ESTIMATING FINITE SAMPLE CRITICAL VALUES FOR UNIT ROOT TESTS USING PURE RANDOM WALK PROCESSES: A NOTE

By Yin-Wong Cheung and Kon S. Lai

University of California at Santa Cruz and California State University at Los Angeles

First version received November 1993

Abstract. Finite sample critical values currently available for the augmented Dickey–Fuller test are commonly obtained via simulations using ARIMA (0, 1, 0) processes. An implicit but critical assumption is that the possible presence of nuisance nonunit roots in general processes does not affect the finite sample size property of the test. The validity of this assumption, though always presumed, has not been verified. This study shows that the use of ARIMA (0, 1, 0) processes for computing the critical values is not so restrictive as it may seem. By estimating empirical size curves as a function of nuisance root parameters, results of Monte Carlo analysis suggest that the empirical test size is not sensitive to nuisance autoregressive (AR) and moving-average (MA) roots over a wide range of their values, except only when the AR or MA root is near unity. The results support, though not unqualifiedly, the reliability and usefulness of finite sample critical values estimated based on simple ARIMA (0, 1, 0) processes.

Keywords. Augmented Dickey–Fuller test; empirical size; nuisance root parameter; Monte Carlo.

1. INTRODUCTION

The augmented Dickey–Fuller (ADF) test has been widely used to test for unit roots in time series. Consider a series $\{x_t\}$ generated by the following data generating process (DGP):

$$(1 - \rho L)x_t = u_t$$

$$\prod_{i=1}^{p}(1 - \phi_i L)u_t = \prod_{j=1}^{q}(1 - \theta_j L)e_t \qquad t = 1, 2, \ldots, T \tag{1}$$

where L is the lag operator; ρ is the largest autoregressive (AR) root; ϕ_i and θ_j are the respective roots of the AR and moving-average (MA) lag polynomials describing the innovation process u_t, with $|\phi_i| < 1$ and $|\theta_j| < 1$; e_t is independent and identically distributed (i.i.d.) $N(0, \sigma^2)$; and the AR and MA parts share no common root. A unit root exists when $\rho = 1$. Let $\Delta = 1 - L$. Using AR(k) as an approximating model, the ADF test for the hypothesis H_0: $\rho = 1$ is conducted by regressing Δx_t on $(1, x_{t-1}, \Delta x_{t-1}, \ldots, \Delta x_{t-k+1})$ and examining the negativity of the coefficient on x_{t-1} based on its regression t ratio τ_k.

0143-9782/95/05 493–498 JOURNAL OF TIME SERIES ANALYSIS Vol. 16, No. 5

Under H_0: $\rho = 1$, the asymptotic distribution of the ADF statistic is given by

$$\tau_k \Rightarrow \left\{\int_0^1 W(r)\,dW(r)\right\}\left\{\int_0^1 W(r)^2\,dr\right\}^{-1/2} \tag{2}$$

where $W(r)$ is a standard Brownian motion over the [0, 1] interval (see Dickey and Fuller, 1979). This distribution is invariant with respect to k (the lag order) and ϕ_i and θ_j (the nonunit root parameters). Since direct derivation of the finite sample distribution is not tractable, finite sample critical values are typically estimated via simulations. Although the independence with respect to k, ϕ_i and θ_j is an asymptotic result only, its validity in finite samples – with which empirical applications always deal—has not been verified but presumed. For example, MacKinnon (1991) provides response surface estimates of finite sample critical values for the ADF test with a fixed $k = 1$ and DGPs given by ARIMA (0, 1, 0) processes. More general DGPs, for which u_t is serially correlated, will contain additional nuisance parameters, ϕ_i and θ_j. Cheung and Lai (1995) extend MacKinnon's (1991) analysis by allowing for the effect of k but still omits these other nuisance parameters. Such omission will not be appropriate if ϕ_i and θ_j can systematically affect the empirical size of the ADF test. This study examines the potential sensitivity of the empirical test size to nonunit roots.

2. MONTE CARLO EXPERIMENTS AND RESULTS

To allow for possibly different effects of the AR and MA roots on the ADF test, two DGPs are considered under H_0: $\rho = 1$. They are ARIMA (1, 1, 0) and ARIMA (0, 1, 1) processes. Following Equation (1), let $\phi = \phi_1$ and $\theta = \theta_1$. The experimental design covers different possible combinations of (k, T, ϕ, θ) with $k = \{2, 4, 6, 8, 10, 12\}$, $T = \{50, 100, 250\}$, $\phi = \{0.95, 0.90, \ldots, -0.95\}$ and $\theta = \{0.95, 0.90, \ldots, -0.95\}$. ADF tests with and without a time trend are conducted. For each given (k, T) combination, finite sample critical values of the 5% and 10% tests are computed using the response surfaces estimated by Cheung and Lai (1995), thereby correcting for the effects of both k and T parameters. In this way, more accurate evaluation of the effects of nuisance AR and MA parameters on the test size can be made by separating them from those of k and T. The design here yields a total of 2808 possible combinations of simulation experiments, each of which is based on 40000 replications.

The Monte Carlo results are summarized using graphs, which can efficiently present a vast amount of data information in a compact space and help reveal general overall patterns. Following Tufte's (1983) idea of 'small multiples,' multiple graphs are arranged in series to facilitate comparison across k, T, ϕ and θ parameters. The results are organized into groups for different ADF tests: with a time trend (Figure 1) and without time trend (Figure 2). Each

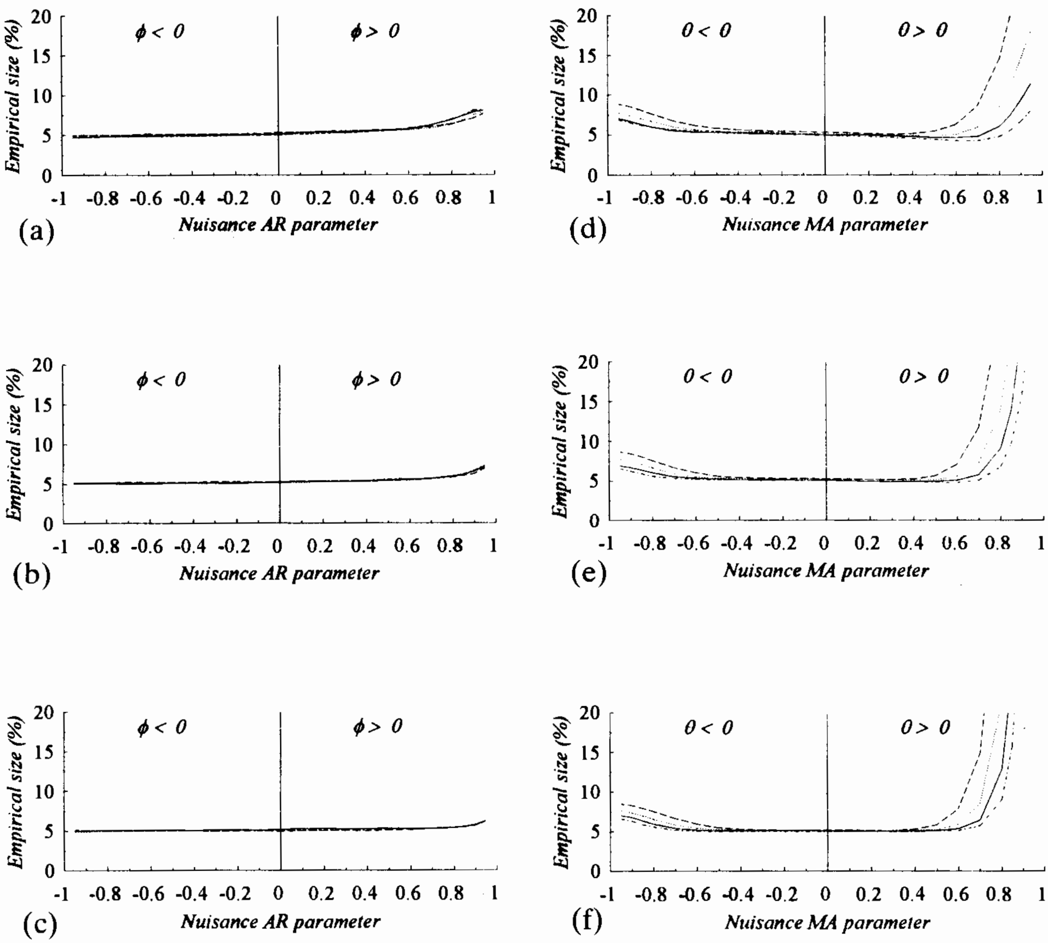

FIGURE 1. Empirical size curves for the ADF test with a time trend: (a) $T = 50$; (b) $T = 100$; (c) $T = 250$; (d) $T = 50$; (e) $T = 100$; (f) $T = 250$. (a)–(c): – – , $k = 2$; · · · ·, $k = 4$; —, $k = 6$; –·–·, $k = 8$. (d)–(f): – – , $k = 6$; · · · ·, $k = 8$; —, $k = 10$; –·–·, $k = 12$.

group consists of a 3×2 matrix of graphs, showing how the empirical test size of the ADF test with various lag choices changes as a function of the nuisance AR or MA root. The results for the 5% test are reported below (similar results were obtained for the 10% test).

A number of interesting results can be observed from Figures 1 and 2. Consider first the sensitivity of the test size to the nuisance AR parameter. The empirical size curves are almost entirely flat at the 5% level – which is the nominal size of the test – indicating that the nonunit AR parameter causes little size distortion in general. The empirical test size may deviate from its nominal level only when the value of the AR root is large and close to unity, a situation in which the DGP is near to having two unit roots. These findings are robust with respect to the type of the ADF test as well as the lag order parameter. The robustness with respect to the latter reflects and confirms the accuracy of the lag-adjusted critical values used.

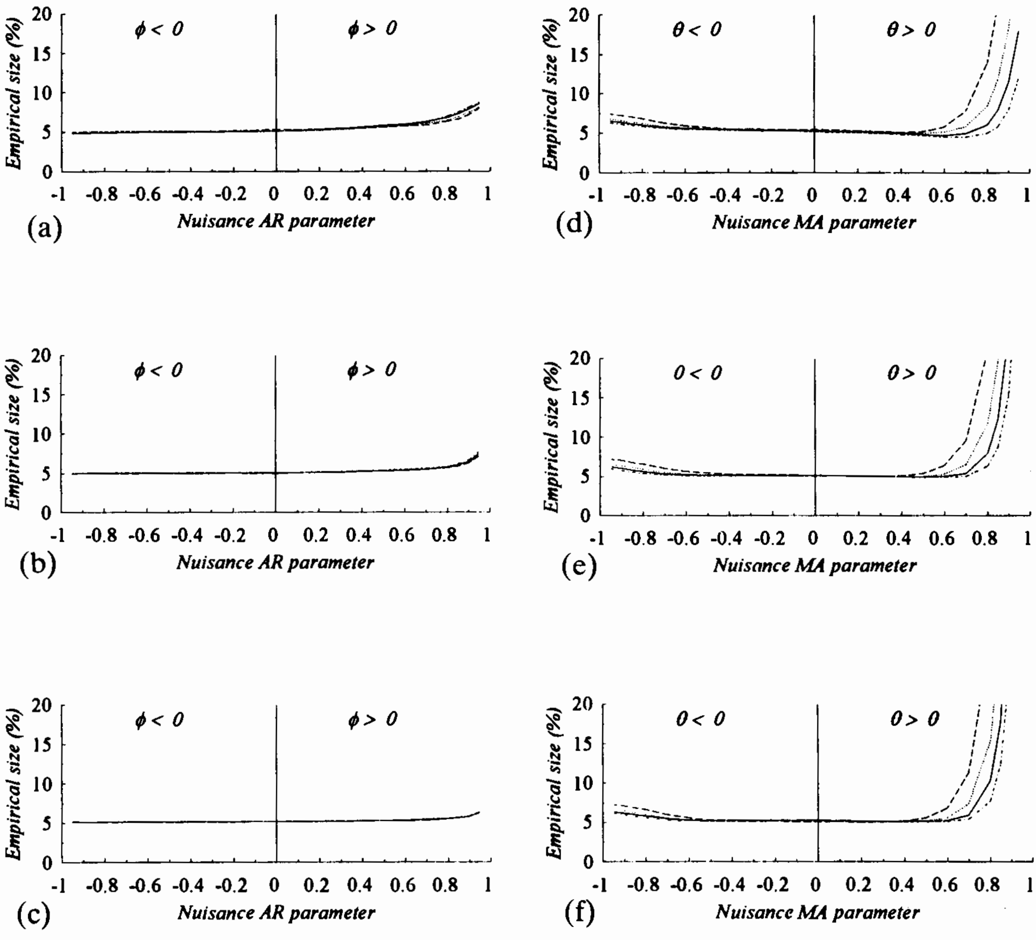

FIGURE 2. Empirical size curves for the ADF test without a time trend: (a) $T = 50$; (b) $T = 100$; (c) $T = 250$; (d) $T = 50$; (e) $T = 100$; (f) $T = 250$. (a)–(c): – –, $k = 2$; · · · ·, $k = 4$; —, $k = 6$; – · – ·, $k = 8$. (d)–(f): – –, $k = 6$; · · · ·, $k = 8$; —, $k = 10$; – · – ·, $k = 12$.

Consider next the sensitivity to the nuisance MA parameter. The empirical size curves remain flat at the 5% level over a broad range of values of the MA root. The range of values over which the empirical size curve is flat appears wider for tests with a higher lag order than those with a lower lag order. This can be explained by the fact that a low AR lag order provides poor approximation to strong MA dependence, resulting in substantial size distortion (see Agiakloglou and Newbold (1992) for the power implications of using different lag orders). Moreover, the size distortion can be particular serious when the value of the MA root is large and positive, a situation in which some root-cancellation effects may be at work. Schwert (1989) observes similar size distortion for DGPs with a large and positive MA root. Compared with Schwert's, our results are actually stronger, given the adjustments here for possible finite sample bias and the lag order effect. In addition, the empirical size curve reported here can provide a clear sense of the potential size distortion as a function of different ranges of the MA root values, instead of just a few parameter values.

The results as a whole show that the empirical size of the ADF test is not sensitive to nuisance AR and MA parameters over a rather wide range of their values, provided that a sufficiently large lag order is employed in the test to capture dependence – the usual condition required for applying the ADF test. It follows that the use of simple random walk processes can still yield reasonable reliable and accurate estimates of finite sample critical values.

3. CONCLUSION

An issue concerning empirical applications of the ADF test has been examined. Finite sample critical values available in the literature for the test are commonly obtained through simulations using ARIMA (0, 1, 0) processes. Such processes appear restrictive in most practical situations. The reliability and usefulness of these finite sample critical values relies on an implicit but critical assumption that the presence of nonunit roots in more general processes does not affect the size property of the test. The validity of this assumption is always presumed in applied work but has not been systematically verified. If the empirical test size is actually sensitive to nuisance nonunit roots, the finite sample critical values currently available will be of limited use, and their lack of general applicability will make practical uses of the ADF test difficult. If the sensitivity is significant, a new set of critical values will have to be computed for every different DGP; this in turn will require accurate knowledge of the exact parametric specification of each given DGP.

This study finds that the use of ARIMA (0, 1, 0) processes is not so restrictive as it may seem. The sensitivity of the empirical size to nuisance nonunit roots is evaluated using the Monte Carlo method, through which empirical size curves as a function of nuisance root parameters are estimated. It is found that the empirical test size is not sensitive to nuisance AR and MA parameters over a wide range of their values, except only when either root is positive and near unity. The results provide support, though not an entirely unqualified one, for the reliability and usefulness of finite sample critical values estimated based on simple ARIMA (0, 1, 0) processes.

REFERENCES

Agiakloglou, C. and Newbold, P. (1992) Empirical evidence on Dickey–Fuller type tests. *J. Time Ser. Anal.* 13, 471–83.

Cheung, Y.-W. and Lai, K. S. (1995) Lag order and critical values of the Augmented Dickey–Fuller test. *J. Bus. Econ. Statist.*, forthcoming.

Dickey, D. A. and Fuller, W. A. (1979) Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Statist. Assoc.* 74, 427–31.

MacKinnon, J. G. (1991) Critical values for cointegration test. In *Long-Run Economic Relationships: Readings in Cointegration*, eds. R. F. Engle and C. W. J. Granger, New York: Oxford University Press.

Schwert, G. W. (1989) Tests for unit roots: A Monte Carlo investigation. *J. Bus. Econ. Statist.* 7, 147–59.

Tufte, E. (1983) *The Visual Display of Quantitative Information*. Cheshire, CT: Graphics Press.