is not affected by the intercept is not entirely new to the literature. Dickey (1984) reports results of an extensive Monte Carlo study which investigated the powers of this and other tests.[4] A careful reading of his clever summary in Table 6.1 (p.491) reveals that the intercept value does not affect the power of Φ_1. Unfortunately, Dickey did not report that these results contradicted the Dickey and Fuller (1981) results, and did not warn the profession about the implications of his results for the use of these tests, perhaps because in 1984 it was not yet evident that the profession would so enthusiastically adopt these tests.

We began our investigation of this by replicating the Dickey-Fuller Monte Carlo results, in which data were generated by equation (1), using a fixed starting value of $y_0 = 0$, regardless of the intercept. Our results, shown in Table I, differ minimally from the Dickey-Fuller Table VII results. We have omitted columns for which $\rho > 1$, and five rows corresponding to other F statistics and variants of the DF t statistic. The second-last row refers to the traditional DF t statistic, two sided, and the last row, added by us, is the traditional one-sided DF t statistic. This last row is reported by Dickey and Fuller in their Table IX (p.1068).

Next, we replicated their results by using a fixed starting value equal to the unconditional mean, namely $y_0 = \alpha/(1-\rho)$, so that it varies with the intercept. The results, shown in Table II, were as we expected – when $\rho < 1$, power was unaffected as the intercept varied. The Dickey-Fuller power values in Table I for $\alpha = 0$ match those of Table II, of course, but when $\alpha \neq 0$ they are very different. Also of note is that the DF t test powers are no longer affected by the intercept. This makes sense – the t test should be unaffected by the intercept value, so its power should also be unaffected.

Finally, we replicated their results using what we regard as the most relevant starting value methodology, a random starting value. Here y_0 was chosen randomly with mean $\alpha/(1-\rho)$, the unconditional mean of y, and variance $\sigma^2/(1-\rho^2)$, the unconditional variance of y. The results, shown in Table III, differ from those in Table II only insofar as power is a bit higher. This is consistent with our earlier discussion of the influence of starting values.

The main point of this paper is that for testing $\rho = 1$ the Φ_1 test is dominated by the (one-sided) DF t test. This is evident in Table III where the estimated powers for Φ_1 are uniformly lower than those of the traditional one-sided DF t test, in contrast to the Dickey-Fuller results reported in Table I. The Dickey-Fuller results are only relevant to a situation in which a researcher knows that in the recent past the value of y was considerably different from its unconditional mean, by an amount affected by the true intercept value. An example would be if we knew, regardless of the intercept value, that the process started recently at zero. This does not to us seem representative of actual applications.

---

[4] We are greatly indebted to David Dickey for drawing this paper to our attention, and for his willingness to help us investigate this issue in a scholarly fashion.

5. Conclusions

The main conclusion of this paper is that the F test Φ_1 should be abandoned in favor of the Dickey-Fuller t test, contradicting recommendations to the contrary in the literature. This should simplify unit root testing strategies, such as those of Perron (1988), Holden and Perman (1994), and Enders (1995), that employ this test in the mistaken belief that it will enhance power.

These results suggest an interesting avenue for future research. If under the null of a unit root the t and Φ_1 tests (adjusted for degrees of freedom) are not equivalent, but are equivalent, or nearly so, under the alternative, perhaps this equivalency could be exploited to develop a unit root test with a null of stationarity.

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

Statistic	$\rho = 0.8$ α			$\rho = 0.9$ α			$\rho = 0.95$ α			$\rho = 0.99$ α			$\rho = 1.0$ α		
	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
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	0.98	1.00
$\hat{\tau}_\mu$ two-sided	0.71	0.76	0.88	0.18	0.27	0.58	0.06	0.13	0.51	0.04	0.08	0.14	0.05	0.27	0.34
$\hat{\tau}_\mu$ 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	0.01	0.00

Power is computed from 10,000 samples

Table II
Empirical Power of 0.05 Tests for Sample Size 100 ($y_0 = \alpha/(1-\rho)$)

Statistic	$\rho = 0.8$ α			$\rho = 0.9$ α			$\rho = 0.95$ α			$\rho = 0.99$ α			$\rho = 1.0$ α		
	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
Φ_1	0.77	0.77	0.77	0.23	0.23	0.23	0.08	0.08	0.08	0.05	0.05	0.05	0.05	0.98	1.00
$\hat{\tau}_\mu$ two-sided	0.71	0.71	0.71	0.18	0.18	0.18	0.06	0.06	0.06	0.04	0.04	0.04	0.05	0.27	0.34
$\hat{\tau}_\mu$ one-sided	0.86	0.86	0.86	0.31	0.31	0.31	0.12	0.12	0.12	0.06	0.06	0.06	0.06	0.01	0.00

Power is computed from 10,000 samples

Table III
Empirical Power of 0.05 Tests for Sample Size 100 [$y_0 \sim N(\alpha/(1-\rho), \sigma^2/(1-\rho^2))$]

Statistic	$\rho = 0.8$ α			$\rho = 0.9$ α			$\rho = 0.95$ α			$\rho = 0.99$ α			$\rho = 1.0$ α		
	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
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	0.98	1.00
$\hat{\tau}_\mu$ two-sided	0.73	0.73	0.73	0.20	0.20	0.20	0.07	0.07	0.07	0.05	0.05	0.05	0.05	0.27	0.34
$\hat{\tau}_\mu$ 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	0.00	0.00

Power is computed from 10,000 samples

References

Ayat, L., and P. Burridge (2000) "Unit Root Tests in the Presence of uncertainty about the Non-stochastic Trend" *Journal of Econometrics* **95**, 71-96.

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

Dickey, D. A. (1984) "Powers of Unit Root Tests" *Proceedings of Business and Economic Statistics Section, American Statistical Association*, 489-493.

Enders, W. (1995*) Applied Econometric Time Series*, Wiley: New York.

Hamilton, J. D. (1994) *Time Series Analysis*, Princeton University Press: Princeton.

Hatanaka, M. (1996) *Time-Series-Based Econometrics: Unit Roots and Cointegration*, Oxford University Press: Oxford.

Holden, D., and R. Perman (1994) "Unit Roots and Cointegration for the Economist" in *Cointegration for the Applied Economist* by B. Bhaskara Rao, Ed., St. Martin's Press: New York, 47-112.

Perron, P. (1988) "Trends and Random Walks in Macroeconomic Time Series" *Journal of Economic Dynamics and Control* **12**, 297-332.