Granger Centre
for time series econometrics

Granger Centre Discussion Paper Series

Testing for unit roots in the presence of uncertainty over both the trend and initial condition

by

David I. Harvey, Stephen J. Leybourne and A. M. Robert Taylor

---

Granger Centre Discussion Paper No. 08/03

The University of Nottingham

Testing for Unit Roots in the Presence of Uncertainty over Both the Trend and Initial Condition*

Prepared for the Conference in Honour of Peter C. B. Phillips

SMU, July 14-15 2008.

David I. Harvey, Stephen J. Leybourne and A. M. Robert Taylor
School of Economics and Granger Centre for Time Series Econometrics
University of Nottingham

May 2008

Abstract

We provide a joint treatment of two major problems that surround testing for a unit root in practice, namely uncertainty as to whether or not a linear deterministic trend is present in the data, and uncertainty as to whether the initial condition of the process is (asymptotically) negligible or not. In earlier work [Harvey, Leybourne and Taylor, 2008] we proposed methods to deal with trend uncertainty when the initial condition is assumed to be (asymptotically) negligible, together with methods to deal with uncertainty over the initial condition when the form of the trend function was taken as known. In each case we recommended a simple union of rejections-based decision rule. In the first case rejecting the unit root null whenever either of the quasi-differenced (QD) detrended or QD demeaned augmented Dickey-Fuller [ADF] unit root tests yields a rejection, and in the second case if either of the QD and OLS detrended/demeaned ADF tests rejects. Both approaches were shown to work well. In this paper we extend these procedures to allow for both trend and initial condition uncertainty, proposing a four-way union of rejections decision rule based on the QD and OLS demeaned and the QD and OLS detrended ADF tests. This is shown to work well but to lack power, relative to the best available test, in some scenarios. A modification of the basic union, based on auxiliary information including linear trend pre-test statistics, is proposed and shown to deliver significant improvements. A by-product of our analysis is that the power functions of the associated trend function pre-tests are shown to be heavily dependent on the initial condition.

Keywords: Unit root test; trend uncertainty; initial condition uncertainty; asymptotic power; union of rejections decision rule; trend tests.
JEL Classification: C22.

*Correspondence to: Robert Taylor, School of Economics, University of Nottingham, University Park, Nottingham NG7 2RD, U.K. Email: robert.taylor@nottingham.ac.uk

1 Introduction

The question of whether or not a time series contains a unit autoregressive root is a long standing one in time-series econometrics. Peter Phillips has made wide-ranging and essential contributions to this literature. Phillips (1987a) was a ground-breaking paper in this literature and was one of the very first papers to introduce the tool of Wiener algebra to econometricians. The quality of the exposition of that paper is such that even today, over twenty years on from its publication, it remains in our view one of the key references in the unit root literature. It is impossible to do justice to the contribution Peter Phillips has made to this literature, but here is small selection of the areas where he has made seminal contributions: local-to-unity asymptotics in Phillips (1987b); the role of the deterministic trend function in unit root testing in, for example, Phillips and Perron (1988) and Phillips (1998); optimal unit root testing in Schmidt and Phillips (1992); Bayesian approaches to unit root testing and issues of model selection in, for example, Phillips (1991a,1991b) and Phillips and Ploberger (1994); unit root testing with infinite variance errors in Phillips (1990); stationarity tests in Kwiatkowski *et al.* (1992); frequency domain approaches to unit root testing in Choi and Phillips (1993); panel unit root testing in, for example, Phillips and Moon (1999); moderate deviations from unity in Phillips and Magdalinos (2007); and with Zhijie Xiao a definitive and comprehensive review of the unit root literature, Phillips and Xiao (1998). The preceding list is not meant to be exhaustive, nor does it include Peter Phillips' important contributions to other areas of econometrics such as (in no particular order): co-integration; exact distribution theory; spurious regressions; continuous-time models; asymptotic distribution theory; approximation theory; instrumental variables and GMM; empirical finance; Bayesian methods; probabalistic theory; non-linear models; fractional integration and long memory; automated methods; inference under heteroskedasticity and HAC estimation, to name but a few.

For economic data, the autoregressive time series process of interest is generally not considered to be observed directly, but is instead assumed to be observed subject to some additive deterministic component. The unit root question cannot be properly analyzed until some characterization of the underlying deterministic component is made, since good practice dictates we should apply a unit root test that yields inference not dependent on whether or not a particular deterministic component is present. In a recent paper, Harvey, Leybourne and Taylor (2008) (HLT) consider this empirically important issue. They address the problem of whether to include a constant or constant and linear trend in a unit root test regression, citing this as the most common form of uncertainty regarding the deterministic component where macroeconomic data are concerned.

Assuming an asymptotically irrelevant initial condition (the deviation of the first observation from the deterministic component), it is well-known that the quasi-difference (QD) demeaned and detrended augmented Dickey-Fuller (ADF) unit root tests of Elliott *et al.* (1996) are efficient relative to their ordinary least squares (OLS) demeaned and detrended counterparts. On this basis HLT examine a strategy based on the union of rejections of the QD demeaned and detrended tests. The union principle exploits the fact that when a trend is absent, both QD tests are correctly sized under the unit root null but under the (locally) stationary alternative the demeaned test is the more likely of the two to signal a rejection of the unit root null in favour of stationarity (around a mean) since its power is not compromised by the inclusion of an irrelevant trend term. When a trend is present, the demeaned test becomes undersized and has trivially low power, to the extent that it is unlikely ever to reject the unit root null. Unit root inference then essentially becomes contingent on the detrended test alone, whose size and power to reject the unit root null in favour of stationarity (around a trend) are unchanged due to its invariance to a trend. Despite its simplicity, this union of rejections procedure is shown to be generally at least as powerful as competing procedures which involve some form of pre-testing for the presence of the trend term as a method to select between the QD demeaned or detrended unit root tests.

A second issue that impacts significantly on the power of unit root tests is the behaviour of the initial condition, as examined by Elliott and Müller (2006). Examination of this issue also stems from empirical considerations. A mean-reverting process with a large (in the absolute sense) initial condition could be used to characterize economic data that just happens to be observed directly after some structural episode, such as a policy shift or political regime change. Conversely, a more modest initial condition might be associated with data observed within a period of comparative economic stability. HLT consider the effects of such initial condition uncertainty, this time for an assumed deterministic specification (that is, either constant or constant and linear trend). While the initial condition plays no role when the unit root null is true, under the stationary alternative, if the initial condition is not asymptotically negligible, the QD tests have powers which decline monotonically towards zero as the absolute value of the initial condition increases, while those of the OLS tests rise monotonically. HLT therefore suggest using a union of rejections of the QD and OLS tests (either both demeaned or both detrended). The union exploits the superior power of the QD tests over the OLS tests for small initial conditions, and simultaneously exploits the reverse relationship for larger initial conditions. Again, this simple procedure is shown to perform well in comparison to other possible procedures.

In their analyses HLT therefore abstract away from any initial condition uncertainty when examining uncertainty regarding the trend component and, equally, abstract away from any trend uncertainty when questioning uncertainty over the initial condition. Since it would be difficult to argue on either theoretical or empirical grounds that these forms of uncertainty should exist in isolation, in this paper we explore a joint treatment of uncertainty of both forms, following up on a suggestion for further research made by HLT. To this end, in section 3 below we begin by describing the asymptotic behaviour of the QD demeaned and detrended unit root tests together with their OLS demeaned and detrended counterparts when both forms of uncertainty can arise. This allows us to examine interactive behaviour not previously considered.

On the basis of our findings, in section 4 we follow up on the suggestion made in HLT's rejoinder to consider a unit root testing strategy involving a size-corrected union of rejections formed from all four of the these unit root tests. Combining the tests in this way represents a fairly natural extension of the individual analyses of

HLT. Asymptotic evidence shows that this simple approach performs reasonably well and, by virtue of its construction, avoids the substantial power losses associated with inappropriate test selection (e.g. the QD demeaned test when there is no trend but a large initial condition). However, its power performance can still fall some way below that of the most appropriate test for a particular situation (e.g. the OLS detrended test when a trend and large initial condition are both present).

In order to address this shortcoming, in section 5 we propose a modification to the four-way union of rejections approach. This modified procedure incorporates extra information gained from auxiliary statistics used to detect the presence of a linear trend component and to detect a large initial condition. These auxiliary statistics are not however used as conventional pre-tests to select between models. For example, an insignificant trend statistic is not taken to imply that no trend is present. Rather, their role is somewhat less assertive and simply used to indicate the possibility that a trend, or large initial condition, may be present, thereby providing information which can be used to tailor the union of rejections in an appropriate fashion. Asymptotic evidence shows that this modified procedure works very well, restoring most of the power losses that can arise with its unmodified counterpart across the differing trend and initial condition specifications.

In addition, an interesting by-product of our analysis is that the behaviour of some recently proposed "robust" statistics designed to detect the presence of a linear trend is very sensitive to the initial condition. Specifically, we show that their power to detect trends that are present in the data can be extremely low when the initial condition is not small. As a consequence, we suggest that extreme caution should be exercised when employing such trend pre-tests to decide whether to exclude a trend term from any unit root test regression, and indeed in other situations.

In section 6 we conduct an assessment of the relative finite sample performance of the unmodified and modified union of rejections procedures. These results quite convincingly demonstrate that the much better power asymptotic properties yielded by the modification should also be accessible in practice. In section 7, using Canadian and U.S. monthly interest rate data, we provide empirical examples which attempt to illustrate the sensitivity of inference of the individual unit root tests when initial conditions change and trends may or may not be present, and also show that the modified union of rejections can provide robust unit root inference in these same circumstances. Section 8 offers some conclusions.

In what follows we use the following notation: '$x := y$' ('$x =: y$') indicates that x is defined by y (y is defined by x); '$\xrightarrow{d}$' denotes weak convergence as the sample size diverges; $\mathbb{I}(\cdot)$ denotes the indicator function, and $\lfloor \cdot \rfloor$ denotes the integer part of its argument.

2 The Model

Consider the case where we have a sample of T observations generated according to the data generating process [DGP]:

$$
\begin{aligned}
y_t &= \mu + \beta_T t + u_t, \quad t = 1, ..., T && (1)\\
u_t &= \rho_T u_{t-1} + \varepsilon_t, \quad t = 2, ..., T. && (2)
\end{aligned}
$$

Within (2), we set $\rho_T := 1 - c/T$ for $0 \leq c < \infty$. Here $c = 0$ corresponds to the unit root case; and $c > 0$ the local alternative..

The innovation process $\{\varepsilon_t\}$ of (2) is taken to satisfy the following conventional (cf. Chang and Park, 2002, and Phillips and Solo, 1992, *inter alia*) stable and invertible linear process-type assumption:

Assumption 1 *The stochastic process* $\{\varepsilon_t\}$ *is such that*

$$\varepsilon_t = C(L)e_t, \ C(L) := \sum_{i=0}^{\infty} C_i L^i, \ C_0 := 1$$

with $C(z) \neq 0$ *for all* $|z| \leq 1$ *and* $\sum_{i=0}^{\infty} i|C_i| < \infty$, *and where* $\{e_t\}$ *is a martingale difference sequence with conditional variance* σ^2 *and* $\sup_t E(e_t^4) < \infty$. *We also define* $\sigma_\varepsilon^2 := E(\varepsilon_t^2)$ *and* $\omega_\varepsilon^2 := \lim_{T\to\infty} T^{-1}E(\sum_{t=1}^{T} \varepsilon_t)^2 = \sigma^2 C(1)^2$.

The next assumption specifies the behaviour of the coefficient on the linear trend term in (1), providing an appropriate Pitman (local) drift for our subsequent asymptotic analyses. This assumption coincides with that employed in HLT.

Assumption 2 *The trend coefficient,* β_T, *in (1) satisfies* $\beta_T := \kappa\omega_\varepsilon T^{-1/2}$, *where* κ *is a finite constant.*

Remark 2.1. It is appropriate that we consider a local trend model in order that the subsequent asymptotic analysis in the paper reflects the uncertainty that exists in finite samples over whether a linear trend is present in the data or not. If one were to assume the trend coefficient was fixed (independent of T) then, for example, the pre-tests for the presence of a trend considered in section 5.1 of this paper would reject the null hypothesis of no trend with probability one in the limit.

Remark 2.2. The scaling of β by ω_ε in Assumption 2 is simply a convenience measure to ensure that ω_ε does not appear in subsequent expressions for the limit distributions.

Our final assumption concerns the initial condition, u_1 in (2).

Assumption 3 *The initial condition,* u_1, *is generated according to* $u_1 = \xi$, *where* $\xi := \alpha\sqrt{\omega_\varepsilon^2/(1-\rho_T^2)}$ *for* $\rho_T := 1 - c/T$, $c > 0$. *For* $c = 0$, *that is under* H_0, *we may set* $\xi = 0$, *without loss of generality, due to the exact similarity of the unit root tests considered in this paper to the initial conditions; see, e.g., Müller and Elliott (2003).*

Remark 2.3. In Assumption 3, α controls the magnitude of u_1 relative to the magnitude of the standard deviation of a stationary $AR(1)$ process with parameter ρ_T and innovation long-run variance ω_ε^2. The form given for ξ is consistent with the analysis of Müller and Elliott (2003), Elliott and Müller (2006).

3 Conventional Unit Root Tests

Our focus in this paper is on testing the unit root null hypothesis $H_0 : c = 0$ in (2) against the local alternative hypothesis $H_1 : c > 0$. To that end, the unit root test statistics we consider are the ADF t-ratio tests of Elliott *et al.* (1996) based on QD demeaning ($DF\text{-}QD^{\mu}$) or detrending ($DF\text{-}QD^{\tau}$); and the ADF t-ratio tests of Dickey and Fuller (1979) and Said and Dickey (1984) based on OLS demeaning ($DF\text{-}OLS^{\mu}$) and detrending ($DF\text{-}OLS^{\tau}$).

The $DF\text{-}QD^i$ test ($i = \mu,\ \tau$) rejects for large negative values of the t-statistic for $\rho = 1$ in the fitted regression equation

$$\tilde{u}_t = \rho\tilde{u}_{t-1} + \sum_{j=1}^{p} \phi_j \Delta\tilde{u}_{t-j} + e_t, \quad t = p+2, ..., T \tag{3}$$

where, on setting $\bar{\rho}_T := 1 - \bar{c}/T$, $\tilde{u}_t := y_t - z_t'\tilde{\theta}$, with $\tilde{\theta}$ obtained from the QD regression of $\mathbf{y}_{\bar{c}} := (y_1, y_2 - \bar{\rho}_T y_1, ..., y_T - \bar{\rho}_T y_{T-1})'$ on $\mathbf{Z}_{\bar{c}} := (z_1, z_2 - \bar{\rho}_T z_1, ..., z_T - \bar{\rho}_T z_{T-1})'$, where $z_t := 1$ for $DF\text{-}QD^{\mu}$, and $z_t := (1,\ t)'$ for $DF\text{-}QD^{\tau}$. The value of the QD parameter, $\bar{c}$, is specified according to the form of the deterministic vector z_t and the desired significance level; see Elliott *et al.* (1996) for details. For $DF\text{-}QD^{\mu}$, they suggest $\bar{c} = 7$, while for $DF\text{-}QD^{\tau}$, $\bar{c} = 13.5$. It is assumed that p is chosen according to some consistent model selection procedure, such as the MAIC procedure of Ng and Perron (2001) and Perron and Qu (2007).

The $DF\text{-}OLS^i$ test ($i = \mu,\ \tau$) rejects for large negative values of the t-statistic for $\rho = 1$ in the fitted regression

$$\hat{u}_t = \rho\hat{u}_{t-1} + \sum_{j=1}^{p} \phi_j \Delta\hat{u}_{t-j} + e_t, \quad t = p+2, ..., T \tag{4}$$

where $\hat{u}_t := y_t - z_t'\hat{\theta}$ is the residual from an OLS regression of y_t on $z_t := 1$, $\theta = \mu$ for $DF\text{-}OLS^{\mu}$ or $z_t := (1,\ t)'$, $\theta = (\mu, \beta)'$ for $DF\text{-}OLS^{\tau}$. Again p is assumed to be chosen via a consistent model selection procedure.

The large sample behaviour of the four unit root tests is summarized in the following lemma.

Lemma 1 *Let $\{y_t\}$ be generated according to (1)-(2) with $\rho_T := 1 - c/T$, $0 \leq c < \infty$, and let Assumptions 1-3 hold. Then,*

$$DF\text{-}QD^{\mu} \xrightarrow{d} \frac{\{\kappa + K_c(1)\}^2 - 1}{2\sqrt{\int_0^1 \{\kappa r + K_c(r)\}^2 dr}} =: \tau_1(c, \kappa, \alpha)$$

$$DF\text{-}QD^{\tau} \xrightarrow{d} \frac{K_c^{\tau,\bar{c}}(1)^2 - 1}{2\sqrt{\int_0^1 K_c^{\tau,\bar{c}}(r)^2 dr}} =: \tau_2(c, \alpha)$$

$$DF\text{-}OLS^{\mu} \xrightarrow{d} \frac{\{\frac{\kappa}{2}+K_c^{\mu}(1)\}^2-\{\frac{-\kappa}{2}+K_c^{\mu}(0)\}^2-1}{2\sqrt{\int_0^1\{\kappa(r-\frac{1}{2})+K_c^{\mu}(r)\}^2dr}} =: \tau_3(c,\kappa,\alpha)$$

$$DF\text{-}OLS^{\tau} \xrightarrow{d} \frac{K_c^{\tau}(1)^2-K_c^{\tau}(0)^2-1}{2\sqrt{\int_0^1 K_c^{\tau}(r)^2dr}} =: \tau_4(c,\alpha)$$

where

$$K_c(r) := \begin{cases} W(r) & c=0 \\ \alpha(e^{-rc}-1)(2c)^{-1/2}+W_c(r) & c>0 \end{cases},$$

and

$$\begin{aligned} K_c^{\mu}(r) &:= \quad K_c(r)-\int_0^1 K_c(s)ds, \\ K_c^{\tau}(r) &:= \quad K_c^{\mu}(r)-12\left(r-\frac{1}{2}\right)\int_0^1\left(s-\frac{1}{2}\right)K_c(s)ds, \\ K_c^{\tau,\bar{c}}(r) &:= \quad K_c(r)-r\left\{\bar{c}^*K_c(1)+3(1-\bar{c}^*)\int_0^1 rK_c(r)dr\right\}, \end{aligned}$$

with $W_c(r) := \int_0^r e^{-(r-s)c}dW(s)$, *where* $W(r)$ *is a standard Wiener process, and* $\bar{c}^* := (1+\bar{c})/(1+\bar{c}+\bar{c}^2/3)$.

Proof: *Under Assumptions 1 and 3, we have that* $T^{-1/2}u_{\lfloor Tr\rfloor} \xrightarrow{d} \omega_{\varepsilon}K_c(r)$; *see, for example, Müller and Elliott (2003). The results then follow directly from Lemma 1 of HLT upon replacing the functional* $W_c(r)$ *of that lemma with* $K_c(r)$.

Remark 3.1. Lemma 1 makes clear the dependence of the limiting distributions of both *DF-QD*$^{\mu}$ and *DF-OLS*$^{\mu}$ on the local trend parameter κ for $c \geq 0$. The limiting distributions of all four tests also depend on the initial condition parameter α whenever $c > 0$. Notice also that the limiting distributions of *DF-QD*$^{\mu}$ and *DF-OLS*$^{\mu}$ do not depend on the variance and serial correlation nuisance parameters related to the linear process of Assumption 1 (σ^2, σ_{ε}^2 and ω_{ε}^2) when a local trend term is omitted. This is because the parametric lagged difference-based (and, for that matter, non-parametric kernel-based estimators) of these quantities are still consistent under the local trend mis-specification.

Figures 1 – 4 about here

Figures 1-4 show the asymptotic power functions of the four tests across $c = \{0,1,2,...,30\}$. These power functions are given for values of the local trend parameter $\kappa = 0,1,2,4$ and initial condition parameter $\alpha = 0,\pm0.5,\pm1,\pm2,\pm4,\pm6$. We use asymptotic critical values appropriate for a nominal 0.05 significance level for a correctly specified model i.e. $\kappa = 0$ for *DF-QD*$^{\mu}$ and *DF-OLS*$^{\mu}$.[1] Critical values

[1] In the following asymptotic simulations we do not need to separately consider negative values for κ. When $\alpha = 0$, our results are the same for negative and positive κ. When $\alpha \neq 0$, results for negative κ and positive (negative) α are the same as those for positive κ and negative (positive) α.

for $DF\text{-}QD^{\mu}$, $DF\text{-}QD^{\tau}$, $DF\text{-}OLS^{\mu}$, $DF\text{-}OLS^{\tau}$ are given by -1.94, -2.85, -2.86 and -3.42, respectively. The results were obtained by direct simulation of the limiting distributions in Lemma 1, approximating the Wiener processes using $NIID(0, 1)$ random variates, and with the integrals approximated by normalized sums of 1000 steps. Here and throughout the paper, simulations were programmed in Gauss 7.0 using 50,000 Monte Carlo replications.

Figure 1 shows the results for the case where $\kappa = 0$. For $\alpha = 0$, a within-test comparison makes clear the emphatic asymptotic power gains achieved by tests which exclude linear trend terms, i.e. $DF\text{-}QD^{\mu}$ and $DF\text{-}OLS^{\mu}$ over their detrended counterparts: $DF\text{-}QD^{\tau}$ and $DF\text{-}OLS^{\tau}$. This is particularly marked when we compare $DF\text{-}QD^{\mu}$ to $DF\text{-}QD^{\tau}$. Comparing across tests, it is also evident that $DF\text{-}OLS^{\mu}$ and $DF\text{-}OLS^{\tau}$ are not competitive in this environment; the power curve of $DF\text{-}OLS^{\mu}$ is, for example, near identical to that of $DF\text{-}QD^{\tau}$. Here then, by some considerable margin, the best inference is provided by $DF\text{-}QD^{\mu}$, which is known to be near-efficient in this case; see ERS.[2] For $\alpha \neq 0$, the results are symmetric in the sign of α. We see that the powers of $DF\text{-}QD^{\mu}$ and $DF\text{-}QD^{\tau}$ decline very rapidly as $|\alpha|$ increases (both have power below size everywhere for $|\alpha| = 4$), whereas the powers of $DF\text{-}OLS^{\mu}$ and $DF\text{-}OLS^{\tau}$ increase steadily in $|\alpha|$. Here $DF\text{-}QD^{\mu}$ is the most powerful test amongst those considered only for $|\alpha| = 0.5$, thereafter $DF\text{-}OLS^{\mu}$ always provides the most powerful test.

Results for $\kappa = 1$ are given in Figure 2. For $\alpha = 0$, both $DF\text{-}QD^{\mu}$ and $DF\text{-}OLS^{\mu}$ are slightly undersized and have power below their size everywhere, which would be expected given that they exclude the local trend term. Rather less expected is that when $|\alpha|$ is large, although size does not alter, $DF\text{-}OLS^{\mu}$ does recover some power, as does $DF\text{-}QD^{\mu}$, albeit to a much lesser extent. Notice also that the behaviour of the demeaned tests, $DF\text{-}QD^{\mu}$ and $DF\text{-}OLS^{\mu}$, is no longer symmetric in the sign of α; there is clearly some interaction between the unattended trend and the initial condition. For $DF\text{-}OLS^{\mu}$, other things being equal, power is higher when α is negative than when it is positive. While much less obvious, the converse would seem to be true for $DF\text{-}QD^{\mu}$. The power profiles of $DF\text{-}QD^{\tau}$and $DF\text{-}OLS^{\tau}$ are of course identical to those in Figure 1 since these tests are exact invariant to β_T.

In Figure 3, when $\kappa = 2$, $DF\text{-}QD^{\mu}$ and $DF\text{-}OLS^{\mu}$ become further undersized and the region where $DF\text{-}OLS^{\mu}$ recovers power is now much contracted relative to that which was seen in Figure 2. Once $\kappa = 4$, as in Figure 4, the sizes of $DF\text{-}QD^{\mu}$ and $DF\text{-}OLS^{\mu}$ are only just above zero and the transient region where $DF\text{-}OLS^{\mu}$ showed power has mostly vanished.

Taking the results of Figures 1-4 together, it becomes clear that we should want $DF\text{-}QD^{\mu}$ to supply inference on the unit root hypothesis when $\kappa = 0$ and $|\alpha|$ is zero or very small, and would want $DF\text{-}OLS^{\mu}$ to fulfill this function for any larger $|\alpha|$. Similarly, when $\kappa \neq 0$, inference would be best based on $DF\text{-}QD^{\tau}$ when $|\alpha|$ is zero or very small and based on $DF\text{-}OLS^{\tau}$ otherwise (dismissing the transient behaviour of the demeaned tests in this case).

[2]Although not formally efficient, in the limit these tests lie arbitrarily close to the asymptotic Gaussian local power envelopes for these testing problems.

4 A Union of Rejections Strategy

Since each of the four tests is capable of providing near-efficient asymptotic unit root inference within some area of the local trend and initial condition parameter space, and also because no test is asymptotically oversized anywhere, then in the absence of any information regarding the magnitudes of the local trend or initial condition parameters, a feasible unit root testing strategy can be based on a four-way union of rejections formed from the tests. That is, we reject the unit root null if any of the four tests rejects the null hypothesis.

More precisely, letting $cv_\gamma^{Q,i}$ and $cv_\gamma^{O,i}$ denote the asymptotic null critical values of $DF\text{-}QD^i$ and $DF\text{-}OLS^i$ respectively, $i = \mu, \tau$, for individual tests conducted at a significance level γ, we employ the decision rule:

$$\begin{aligned}\text{Reject } H_0 \text{ if } \big\{ & DF\text{-}QD^\mu < \tau_\gamma cv_\gamma^{Q,\mu} \text{ or } DF\text{-}QD^\tau < \tau_\gamma cv_\gamma^{Q,\tau} \\ & \text{or } DF\text{-}OLS^\mu < \tau_\gamma cv_\gamma^{O,\mu} \text{ or } DF\text{-}OLS^\tau < \tau_\gamma cv_\gamma^{O,\tau} \big\}\end{aligned} \tag{5}$$

with τ_γ (a scaling constant whose relevance will be made clear shortly) set to unity in (5). Observe that, as in HLT's rejoinder, the decision rule in (5) can be written in terms of the composite statistic

$$\begin{aligned} UR_4(\gamma) := & \; DF\text{-}QD^\mu \mathbb{I}(DF\text{-}QD^\mu < \tau_\gamma cv_\gamma^{Q,\mu}) \\ + DF\text{-}QD^\tau \mathbb{I}(DF\text{-}QD^\mu \geq & \; \tau_\gamma cv_\gamma^{Q,\mu}) \mathbb{I}(DF\text{-}QD^\tau < \tau_\gamma cv_\gamma^{Q,\tau}) \\ + DF\text{-}OLS^\mu \mathbb{I}(DF\text{-}QD^\mu \geq & \; \tau_\gamma cv_\gamma^{Q,\mu}) \mathbb{I}(DF\text{-}QD^\tau \geq \tau_\gamma cv_\gamma^{Q,\tau}) \mathbb{I}(DF\text{-}OLS^\mu < \tau_\gamma cv_\gamma^{O,\mu}) \\ + DF\text{-}OLS^\tau \mathbb{I}(DF\text{-}QD^\mu \geq & \; \tau_\gamma cv_\gamma^{Q,\mu}) \mathbb{I}(DF\text{-}QD^\tau \geq \tau_\gamma cv_\gamma^{Q,\tau}) \mathbb{I}(DF\text{-}OLS^\mu \geq \tau_\gamma cv_\gamma^{O,\mu}) \end{aligned} \tag{6}$$

where a rejection of H_0 is recorded when $UR_4(\gamma) = DF\text{-}QD^\mu < \tau_\gamma cv_\gamma^{Q,\mu}$ or $UR_4(\gamma) = DF\text{-}QD^\tau < \tau_\gamma cv_\gamma^{Q,\tau}$ or $UR_4(\gamma) = DF\text{-}OLS^\mu < \tau_\gamma cv_\gamma^{O,\mu}$ or $UR_4(\gamma) = DF\text{-}OLS^\tau < \tau_\gamma cv_\gamma^{O,\tau}$. In the limit, we therefore have from Lemma 1 and applications of the continuous mapping theorem that

$$\begin{aligned} UR_4(\gamma) \xrightarrow{d} & \; \tau_1(c,\kappa,\alpha) \mathbb{I}(\tau_1(c,\kappa,\alpha) < \tau_\gamma cv_\gamma^{Q,\mu}) \\ + \tau_2(c,\alpha) \mathbb{I}(\tau_1(c,\kappa,\alpha) \geq & \; \tau_\gamma cv_\gamma^{Q,\mu}) \mathbb{I}(\tau_2(c,\alpha) < \tau_\gamma cv_\gamma^{Q,\tau}) \\ + \tau_3(c,\kappa,\alpha) \mathbb{I}(\tau_1(c,\kappa,\alpha) \geq & \; \tau_\gamma cv_\gamma^{Q,\mu}) \mathbb{I}(\tau_2(c,\alpha) \geq \tau_\gamma cv_\gamma^{Q,\tau}) \mathbb{I}(\tau_3(c,\kappa,\alpha) < \tau_\gamma cv_\gamma^{O,\mu}) \\ + \tau_4(c,\alpha) \mathbb{I}(\tau_1(c,\kappa,\alpha) \geq & \; \tau_\gamma cv_\gamma^{Q,\mu}) \mathbb{I}(\tau_2(c,\alpha) \geq \tau_\gamma cv_\gamma^{Q,\tau}) \mathbb{I}(\tau_3(c,\kappa,\alpha) \geq \tau_\gamma cv_\gamma^{O,\mu}). \end{aligned} \tag{7}$$

As is clear from the representation in (7), while the asymptotic size of the strategy given in (5) is not dependent on α, it is, however, dependent on κ. This dependence arises through the presence of $DF\text{-}QD^\mu$ and $DF\text{-}OLS^\mu$, both of whose sizes are decreasing functions of κ. As a consequence, the maximum size of the strategy obtains when $\kappa = 0$ and then declines monotonically in κ. Of course, this maximum size well exceeds the size γ of the individual tests at this point.[3] A conservative strategy which

[3]For example, when $\kappa = 0$, if the individual tests each have asymptotic size 0.05, then the union with $\tau_\gamma = 1$ has asymptotic size 0.145. This is somewhat lower than the Bonferroni upper bound of 0.200 of this union, due to the high positive correlation between $DF\text{-}QD^\mu$ and $DF\text{-}OLS^\tau$.

yields a maximum size of γ across κ can therefore be obtained by choosing τ_γ such that (5) yields size γ when $\kappa = 0$. As noted in HLT's rejoinder, an alternative way of representing the decision rule in (5) is given by:

$$\text{Reject } H_0 \text{ if } \min\Big\{DF\text{-}QD^\mu, \tfrac{cv_\gamma^{Q,\mu}}{cv_\gamma^{Q,\tau}} DF\text{-}QD^\tau, \\ \tfrac{cv_\gamma^{Q,\mu}}{cv_\gamma^{O,\mu}} DF\text{-}OLS^\mu, \tfrac{cv_\gamma^{Q,\mu}}{cv_\gamma^{O,\tau}} DF\text{-}OLS^\tau\Big\} < \tau_\gamma cv_\gamma^{Q,\mu}, \tag{8}$$

from which it is straightforward to determine the appropriate value of τ_γ. Setting $c = 0$ and $\kappa = 0$, we simulate the decision rule in (8) using the (joint) limit distributions of *DF-OLS*$^\mu$, *DF-QD*$^\mu$, *DF-OLS*$^\tau$ and *DF-QD*$^\tau$ as given in Lemma 1, obtain an asymptotic γ-level critical value from this empirical CDF, cv_γ^m say, and then $\tau_\gamma = cv_\gamma^m / cv_\gamma^{Q,\mu}$. Panel A of Table reports τ_γ for the usual significance levels $\gamma = 0.10, 0.05, 0.01$. In what follows we will refer to this conservative four-way union of rejections strategy as UR(*DF-QD*$^\mu$,*DF-QD*$^\tau$,*DF-OLS*$^\mu$,*DF-OLS*$^\tau$), or more simply as UR when no ambiguity arises.

Figures 1-4 also show the asymptotic behaviour of UR, again for $\gamma = 0.05$. For $\kappa = 0$, it is quite clear from Figure 1 that UR does a very decent job of mimicking the power profile of the best test, *DF-QD*$^\mu$, when $|\alpha|$ is zero or very small, and as a consequence fairly comprehensively outperforms the other three tests. For the larger values of $|\alpha|$, while UR is generally much better than *DF-OLS*$^\tau$ (and therefore obviously far better than *DF-QD*$^\mu$ and *DF-QD*$^\tau$), it still falls some way behind the best test, *DF-OLS*$^\mu$. When $\kappa \neq 0$ (Figures 2-4), while UR always avoids the very low power typically associated with *DF-QD*$^\mu$ and *DF-OLS*$^\mu$ and, for larger $|\alpha|$, *DF-QD*$^\tau$, it struggles to match the high power of *DF-QD*$^\tau$ for small $|\alpha|$ or the high power of *DF-OLS*$^\tau$ for larger $|\alpha|$. In fact, *DF-OLS*$^\tau$ would almost always be preferred to UR when $\kappa \neq 0$.

5 A Modified Union of Rejections Strategy

Given the mixed performance of the UR strategy seen in the previous section, we might then consider whether it is possible to modify the procedure to make it more competitive with the best performing tests outside of the cases where $\kappa = 0$ and $|\alpha|$ is very small.[4] As it stands, UR does not incorporate any information that may be accessible on the magnitudes of the local trend and initial condition parameters. Since it is not possible to consistently estimate κ or α in the current framework, no consistent pre-test of, for example, $\kappa = 0$ or $\alpha = 0$ can be constructed. For the purposes of UR, however, we are simply interested in detecting whether $|\kappa|$ or $|\alpha|$ is large in the sense that $|\kappa|$ is of sufficient magnitude to seriously compromise the powers of *DF-QD*$^\mu$ and *DF-OLS*$^\mu$, and $|\alpha|$ is of sufficient magnitude to seriously compromise the powers of *DF-QD*$^\mu$ and *DF-QD*$^\tau$.

[4] In the context of uncertainty with regard to the trend alone, information gleaned from trend tests can be used to improve the power of the simple union of rejections strategy proposed by HLT; see Breitung's commentary on HLT and HLT's rejoinder.

Let s_β and s_α represent statistics whose distributions, other things being equal, shift further rightward as $|\kappa|$ and $|\alpha|$ increase, respectively. We then assume that if $s_\beta > c_\beta$ (where c_β is some critical value) this may indicate a large value of $|\kappa|$, and, similarly $s_\alpha > c_\alpha$ (where c_α is a critical value) may indicate a large value of $|\alpha|$. However, it is important to stress that we do not assume that $s_\beta \leq c_\beta$ or $s_\alpha \leq c_\alpha$ indicates that $|\kappa|$ or $|\alpha|$ are negligibly small. Then we can consider the following modified union of rejections strategy, denoted $UR(s_\beta, s_\alpha)$ in what follows:

Definition 1 *The modified union of rejections strategy, $UR(s_\beta, s_\alpha)$, is as follows:*

(i) If $s_\beta \leq c_\beta$ and $s_\alpha \leq c_\alpha$, then use the conservative decision rule, $UR(DF\text{-}QD^\mu, DF\text{-}QD^\tau, DF\text{-}OLS^\mu, DF\text{-}OLS^\tau)$ of section 4;
(ii) If $s_\beta \leq c_\beta$ and $s_\alpha > c_\alpha$, then use the decision rule, $UR(DF\text{-}OLS^\mu, DF\text{-}OLS^\tau)$;
(iii) If $s_\beta > c_\beta$ and $s_\alpha \leq c_\alpha$, then use the decision rule, $UR(DF\text{-}QD^\tau, DF\text{-}OLS^\tau)$;
(iv) If $s_\beta > c_\beta$ and $s_\alpha > c_\alpha$, then use the decision rule, Reject H_0 if $DF\text{-}OLS^\tau < cv_\gamma^{O,\tau}$.

In *(ii)* of Definition 1, $UR(DF\text{-}OLS^\mu, DF\text{-}OLS^\tau)$ represents the decision rule

$$\text{Reject } H_0 \text{ if } \left\{DF\text{-}OLS^\mu < \tau'_\gamma cv_\gamma^{O,\mu} \text{ or } DF\text{-}OLS^\tau < \tau'_\gamma cv_\gamma^{O,\tau}\right\}$$

where the scaling constant τ'_γ ensures that this two-way union of rejections strategy is conservative, with a maximum asymptotic size of γ across κ; these constants are provided in Panel B of Table 1. This decision rule can also be written as:

$$\text{Reject } H_0 \text{ if } \min\left\{DF\text{-}OLS^\mu, \frac{cv_\gamma^{O,\mu}}{cv_\gamma^{O,\tau}} DF\text{-}OLS^\tau\right\} < \tau'_\gamma cv_\gamma^{O,\mu}.$$

Similarly, in *(iii)* of Definition 1, $UR(DF\text{-}QD^\tau, DF\text{-}OLS^\tau)$ represents the decision rule

$$\text{Reject } H_0 \text{ if } \left\{DF\text{-}QD^\tau < \tau''_\gamma cv_\gamma^{Q,\tau} \text{ or } DF\text{-}OLS^\tau < \tau''_\gamma cv_\gamma^{O,\tau}\right\}$$

where the scaling constant τ''_γ is again used to ensure that the asymptotic size is γ; these constants are given in Panel C of Table 1. Notice that $UR(DF\text{-}QD^\tau, DF\text{-}OLS^\tau)$ is the two-way union of rejections decision rule considered in HLT to deal with uncertainty over the initial condition, u_1, when a linear trend is assumed to be present in the DGP.[5] Again, this can be written in the form

$$\text{Reject } H_0 \text{ if } \min\left\{DF\text{-}QD^\tau, \frac{cv_\gamma^{Q,\tau}}{cv_\gamma^{O,\tau}} DF\text{-}OLS^\tau\right\} < \tau''_\gamma cv_\gamma^{Q,\tau}. \tag{9}$$

The rationale behind the modified strategy in Definition 1 is as follows. Under *(i)* we have no reason to suggest that either the local trend parameter or the initial condition parameter is large, but nor can we assume that they are necessarily small. As such, an obvious strategy is therefore to perform the conservative four-way union of rejections, UR, from section 4. Under *(ii)* there may be some evidence of a large initial

[5] Notice that this last strategy is not conservative across κ, since the sizes of $DF\text{-}QD^\tau$ and $DF\text{-}OLS^\tau$ are (exact) invariant to the value of the local trend parameter.

condition in which case $DF\text{-}QD^{\mu}$ and $DF\text{-}QD^{\tau}$ are known to have very low power. We therefore remove them from the union of rejections. However, since we cannot be sure that the local trend is small, we still need to consider both $DF\text{-}OLS^{\mu}$ and $DF\text{-}OLS^{\tau}$. Under *(iii)* there is some evidence of a large local trend, so that $DF\text{-}QD^{\mu}$ and $DF\text{-}OLS^{\mu}$ may have very low power and so they are consequently removed from the union of rejections. Since we cannot now be sure that the initial condition is small, we still need to consider both $DF\text{-}QD^{\tau}$ and $DF\text{-}OLS^{\tau}$. Under *(iv)* there may be evidence of both a large local trend and initial condition, both of which would cause $DF\text{-}QD^{\mu}$ to have low power, while the former would cause $DF\text{-}OLS^{\mu}$ to lack power and the latter would cause $DF\text{-}QD^{\tau}$ to lack power. We therefore move to a consideration of $DF\text{-}OLS^{\tau}$ alone. Although s_{β} and s_{α} might be considered as being employed as local trend and initial condition pre-tests within $UR(s_{\beta}, s_{\alpha})$, it is important to reiterate that their role within our procedure is only as an indicator of possibilities and *not* one of model selection. For example, if $s_{\beta} \leq c_{\beta}$ we do *not* conclude that $\kappa = 0$, and subsequently only consider demeaned unit root tests.

In order to make $UR(s_{\beta}, s_{\alpha})$ of Definition 1 operational, we need to select statistics to fulfill the roles of s_{β} and s_{α}. This is the subject matter of the following two subsections.

5.1 Trend Detection

For s_{β} we will consider four candidate statistics, each of which can be used to test the null hypothesis $\beta_T = 0$ against alternative $\beta_T \neq 0$ in (1), although again it should be stressed that we are not using these statistics for this purpose here. Each of these belongs to the class of so-called "robust" tests for trend in the sense that the asymptotic critical values for testing $\beta_T = 0$ are the same regardless of whether u_t contains a unit root or is stationary. Specifically, the statistics we consider are the t_{λ} and t_{λ}^{m2} statistics of Harvey *et al.* (2007), the *t-PSW*[1] statistic of Vogelsang (1998) and the *Dan-J* statistic of Bunzel and Vogelsang (2005).

The t_{λ} statistic of Harvey *et al.* (2007) is a switching-based strategy that attains the local limiting Gaussian power envelope for testing $\beta_T = 0$ against $\beta_T \neq 0$ in (1) irrespective of whether u_t contains a unit root or is stationary (when an asymptotically negligible initial condition is assumed). The test statistic is also asymptotically standard normal under the null hypothesis $\beta_T = 0$ in both cases. It is calculated as

$$t_{\lambda} := (1 - \lambda^*)t_0 + \lambda^* t_1$$

where

$$t_0 := \frac{\hat{\beta}_T}{\sqrt{\hat{\omega}_u^2 / \sum_{t=1}^{T}(t - \bar{t})^2}} \quad \text{and} \quad t_1 := \frac{\breve{\beta}_T}{\sqrt{\breve{\omega}_v^2/(T-1)}}. \tag{10}$$

Here, $\hat{\beta}_T$ denotes the OLS estimator of β_T from (1) and $\hat{\omega}_u^2$ is a long run variance estimator formed using $\hat{u}_t := y_t - \hat{\mu} - \hat{\beta}_T t$. Also, $\breve{\beta}_T$ is the OLS estimator of β_T from (1) estimated in first differences i.e. from $\Delta y_t = \beta_T + v_t,\ \ t = 2, ..., T$ and $\breve{\omega}_v^2$ is a long run variance estimator based on $\breve{v}_t := \Delta y_t - \breve{\beta}_T$. The long run variance estimators are computed using the quadratic spectral kernel with Newey and West (1994) automatic bandwidth selection adopting a non-stochastic prior bandwidth of $\lfloor 4(T/100)^{2/25} \rfloor$. The weight function λ^* is defined as

$$\lambda^* := \exp\left(-0.00025\left(\frac{\mathit{DF\text{-}QD}^\tau}{\mathit{KPSS}^\tau}\right)^2\right)$$

where KPSS^τ is the Kwiatkowski *et al.* (1992) stationarity test statistic

$$\mathit{KPSS}^\tau := \frac{\sum_{t=1}^{T}\left(\sum_{i=1}^{t}\hat{u}_i\right)^2}{T^2\hat{\omega}_u^2}.$$

Harvey *et al.* (2007) show that a modified variant of t_λ, denoted t_λ^{m2}, can provide a more powerful test of the trend hypothesis than t_λ when u_t contains a near unit root. This replaces t_1 with $t_1^{m\,2} := \delta_\gamma R_2 t_1$ where

$$R_2 := \left(\frac{\check{\omega}_v^2}{T^{-1}\hat{\sigma}_u^2}\right)^2$$

and $\hat{\sigma}_u^2 := (T-2)^{-1}\sum_{t=1}^{T}\hat{u}_t^2$. Here δ_γ is a constant chosen so that, at a given significance level γ, t_λ^{m2} has a standard normal critical value for unit root and stationary u_t. For a two tailed 0.05 level test, $\delta_\gamma = 0.00115$.

The $t\text{-}PSW_T^1$ statistic of Vogelsang (1998) is calculated as

$$t\text{-}PSW^1 := t_0' \exp(-c_\gamma' J)$$

where t_0' is t_0 as defined in (10) but with the long run variance estimator, $\hat{\omega}_u^2$, replaced by $100T^{-1}s_z^2$ where $s_z^2 := T^{-1}\sum_{t=1}^{T}\hat{\eta}_t^2$, $\hat{\eta}_t$ denoting the OLS residuals obtained from estimating the partially summed regression

$$\sum_{i=1}^{t} y_i = \mu t + \beta_T \sum_{i=1}^{t} i + \eta_t.$$

In the expression for the $t\text{-}PSW^1$ statistic, the t_0' statistic is scaled by a function of the J unit root test statistic of Park (1990) and Park and Choi (1988). The constant c_γ' is chosen so that for a significance level γ, $t\text{-}PSW^1$ has the same critical value under both $I(0)$ and $I(1)$ errors; for a two tailed 0.05 level test, the asymptotic critical value is 1.015, and $c_\gamma' = 1.036$.

The *Dan-J* statistic of Bunzel and Vogelsang (2005) is essentially a modified version of $t\text{-}PSW^1$ that employs a long run variance estimator based on the "fixed-b" asymptotics of Keifer and Vogelsang (2002). Specifically, the statistic is

$$\mathit{Dan\text{-}J} := t_0'' \exp(-c_\gamma'' J)$$

where t_0'' is t_0 as defined in (10) but with the long run variance estimator, $\hat{\omega}_u^2$, constructed using the Daniell kernel with a data-dependent bandwidth. The bandwidth is given by $\max(\hat{b}_{\text{opt}}T, 2)$, where $\hat{b}_{\text{opt}} = b_{\text{opt}}(\hat{c})$. Here, $\hat{c} := T(1-\hat{\rho})$ with $\hat{\rho}$ obtained by OLS estimation of (1) and (2); and $b_{\text{opt}}(.)$ is a step function given in Bunzel and Vogelsang (2005). In the expressions for *Dan-J*, the t_0'' statistic is scaled by a function of the same J unit root test statistic used in the *t-PSW*^{1} statistic. Again c_γ'' is a constant chosen so that for a significance level γ, *Dan-J* has the same critical value under both $I(0)$ and $I(1)$ errors. The value of c_γ'' depends on $\hat{b}_{\text{opt}}$; Bunzel and Vogelsang (2005) provide a response surface for determining c_γ'' for a given significance level, and $\hat{b}_{\text{opt}}$. The critical values for the test also depend on $\hat{b}_{\text{opt}}$, and again a response surface is provided by the authors for a variety of significance levels. Because c is not consistently estimated using $\hat{c}$, Bunzel and Vogelsang (2005) only provide a limiting distribution for *Dan-J* when it is assumed that c is known in the calculation of $\hat{b}_{\text{opt}}$. That is, when $\hat{b}_{\text{opt}} = b_{\text{opt}}(\hat{c})$ is replaced by $b_{\text{opt}}(c)$. Although this strictly means that their asymptotic results are based on the limiting behaviour of an infeasible test, for the purposes of making comparisons tractable, in what follows the limit distribution for *Dan-J* is that using $b_{\text{opt}}(c)$.

The asymptotic properties of the four trend statistics are summarized in the following lemma.

Lemma 2 *Let the conditions of Lemma 1 hold. Then,*

$$
\begin{aligned}
t_\lambda &\xrightarrow{d} \kappa + K_c(1) \\
t_\lambda^{m2} &\xrightarrow{d} \delta_\gamma \frac{\kappa + K_c(1)}{\left\{\int_0^1 K_c^\tau(r)^2 dr\right\}^2} \\
t\text{-}PSW^1 &\xrightarrow{d} \frac{\kappa + 12\int_0^1 \left(r-\frac{1}{2}\right) K_c(r)dr}{\sqrt{12}\sqrt{100 H_c(r)'\left\{\int_0^1 G_c(r)G_c(r)'dr\right\}^{-1} H_c(r)}} \exp\left(-c_\gamma' A_c(r)\right) \\
Dan\text{-}J &\xrightarrow{d} \frac{\kappa + 12\int_0^1 \left(r-\frac{1}{2}\right) K_c(r)dr}{\sqrt{12}\sqrt{\int_0^1\int_0^1 -k^{*\prime\prime}(r-s)\left\{\int_0^r K_c^\tau(t)dt\right\}\left\{\int_0^s K_c^\tau(t)dt\right\} drds}} \exp\left(-c_\gamma'' A_c(r)\right)
\end{aligned}
$$

where $K_c(r)$ and $K_c^\tau(r)$ are as defined in Lemma 1, $k^{*\prime\prime}(x)$ denotes the second derivative with respect to x of the Daniell kernel evaluated at $x/b_{opt}(c)$,

$$
\begin{aligned}
G_c(r) &:= \left[\begin{array}{cc} r, & r^2 \end{array}\right]' \\
H_c(r) &:= \left[\begin{array}{cc} \int_0^1 r\int_0^r K_c(s)dsdr, & \frac{1}{2}\int_0^1 r^2\int_0^r K_c(s)dsdr \end{array}\right]'
\end{aligned}
$$

and

$$
A_c(r) := \frac{\int_0^1 K_c^\tau(r)^2 dr}{\int_0^1 L_c(r)^2 dr} - 1,
$$

with $L_c(r)$ used to denote the continuous time projection of $K_c(r)$ onto the space spanned by $\{1, r, r^2, ..., r^9\}$.

Proof: *The results for t_λ, t_λ^{m2} and Dan-J follow directly from Lemma 3 of HLT on replacing the functional $W_c(r)$ with $K_c(r)$. The result for t-PSW^1 is shown along the lines of Vogelsang (1998), Theorem 2, for the case when $q = 1$, on replacing the standard Ornstein-Uhlenbeck process in that theorem with $K_c(r)$.*

It is immediately clear from the representations given in Lemma 2 that all four statistics depend in the limit on the initial condition parameter α whenever $c > 0$. Figures 5-8 show the corresponding asymptotic power functions of the four tests across $\kappa = \{0, 0.1, 0.2, ..., 4.0\}$. These power functions are given for $c = 5, 10, 20, 30$ and, as in Figures 1-4, $\alpha = 0, \pm 0.5, \pm 1, \pm 2, \pm 4, \pm 6$. We use asymptotic critical values appropriate for a nominal 0.05 significance level for the two-tailed tests $|t_\lambda|$, $|t_\lambda^{m2}|$, $|t\text{-}PSW^1|$ and $|Dan\text{-}J|$.

Figures 5 – 8 about here

Examining first the case of $c = 5$ in Figure 5, we see that consistent with the findings of Harvey *et al.* (2007), all of the tests, but in particular $|t_\lambda|$, are under-sized when $\kappa = \alpha = 0$. The power performance of the four tests when $\kappa > 0$ varies little across $\alpha = 0, \pm 0.5, \pm 1$. Here $|t_\lambda|$ is unambiguously the most powerful test for larger values of κ, although it has very low power for small κ. Relative to a benchmark of $\alpha = 0$, we see that for the larger values of $|\alpha|$, the powers of $|t_\lambda^{m2}|$, $|t\text{-}PSW^1|$ and $|Dan\text{-}J|$ start to decline quite rapidly. There is also some asymmetry in the observed powers of these three tests here; other things being equal, positive values of α reduce power to a greater extent than do negative values. Once $\alpha = 6$, all of these three tests have power *below* the nominal size across all values of κ considered. For example, for $\kappa = 4.0$, when $\alpha = 0$, $|t_\lambda^{m2}|$ has power above 0.60; this drops to below 0.03 when $\alpha = 6$.

For the larger values of $|\alpha|$, $|t_\lambda|$ behaves somewhat differently to the other three tests. While its power is still generally much lower than the $\alpha = 0$ benchmark for positive α, it is well above this benchmark for negative α. It is also badly oversized (size is around 0.40) when $|\alpha| = 6$. The logic for these phenomena is as follows. Under H_c, t_λ is asymptotically equivalent to t_1 in (10); see Harvey *et al.* (2007). Now observing that $\breve{\beta}_T = (y_T - y_1)/(T - 1)$, it should be clear that when $\kappa = 0$ a very large initial condition can result in $|y_T - y_1|$ being sufficiently large to cause $|t_\lambda|$ to reject too often. In such cases, where $\kappa > 0$ the power of $|t_\lambda|$ would be expected to increase beyond the null rejection frequency when the initial condition, u_1, reinforces the underlying direction of the trend (i.e. when α is negative), and to first decline for smaller values of κ and then later increase for larger values of κ when the initial condition runs counter to the underlying trend.

A qualitatively similar picture is seen when $c = 10$ in Figure 6. The asymptotic local powers of the $|t_\lambda^{m2}|$, $|t\text{-}PSW^1|$ and $|Dan\text{-}J|$ tests decline in $|\alpha|$, again more so for the positive values of α. Other things being equal, the powers of $|t\text{-}PSW^1|$ and $|Dan\text{-}J|$ tend to be somewhat higher overall than for $c = 5$. This is because these tests are less under-sized for $c = 10$ than for $c = 5$. The power of $|t_\lambda|$ again decreases in positive α but increases in negative α. Interestingly, however, the over-sizing seen in $|t_\lambda|$ for

$|\alpha| = 6$ in Figure 5 is no longer evident. This is because for $c = 10$ a larger value of α would be needed, other things equal, to offset the under-sizing in $|t_\lambda|$ when $\alpha = 0$. The same basic patterns continue through $c = 20$ (Figure 7) and $c = 30$ (Figure 8), with the exception of $|t_\lambda^{m2}|$ which is over-sized for $|\alpha| = 4, 6$ when $c = 20$ and $|\alpha| = 1, 2, 4, 6$ when $c = 30$. For example, when $c = 30$ and $|\alpha| = 4$, the empirical size of $|t_\lambda^{m2}|$ is about 0.48. Notice also that in such cases, where $\kappa > 0$ the power of $|t_\lambda^{m2}|$ displays qualitatively similar patterns to those seen for $|t_\lambda|$ in Figure 5(j),(k) for $c = 5$ and $|\alpha| = 6$. A similar explanation seems likely.

What is very clear from these results is that, in the presence of uncertainty about the initial condition, it would be inadvisable to use any of $|t_\lambda|$, $|t_\lambda^{m2}|$, $|t\text{-}PSW^1|$ and $|Dan\text{-}J|$ as a pre-test of $\kappa = 0$ to decide whether to include a constant or constant and linear trend in the model specification prior to unit root testing; i.e., in deciding whether to apply a demeaned or detrended variant of a unit root test. For example, suppose we wish to use these pre-tests to decide between applying $DF\text{-}OLS^\mu$ or $DF\text{-}OLS^\tau$. From Figure 6, which considers $c = 10$, when $\alpha = 6$ and $\kappa = 2$, each of these candidate pre-tests has (to all intents) zero power, so that $DF\text{-}OLS^\mu$ is always erroneously selected. From Figure 3, which considers $\kappa = 2$, we find that when $\alpha = 6$ and $c = 10$, the power of $DF\text{-}OLS^\mu$ is approximately zero. Thus, in this situation, the power of a unit root testing with trend pre-test strategy is effectively zero. In contrast, had we abandoned any pre-test strategy and simply applied $DF\text{-}OLS^\tau$, power would have been roughly 0.44. This example highlights why we incorporate information from trend pre-tests into the modified union of rejections strategy only in a risk-averse fashion.

5.2 Initial Condition Detection

In the last subsection we saw that inference on the local trend cannot be made independent of the behaviour of the initial condition when $c > 0$. In the same circumstances, we can, however, make inference on the initial condition not depend on the local trend simply by considering QD or OLS detrended data. Recall from Figure 1 that the $DF\text{-}QD^\tau$ test reacts completely differently to $DF\text{-}OLS^\tau$ in the presence of a large initial condition. More specifically, other things being equal, $DF\text{-}QD^\tau$ tends to become increasingly less negative as $|\alpha|$ increases, while $DF\text{-}OLS^\tau$ tends to become more negative. This observation, together with the structure of (9), suggests that one possibility for s_α would be to consider

$$s_\alpha := DF\text{-}QD^\tau - \frac{cv_\gamma^{Q,\tau}}{cv_\gamma^{O,\tau}} DF\text{-}OLS^\tau \tag{11}$$

with large values in the upper tail of the distribution of s_α being indicative that $|\alpha|$ is large. Since the joint limit distributions of $DF\text{-}QD^\tau$ and $DF\text{-}OLS^\tau$ are symmetric in α and do not depend on ω_ε^2, other things equal, large values of $|\alpha|$ are then associated with large values of s_α. From a practical perspective, this specification for s_α has the attractive feature that it requires no further computation beyond the calculation of the unit root statistics $DF\text{-}QD^\tau$ and $DF\text{-}OLS^\tau$. The large sample behaviour of s_α is given in the following corollary, which follows directly from Lemma 1 using applications of the continuous mapping theorem.

Corollary 1 *Let the conditions of Lemma 1 hold. Then,*

$$s_\alpha \xrightarrow{d} \frac{K_c^{\tau,\bar{c}}(1)^2 - 1}{2\sqrt{\int_0^1 K_c^{\tau,\bar{c}}(r)^2 dr}} - \frac{cv_\gamma^{Q,\tau}}{cv_\gamma^{O,\tau}} \frac{K_c^\tau(1)^2 - K_c^\tau(0)^2 - 1}{2\sqrt{\int_0^1 K_c^\tau(r)^2 dr}}.$$

Figures 9 – 13 about here

Since s_α does not depend on $|\alpha|$ when $c = 0$, calculating asymptotic critical values for s_α at conventional significance levels γ at the point $c = 0$ does not really make a great deal of sense. Instead we calibrate critical values for $|\alpha| = 0$ when $c = 30$ (the largest value of c we consider). These are given in Table 2. Figure 9 shows the asymptotic power profile of s_α across $|\alpha| = \{0, 0.1, 0.2, ..., 4.0\}$ for $c = 5, 10, 20, 30$ using a nominal 0.05 significance level. The statistic is only correctly sized for $c = 30$, becoming increasingly oversized as c decreases. However, in the current context this is not an issue of great importance since we simply desire s_α to detect large initial conditions when they are actually present. From Figure 9 we see that, whatever the value of c, s_α rejects with probability of about 0.60 for $|\alpha| = 1.2$. Larger values of $|\alpha|$ are detected even more readily. Rejection probabilities reach 1.00 for (roughly) $|\alpha| = 5.7$ when $c = 5$; they reach 1.00 for (roughly) $|\alpha| = 2.5$ when $c = 30$. This is as we would expect, the farther c is from zero, the easier a large initial condition is to detect.

5.3 Asymptotic Performance of The Modified Strategy

Because it is not immediately clear from the results in Figures 5-8 which is the best trend pre-test to select, since each of these can form the most powerful trend pre-test over some part of the κ and α parameter space, we will subsequently consider all of $|t_\lambda|$, $|t_\lambda^{m2}|$ and $|Dan\text{-}J|$ as candidates for s_β in $UR(s_\beta, s_\alpha)$ (we do not further consider $|t\text{-}PSW^1|$ as it performs very similarly to $|Dan\text{-}J|$). In Figures 10-13 we repeat the asymptotic simulation exercises underlying Figures 1-4. Here we present the modified unions of rejections strategies $UR(|t_\lambda|, s_\alpha)$, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$ and, for ease of comparison, repeat the plots for the unmodified strategy UR. Instead of also repeating all the power plots for the individual tests $DF\text{-}QD^\mu$, $DF\text{-}QD^\tau$, $DF\text{-}OLS^\mu$ and $DF\text{-}OLS^\tau$, when $\kappa = 0$ we simply report the power plot of $DF\text{-}QD^\mu$ alone for small $|\alpha|$ and that of $DF\text{-}OLS^\mu$ alone for the larger $|\alpha|$. When $\kappa \neq 0$, we give the power plot of just $DF\text{-}QD^\tau$ for small $|\alpha|$ and that of just $DF\text{-}OLS^\tau$ for the larger $|\alpha|$. This combination of plots can be considered to provide an informal envelope for the unit root tests across the κ and α parameter space.

Throughout Figures 10-13, all tests are run at the nominal asymptotic 0.05 significance level. While the conservative strategy, UR, is correctly sized (for $\kappa = 0$), this is not the case for the modified strategy $UR(s_\beta, s_\alpha)$ due to its dependence on the inferences from s_β and s_α. Since the maximum sizes of $UR(|t_\lambda|, s_\alpha)$, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$ across κ are only 0.069, 0.078 and 0.075, respectively (recall that α

does not influence size), we will simply assimilate these relatively modest upward size distortions and not consider further size corrections.[6]

Figure 10 gives the results for $\kappa = 0$. Since $UR(|t_\lambda|, s_\alpha)$, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$ all behave very similarly across α (particularly the latter two), here we will just refer to them generically as $UR(s_\beta, s_\alpha)$. For $|\alpha| \leq 0.05$, $UR(s_\beta, s_\alpha)$ performs almost as well as UR and, as a result, has power fairly close to that of the (informal) envelope test *DF-QD*$^\mu$. For the larger values of $|\alpha|$, $UR(s_\beta, s_\alpha)$ is more powerful than UR, often substantially so, and is always reasonably close to the envelope test, *DF-OLS*$^\mu$.

In Figure 11, when $\kappa = 1$, we start to see some fairly obvious differences appearing between $UR(|t_\lambda|, s_\alpha)$, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$. The $UR(|t_\lambda|, s_\alpha)$ strategy overall performs the most poorly of the three and, although the rankings change depending on α, $UR(|t_\lambda^{m2}|, s_\alpha)$ appears generally more powerful than $UR(|Dan\text{-}J|, s_\alpha)$; however, the differences are relatively small, especially when one bears in mind the marginally higher maximum size of the former noted above. Both these strategies easily outperform UR everywhere (as does $UR(|t_\lambda|, s_\alpha)$ for the most part), and get very close to the informal envelope *DF-QD*$^\tau$ when $|\alpha|$ is small and very close to the informal envelope *DF-OLS*$^\tau$ for the larger $|\alpha|$. Indeed, for $\alpha = -6.0$ their powers actually lie well above those of the informal envelope for the first half of values of c; of course in this region the union of rejections strategies are capturing some of the rejections by *DF-OLS*$^\mu$ seen in Figure 2(j).

Very similar remarks also apply to the case of $\kappa = 2$ in Figure 12. If anything, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$ behave even more similarly here. Their power profiles also both lie even closer to that given by the informal envelope. Here $UR(|t_\lambda|, s_\alpha)$ also starts to become more competitive. These progressions continue through $\kappa = 4$ in Figure 13. For the most part, the power profiles of all three modified unions of rejections are now near indistinguishable from the informal envelope and therefore they comprehensively dominate UR.

Our asymptotic simulation evidence seems to make a rather convincing case for employing any of the modified union of rejections strategies above the unmodified counterpart. Moreover, although close, the case would appear marginally stronger for the $UR(|t_\lambda^{m2}|, s_\alpha)$ variant. Needless to say, asymptotic evidence can only be considered indicative of what may occur in a finite sample environment. Hence, we now turn to an examination of the behaviour of the strategies as they might be applied in practice, where only a relatively small sample is available.

6 Finite Sample Comparisons

Our finite sample simulations are based on the DGP (1)-(2) with $\varepsilon_t \sim NIID(0,1)$ and a sample size of $T = 150$. We set $\mu = 0$ without loss of generality and consider $\beta_T = \kappa T^{-1/2}$ with $\kappa = 0, 1, 2, 4$, so that the values of β_T correspond to the local trend settings considered in the preceding asymptotic analysis. We consider

[6] For tests run at the nominal 0.10 significance level the maximum sizes of $UR(|t_\lambda|, s_\alpha)$, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$ are 0.126, 0.143 and 0.138, respectively.

$\alpha = 0, \pm 0.5, \pm 1, \pm 2, \pm 4, \pm 6$, as before. The *DF-QD*^{i} and *DF-OLS*^{i}, $i = \mu, \tau$, tests are conducted at the nominal asymptotic 0.05 significance level, while the union of rejections strategies make use of the asymptotic scaling constants and s_α critical values reported in Tables 1 and 2. The number of lagged difference terms, p, included in the unit root test regressions (3) and (4) is determined by application of the MAIC procedure of Ng and Perron (2001) with maximum lag length set at $p_{\max} = \lfloor 12(T/100)^{1/4} \rfloor$, using the modification suggested by Perron and Qu (2007).

Figures 14 – 17 about here

Figures 14-17 report the power functions across $c = \{0, 1, 2, ..., 30\}$ for the same tests and strategies as were presented in Figures 10-13; i.e. UR, $UR(|t_\lambda|, s_\alpha)$, $UR(|t_\lambda^{m2}|, s_\alpha)$, $UR(|Dan\text{-}J|, s_\alpha)$, plus the test that constitutes the informal envelope for each combination of κ and α. First, we can observe that the empirical sizes of the union of rejections strategies are all close to nominal size, with only modest size distortions displayed for this relatively small sample size. The maximum sizes observed for UR, $UR(|t_\lambda|, s_\alpha)$, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$, across all the settings considered, are 0.046, 0.070, 0.069 and 0.067, respectively, so that size behaviour does not appear to be a particular concern in the practical implementation of these strategies.

The relative finite sample power performance of the unit root testing strategies largely mirrors that observed in the limit, with the three modified union of rejections strategies displaying quite similar rejection frequencies for most parameter settings, with power performance that is competitive in comparison to the informal envelope. The simple UR approach is outperformed by the modified strategies in most cases, and on many occasions substantially so. Of the three modified union of rejection strategies, when $\kappa = 0$ the powers are essentially identical for $|\alpha| \leq 4$, while for $|\alpha| = 6$, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$ somewhat outperform $UR(|t_\lambda|, s_\alpha)$. For $\kappa > 0$, the power functions of the three strategies are closer than was observed in the limit. In particular, $UR(|t_\lambda|, s_\alpha)$ performs considerably better than the asymptotic results predict, to the extent that in the cases where power differs among the three approaches, $UR(|t_\lambda|, s_\alpha)$ is generally the most powerful testing strategy ($\alpha = -6$ is an obvious exception), with $UR(|t_\lambda^{m2}|, s_\alpha)$ generally ranked second above $UR(|Dan\text{-}J|, s_\alpha)$. This finding is consistent with the results of Harvey *et al.* (2007), where t_λ was found to have superior power compared to t_λ^{m2} and *Dan-J* in small to moderately sized samples, in contrast to the predictions of the local-to-unity asymptotic analysis.

Taking our asymptotic and finite sample results together we find that the modified union of rejections strategies we consider in this paper provide approaches to testing for a unit root that have decent asymptotic and finite sample size control, and offer good robust power performance in the presence of uncertainty regarding both the presence of a trend and the magnitude of the initial condition. Although there is rather little to choose between the three variants considered, on balance the $UR(|t_\lambda|, s_\alpha)$ and $UR(|t_\lambda^{m2}|, s_\alpha)$ appear to display marginally superior performance to $UR(|Dan\text{-}J|, s_\alpha)$.

7 Empirical Illustrations

In this section we provide two empirical examples to illustrate the behaviour of the unit root tests considered above, using U.S. and Canadian monthly interest rate series over the period 1980M1–2006M12 (324 observations).[7] In both cases, the *DF-QD*^{i} and *DF-OLS*^{i}, $i = \mu, \tau$, tests are conducted at the nominal asymptotic 0.05 level, and, as in the previous section, lag augmentation was performed using the MAIC approach of Ng and Perron (2001) and Perron and Qu (2007), with $p_{\max} = \lfloor 12(T/100)^{1/4} \rfloor$.

Tables 3 – 4 about here

In the first example, we focus on illustrating the role of the initial condition in the relative performance of the tests. We do this by applying the individual unit root tests, and the unmodified and modified union of rejections strategies, to a single series repeatedly, each time moving the start date to examine the sensitivity or robustness of the test outcomes to the initial observation. Specifically, we apply the *DF-QD*$^\mu$, *DF-QD*$^\tau$, *DF-OLS*$^\mu$ and *DF-OLS*$^\tau$ tests, together with the UR and $UR(s_\beta, s_\alpha)$ strategies, to data on Canadian long-term government bond yields using 48 consecutive start dates: 1980M1–1983M12.

The null rejections implied by these tests are reported in Table 3. For each start date, the three modified union of rejections strategies $UR(|t_\lambda|, s_\alpha)$, $UR(|t_\lambda^{m2}|, s_\alpha)$ and $UR(|Dan\text{-}J|, s_\alpha)$ all gave identical inferences, thus we simply record these in a single column in the table, labelled generically as $UR(s_\beta, s_\alpha)$. The tests that do not admit a deterministic trend result in (almost) no rejections across all possible start dates, while rejections are often obtained for the *DF-QD*$^\tau$ and *DF-OLS*$^\tau$ tests, suggesting that a trend is most likely present in the data. Focusing on these latter two procedures, the results clearly highlight the sensitivity of the test outcomes to the initial observation: while both tests (broadly) reject for the first 15 of the start dates considered, only *DF-OLS*$^\tau$ rejects for the next 18 start dates, while only *DF-QD*$^\tau$ rejects for the last 15 start dates. In contrast, the $UR(s_\beta, s_\alpha)$ strategies consistently reject the unit root null across the full range of start dates; indeed, there is only one occasion out of the 48 start dates considered where the $UR(s_\beta, s_\alpha)$ strategies fail to reject. As would be expected from the simulation results, UR rejects less frequently than $UR(s_\beta, s_\alpha)$. The dependence of the power of the individual unit root tests on the magnitude of the initial condition, and the relative robustness of the $UR(s_\beta, s_\alpha)$ strategies, are therefore clearly illustrated in this application.

Our second example illustrates the relative robustness of the $UR(s_\beta, s_\alpha)$ strategies to uncertainty regarding the trend component. Table 4 reports results of the individual tests and the union of rejections strategies applied to Canadian and U.S. short-term Treasury bill rates. It is difficult to have confidence, *a priori*, in deciding whether or not to include a trend component in unit root tests applied to interest rate data, and this uncertainty is borne out in the test results. For the Canadian series, it is a test that includes a trend (*DF-QD*$^\tau$) that provides the only rejection of the null, while

[7]The data are obtained from the International Financial Statistics database.

for the U.S. series, a rejection is only obtained for a test that incorporates a constant term alone (*DF-OLS*$^\mu$). The advantage of using our proposed $UR(s_\beta, s_\alpha)$ strategies is clearly evident here, with rejections of the null being obtained for both countries by all three modified union of rejections strategies (again labelled generically for simplicity), regardless of the fact that one series requires a trend to be included, and the other requires the trend to be excluded, in order for the individual unit root tests to yield null rejections.

8 Conclusions

In this paper we have considered the problem of testing for a unit root when uncertainty exists over both the magnitude of the initial condition, and as to whether or not a linear trend is present in the deterministic component of the series. Building on earlier work in HLT, who develop testing procedures based on separate two-way union of rejections decision rules in the cases where either there is uncertainty over the trend but not over the initial condition, or *vice versa*, we have developed a new procedure which attempts to retain good power properties in the presence of both forms of uncertainty.

We initially investigated a procedure based around a simple four-way union of rejections decision rule, rejecting if any of the QD detrended and demeaned ADF test statistics, *DF-QD*$^\tau$ and *DF-QD*$^\mu$, respectively, of Elliott *et al.* (1996), and the corresponding OLS detrended and demeaned ADF test statistics, *DF-OLS*$^\tau$ and *DF-OLS*$^\mu$, respectively, reject the unit root null hypothesis. The power properties of this basic union of rejections rule were shown to often fall someway short of the power of the best of the four individual tests, however. We consequently proposed a modification of the basic union of rejections rule which incorporated additional sample information gained from auxiliary statistics used to detect the presence of a linear trend and to detect a large initial condition. Reported asymptotic and finite sample evidence suggested that our modified union of rejections procedure displayed decent asymptotic and finite sample size control and offered good robust power performance in the presence of uncertainty over both the presence of a trend and the magnitude of the initial condition. Despite its excellent performance our modified decision rule is very easy to implement, requiring the practitioner to compute only standard unit root test statistics together with a test statistic for the presence of a linear trend. The empirical potential of our proposed approach was illustrated using U.S. and Canadian interest rate data.

An interesting by-product of our numerical analysis was that, and in contrast to what had previously been argued in the literature (see, for example, Vogelsang, 1998, p.136), the behaviour of some recently proposed tests for the presence of a linear trend can be highly sensitive to the magnitude of the initial condition, potentially displaying very low power when the initial condition is not small. We therefore strongly advise against the practice of using these as pre-tests for choosing whether to use a demeaned or detrended variant of a given unit root test.

References

Bunzel, H. and Vogelsang, T.J. (2005). Powerful trend function tests that are robust to strong serial correlation with an application to the Prebisch-Singer hypothesis. Journal of Business and Economic Statistics 23, 381-94.

Chang, Y. and Park, Y.J. (2002). On the asymptotics of ADF tests for unit roots. Econometric Reviews 21, 431–47.

Choi, I. and Phillips, P.C.B. (1993). Testing for a Unit Root by Frequency Domain Regression. Journal of Econometrics 59, 263-286.

Dickey, D.A. and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427–31.

Elliott, G. and Müller, U.K. (2006). Minimizing the impact of the initial condition on testing for unit roots. Journal of Econometrics 135, 285-310.

Elliott, G., Rothenberg, T.J. and Stock, J.H. (1996). Efficient tests for an autoregressive unit root. Econometrica 64, 813-36.

Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2007). A simple, robust and powerful test of the trend hypothesis. Journal of Econometrics 141, 1302-30.

Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2008). Unit Root Testing in Practice: Dealing with Uncertainty over the Trend and Initial Condition (with commentaries and rejoinder). Econometric Theory, forthcoming.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? Journal of Econometrics 54, 159-78.

Müller, U.K. and Elliott, G. (2003). Tests for unit roots and the initial condition. Econometrica 71, 1269–86.

Ng, S. and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. Econometrica 69, 1519-54.

Park, J.Y. (1990). Testing for unit roots and cointegration by variable addition, in Fomby, T. and Rhodes, F. (eds.) Advances in Econometrics: Cointegration, Spurious Regression and Unit Roots. Jai Press: Greenwich.

Park, J.Y. and Choi, B. (1988). A new approach to testing for a unit root. CAE Working Paper 88-23, Cornell University.

Perron, P. and Qu, Z. (2007). A simple modification to improve the finite sample properties of Ng and Perron's unit root tests. Economics Letters 94, 12-19.

Phillips P.C.B. (1987a). Time series regression with a unit root. Econometrica 55, 277-301.

Phillips, P.C.B. (1987b). Towards a unified asymptotic theory for autoregression. Biometrika 74, 535–47.

Phillips, P.C.B. (1990). Time series regression with a unit root and infinite variance errors. Econometric Theory 6, 44-62.

Phillips, P.C.B. (1991a). To criticize the critics: an objective Bayesian analysis of stochastic trends. Journal of Applied Econometrics 6, 333-364.

Phillips, P.C.B. (1991b). Bayesian routes and unit roots: de rebus prioribus semper est disputandum. Journal of Applied Econometrics 6, 435-473.

Phillips, P.C.B. (1998). New tools for understanding spurious regressions. Econometrica 66, 1299-1326.

Phillips, P.C.B. and Magdalinos, T. (2007). Limit theory for moderate deviations from unity. Journal of Econometrics 136, 115-130.

Phillips, P.C.B. and Moon, H.R. (1999). Linear regression limit theory for nonstationary panel data. Econometrica 67, 1057-1111.

Phillips, P.C.B. and Perron, P. (1988). Testing for a unit root in time series regression. Biometrika 75, 335-46.

Phillips, P.C.B. and Ploberger, W. (1994). Posterior odds testing for a unit root with data-based model selection. Econometric Theory 10, 774-808.

Phillips P.C.B. and Solo, V. (1992). Asymptotics for linear processes. Annals of Statistics 20, 971-1001.

Phillips, P.C.B. and Xiao, Z. (1998). A primer on unit root testing. Journal of Economic Surveys 12, 423-70.

Schmidt, P. and Phillips, P.C.B. (1992). LM tests for a unit root in the presence of deterministic trends. Oxford Bulletin of Economics and Statistics 54, 257-87.

Said, S.E. and D.A. Dickey (1984). Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 89, 1420–1437.

Vogelsang, T.J. (1998). Trend function hypothesis testing in the presence of serial correlation. Econometrica 66, 123-48.

Table 1. Asymptotic scaling constants for union of rejections strategies at the γ significance level.

	$\gamma = 0.10$	$\gamma = 0.05$	$\gamma = 0.01$
Panel A. $UR(DF\text{-}QD^{\mu}, DF\text{-}QD^{\tau}, DF\text{-}OLS^{\mu}, DF\text{-}OLS^{\tau})$			
τ_{γ}	1.197	1.157	1.113
Panel B. $UR(DF\text{-}OLS^{\mu}, DF\text{-}OLS^{\tau})$			
τ'_{γ}	1.078	1.064	1.044
Panel C. $UR(DF\text{-}QD^{\tau}, DF\text{-}OLS^{\tau})$			
τ''_{γ}	1.070	1.058	1.043

Table 2. Asymptotic critical values for s_{α} at the γ significance level.

$\gamma = 0.10$	$\gamma = 0.05$	$\gamma = 0.01$
-0.374	-0.181	0.260

Table 3. Application of unit root tests to Canadian long-term interest rates: 1980M1–2006M12

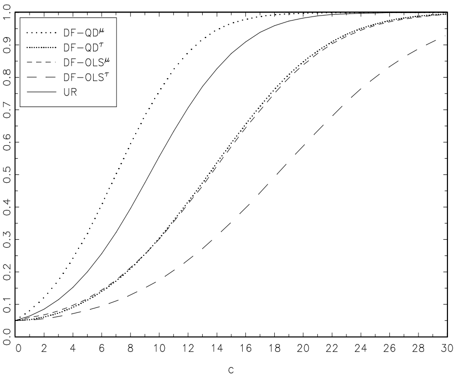
Start date	$DF\text{-}QD^{\mu}$	$DF\text{-}QD^{\tau}$	$DF\text{-}OLS^{\mu}$	$DF\text{-}OLS^{\tau}$	UR	$UR(s_\beta, s_\alpha)$
1980M1	-	R	-	R	-	R
1980M2	-	R	-	R	R	R
1980M3	-	R	-	R	R	R
1980M4	-	R	-	R	-	R
1980M5	-	-	-	R	-	R
1980M6	-	-	-	R	-	R
1980M7	-	R	-	R	-	R
1980M8	-	R	-	R	-	R
1980M9	-	R	-	R	R	R
1980M10	-	R	-	R	R	R
1980M11	-	R	-	R	R	R
1980M12	-	R	-	R	R	R
1981M1	-	R	-	R	R	R
1981M2	-	R	-	R	R	R
1981M3	-	R	-	R	R	R
1981M4	-	-	-	R	-	R
1981M5	-	-	-	R	-	R
1981M6	-	-	-	R	-	R
1981M7	-	-	-	R	R	R
1981M8	-	-	-	R	R	R
1981M9	-	-	R	R	R	R
1981M10	-	-	-	R	R	R
1981M11	-	-	-	R	-	R
1981M12	-	-	-	R	-	R
1982M1	-	-	-	R	R	R
1982M2	-	-	-	R	-	R
1982M3	-	-	-	R	R	R
1982M4	-	-	-	R	-	R
1982M5	-	-	-	R	R	R
1982M6	-	-	-	R	R	R
1982M7	-	-	-	R	R	R
1982M8	-	-	-	R	-	R
1982M9	-	-	-	R	-	R
1982M10	-	R	-	-	-	R
1982M11	-	R	-	-	R	R
1982M12	-	R	-	-	-	R
1983M1	-	R	-	-	-	R
1983M2	-	R	-	-	-	R
1983M3	-	R	-	-	-	R
1983M4	-	R	-	-	-	-
1983M5	-	R	-	-	-	R
1983M6	-	R	-	-	-	R
1983M7	-	R	-	-	-	R
1983M8	-	R	-	-	-	R
1983M9	-	R	-	-	-	R
1983M10	-	R	-	-	-	R
1983M11	-	R	-	-	-	R
1983M12	-	R	-	-	-	R

Note: "R" ("-") denotes rejection (non-rejection) of the unit root null at the 0.05-level.

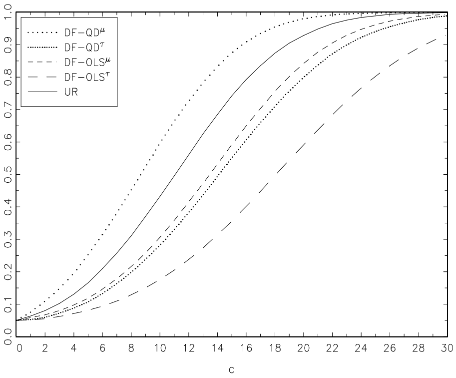
Table 4. Application of unit root tests to US and Canadian short-term interest rates: 1980M1–2006M12

	$DF\text{-}QD^{\mu}$	$DF\text{-}QD^{\tau}$	$DF\text{-}OLS^{\mu}$	$DF\text{-}OLS^{\tau}$	UR	$UR(s_\beta, s_\alpha)$
Canada	-	R	-	-	-	R
U.S.	-	-	R	-	R	R

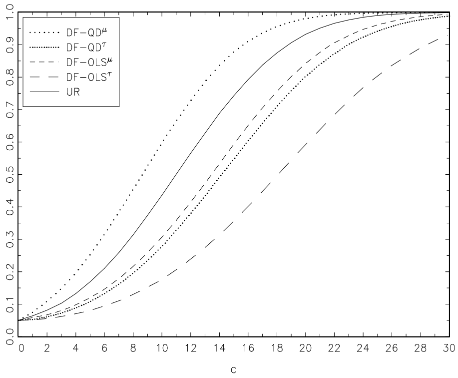
Note: "R" ("-") denotes rejection (non-rejection) of the unit root null at the 0.05-level.



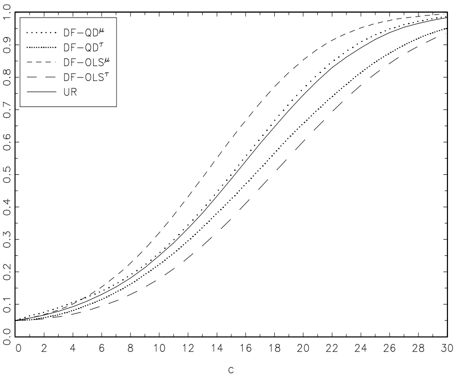
(a) $\alpha = 0$



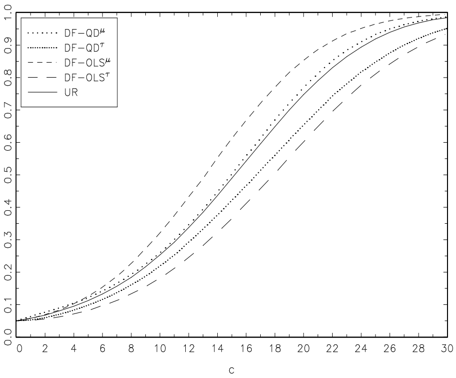
(b) $\alpha = -0.5$



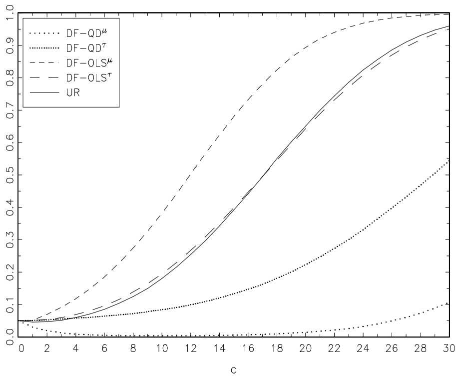
(c) $\alpha = 0.5$



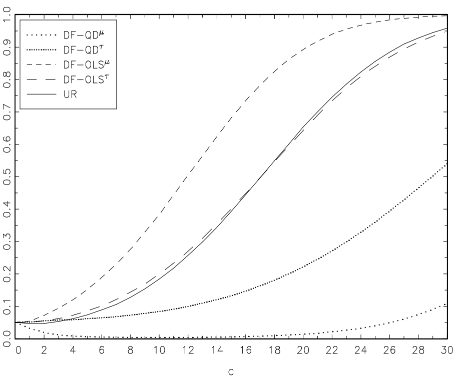
(d) $\alpha = -1$



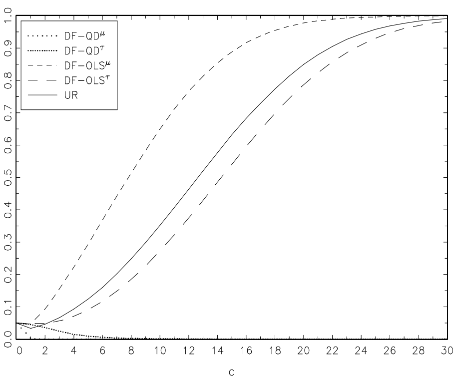
(e) $\alpha = 1$



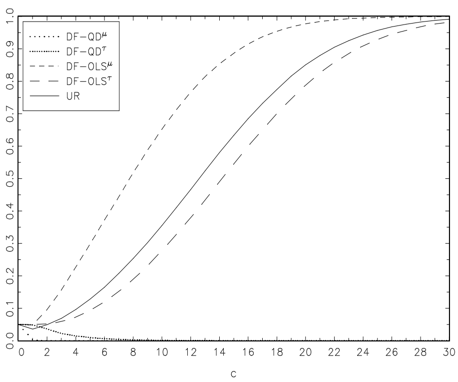
(f) $\alpha = -2$



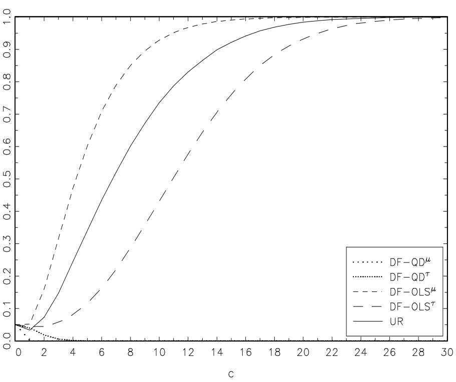
(g) $\alpha = 2$



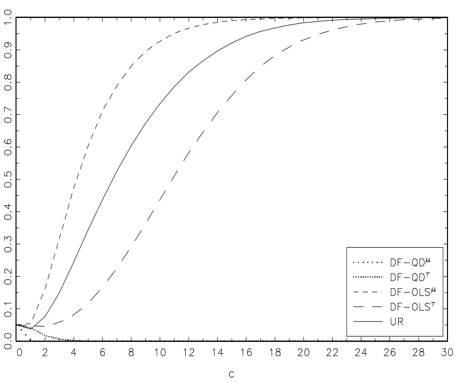
(h) $\alpha = -4$



(i) $\alpha = 4$

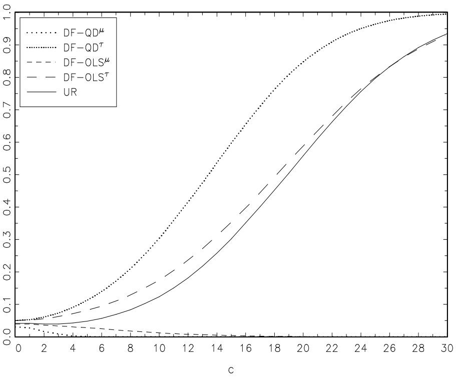


(j) $\alpha = -6$

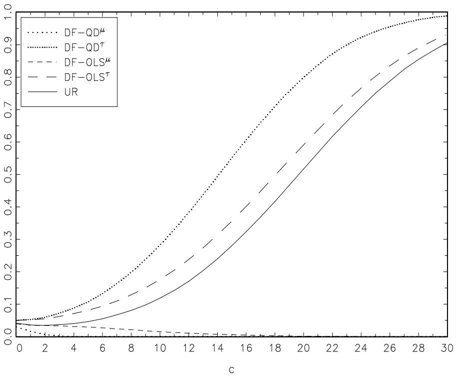


(k) $\alpha = 6$

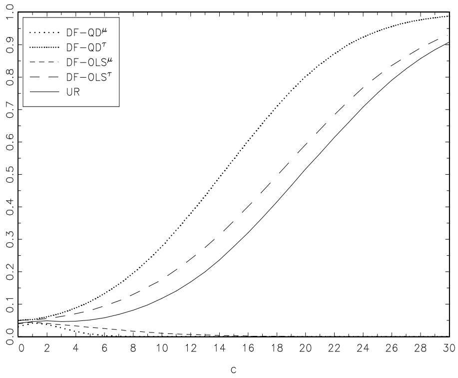
Figure 1. Asymptotic size and local power of unit root tests: $\kappa = 0$



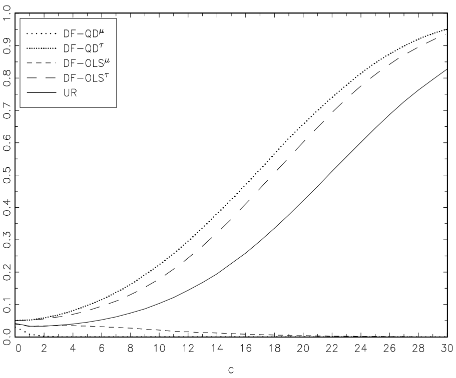
(a) $\alpha = 0$



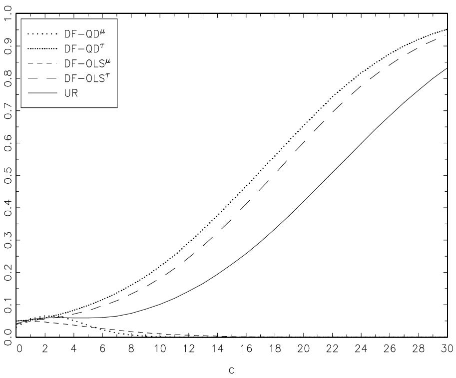
(b) $\alpha = -0.5$



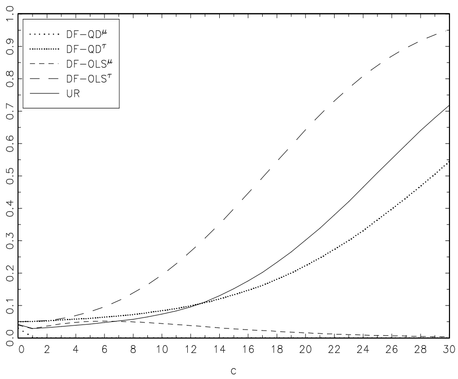
(c) $\alpha = 0.5$



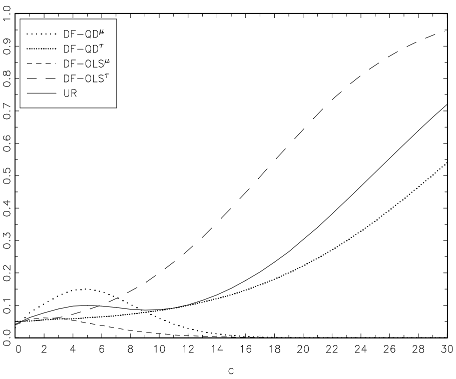
(d) $\alpha = -1$



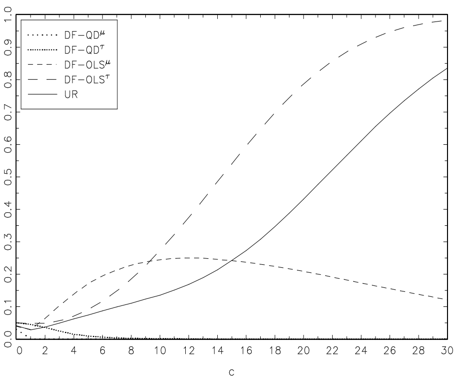
(e) $\alpha = 1$



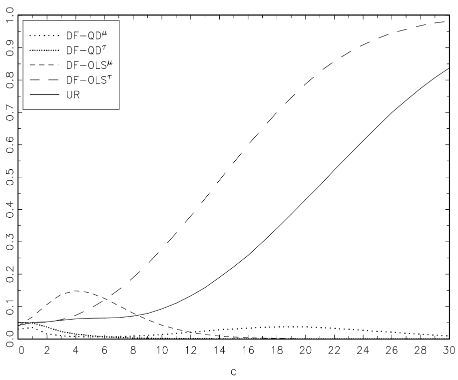
(f) $\alpha = -2$



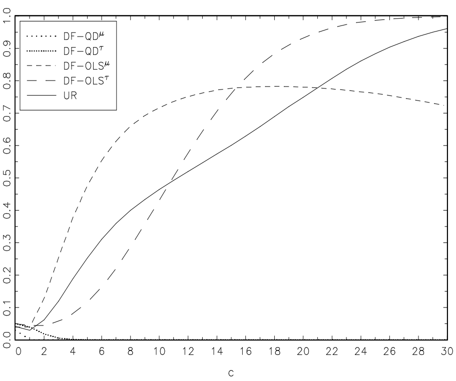
(g) $\alpha = 2$



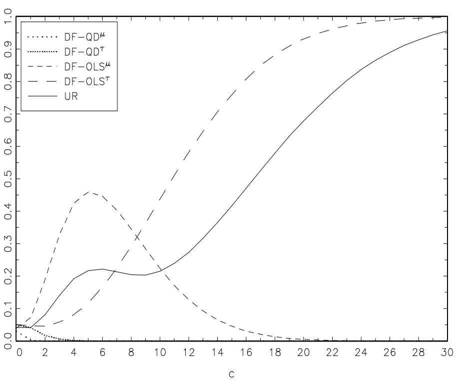
(h) $\alpha = -4$



(i) $\alpha = 4$

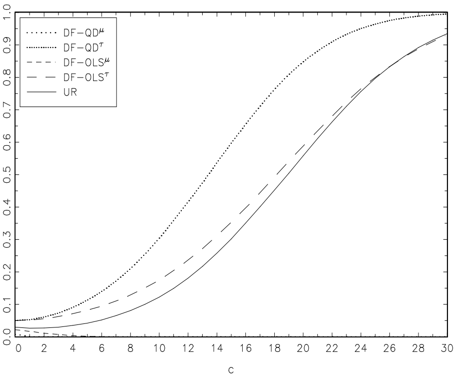


(j) $\alpha = -6$

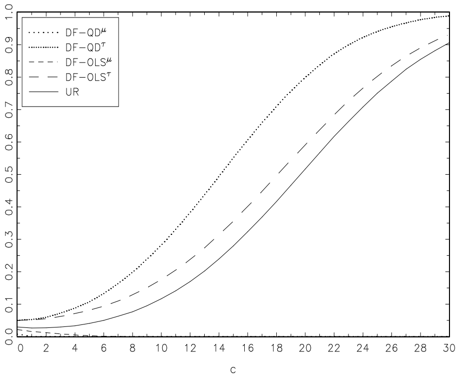


(k) $\alpha = 6$

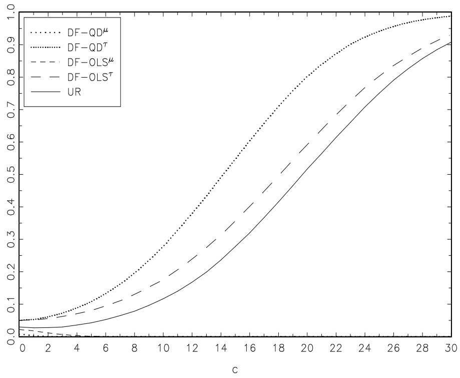
Figure 2. Asymptotic size and local power of unit root tests: $\kappa = 1$



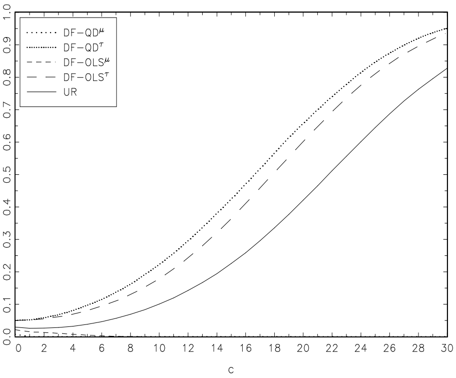
(a) $\alpha = 0$



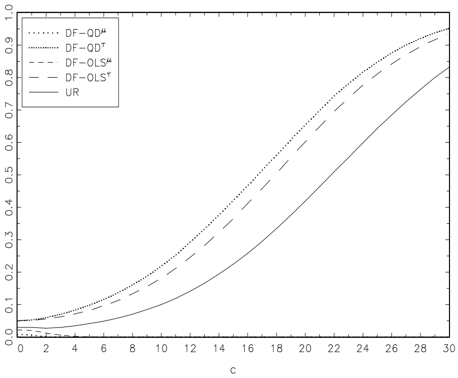
(b) $\alpha = -0.5$



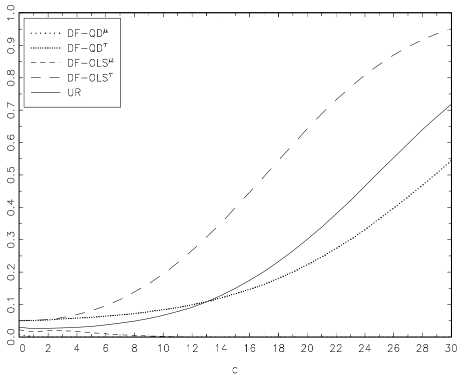
(c) $\alpha = 0.5$



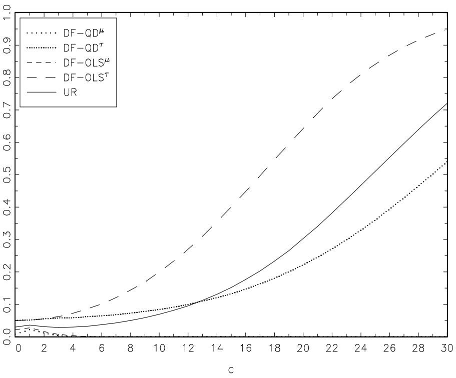
(d) $\alpha = -1$



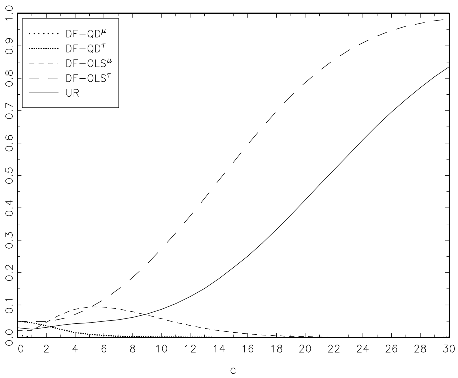
(e) $\alpha = 1$



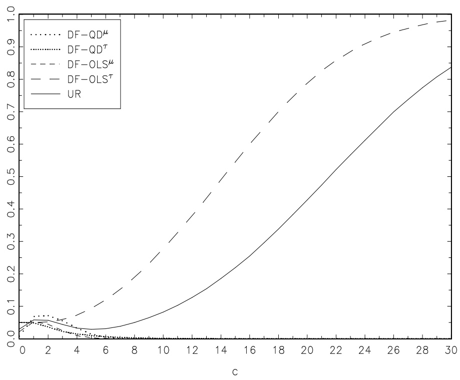
(f) $\alpha = -2$



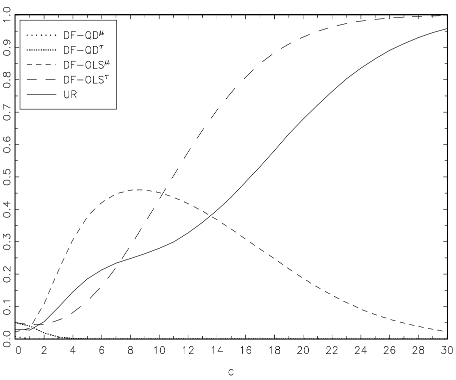
(g) $\alpha = 2$



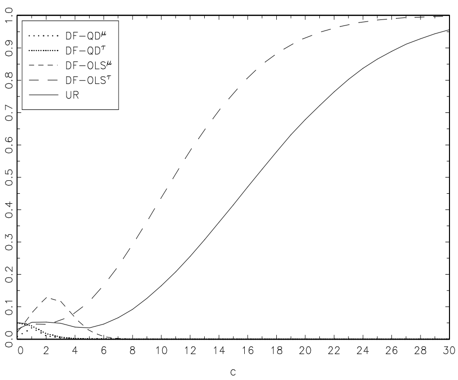
(h) $\alpha = -4$



(i) $\alpha = 4$

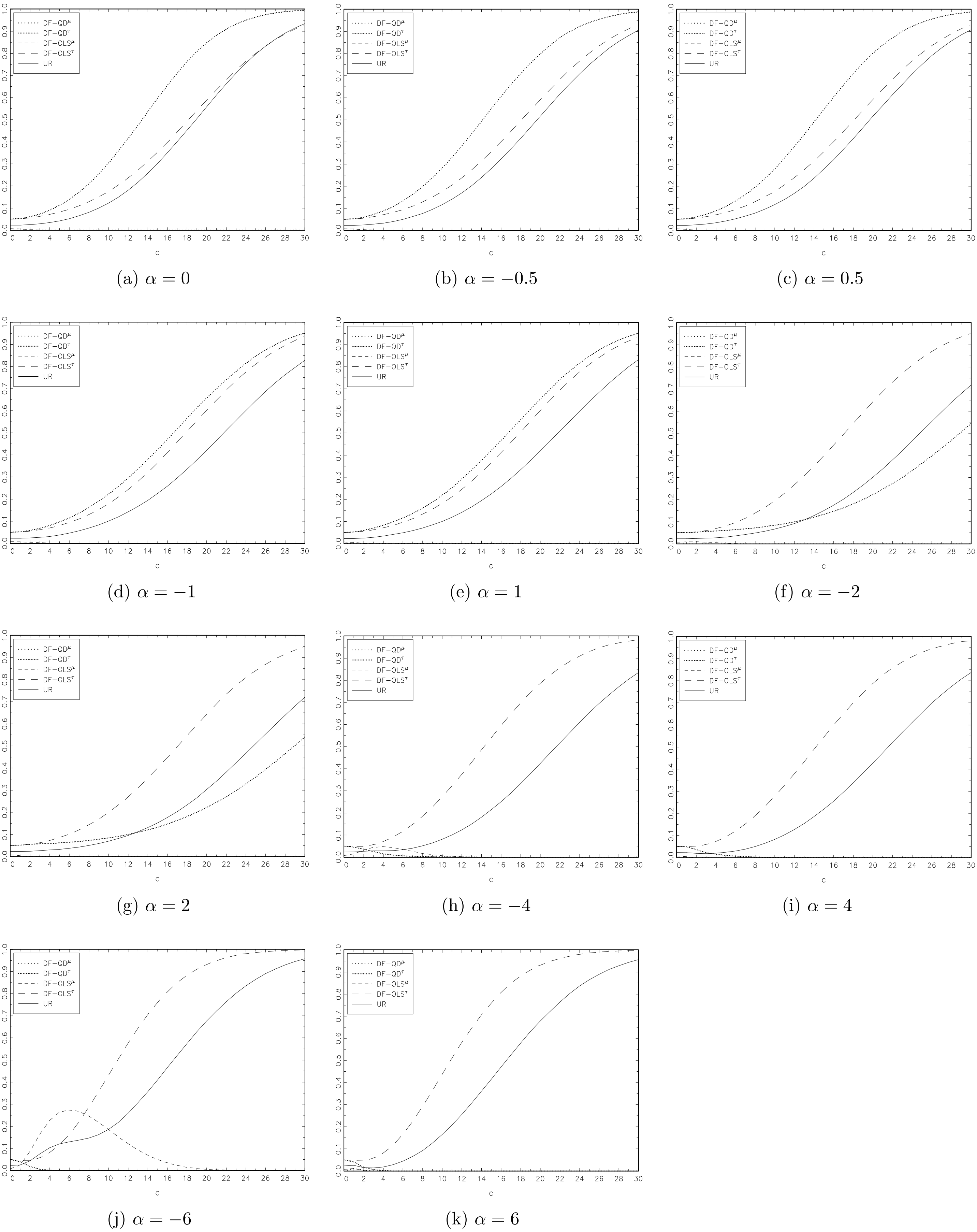


(j) $\alpha = -6$



(k) $\alpha = 6$

Figure 3. Asymptotic size and local power of unit root tests: $\kappa = 2$



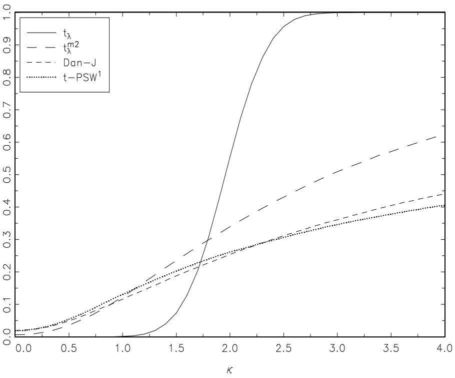
(a) $\alpha = 0$ (b) $\alpha = -0.5$ (c) $\alpha = 0.5$

(d) $\alpha = -1$ (e) $\alpha = 1$ (f) $\alpha = -2$

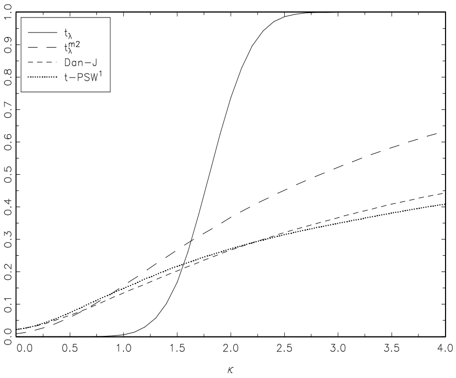
(g) $\alpha = 2$ (h) $\alpha = -4$ (i) $\alpha = 4$

(j) $\alpha = -6$ (k) $\alpha = 6$

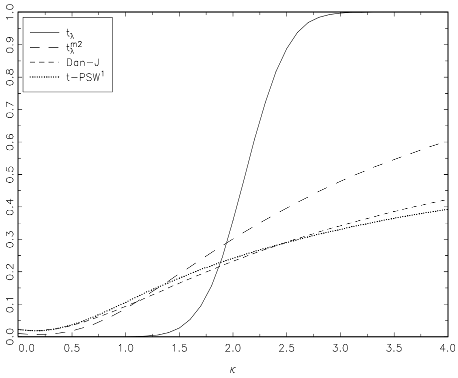
Figure 4. Asymptotic size and local power of unit root tests: $\kappa = 4$



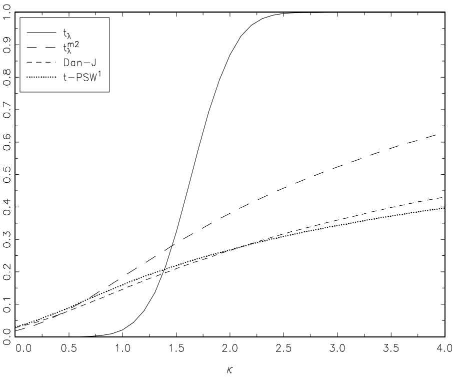
(a) $\alpha = 0$



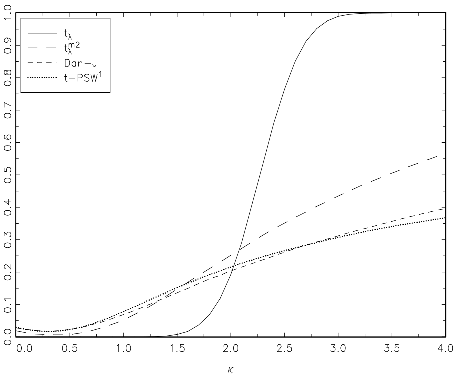
(b) $\alpha = -0.5$



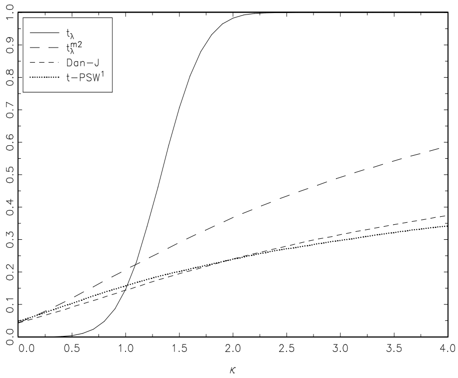
(c) $\alpha = 0.5$



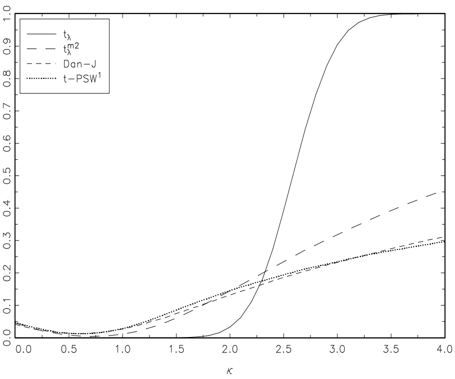
(d) $\alpha = -1$



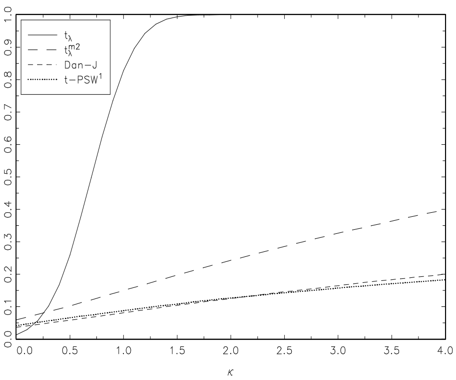
(e) $\alpha = 1$



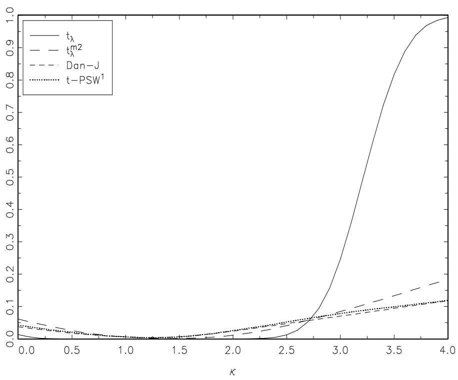
(f) $\alpha = -2$



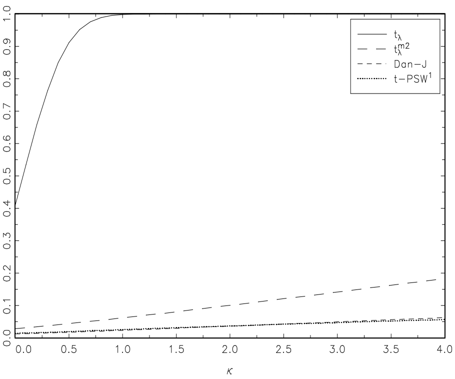
(g) $\alpha = 2$



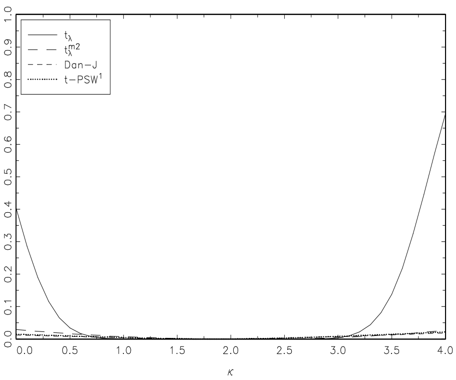
(h) $\alpha = -4$



(i) $\alpha = 4$

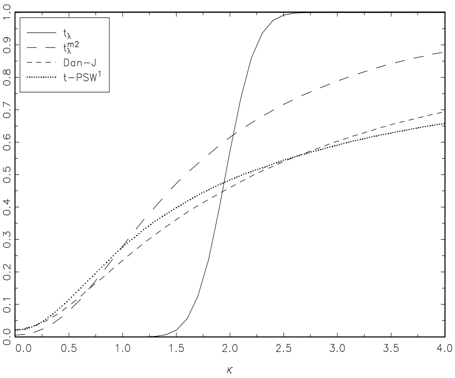


(j) $\alpha = -6$

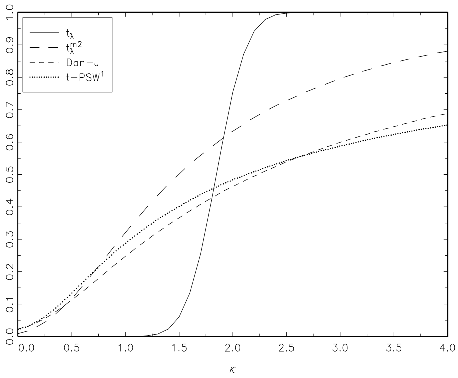


(k) $\alpha = 6$

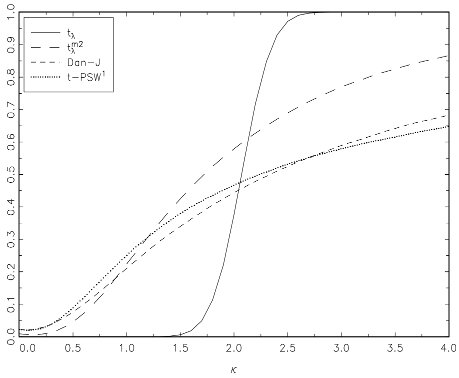
Figure 5. Asymptotic size and local power of s_β: $c = 5$



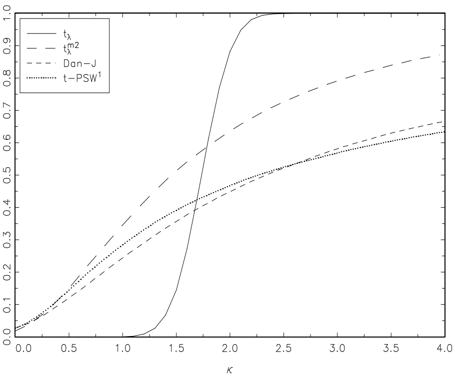
(a) $\alpha = 0$



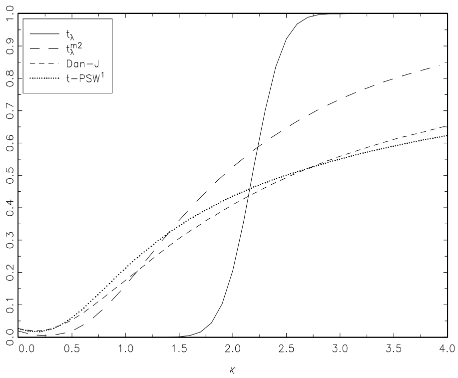
(b) $\alpha = -0.5$



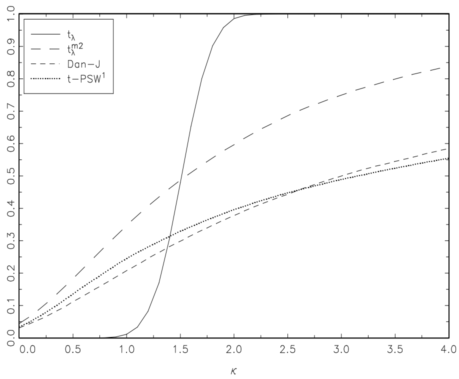
(c) $\alpha = 0.5$



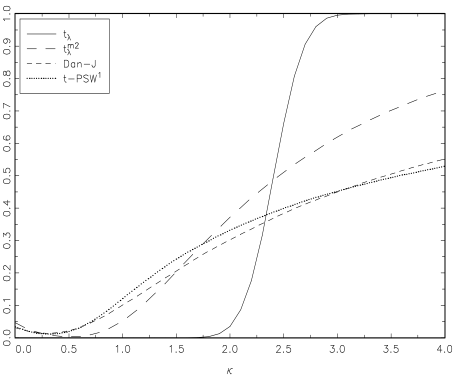
(d) $\alpha = -1$



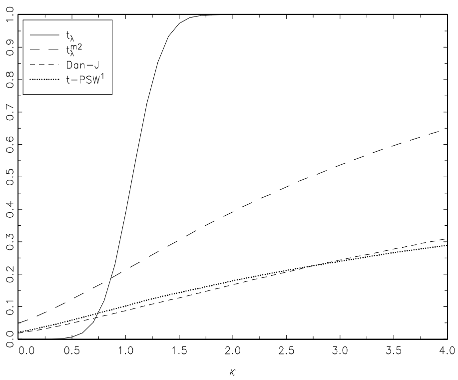
(e) $\alpha = 1$



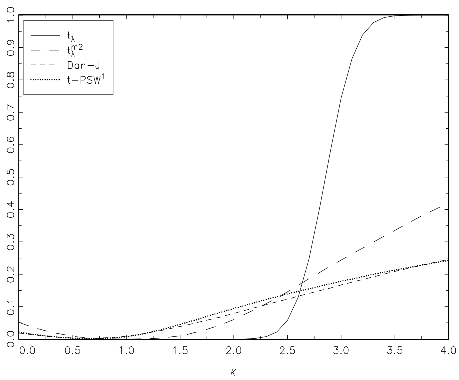
(f) $\alpha = -2$



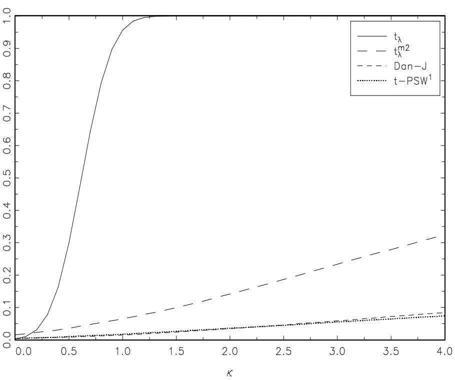
(g) $\alpha = 2$



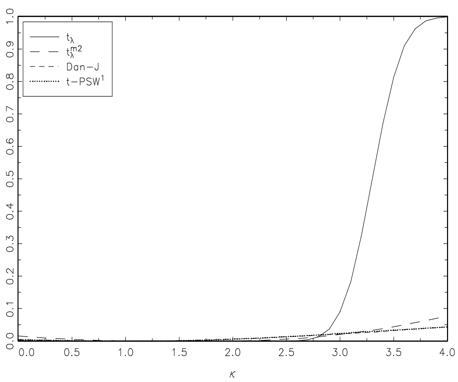
(h) $\alpha = -4$



(i) $\alpha = 4$

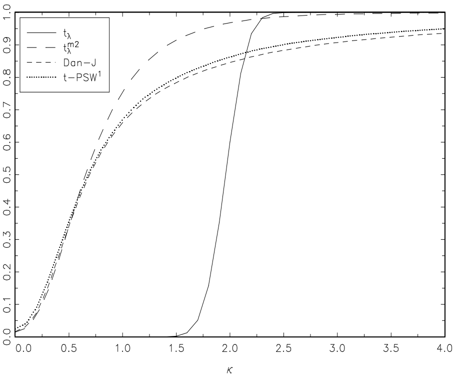


(j) $\alpha = -6$

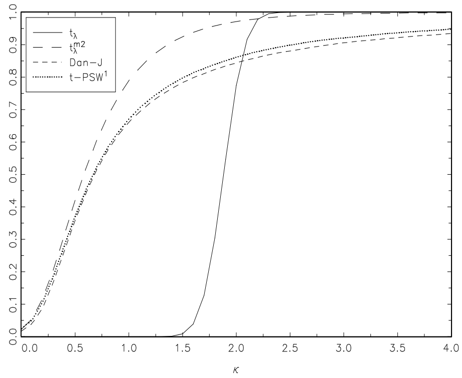


(k) $\alpha = 6$

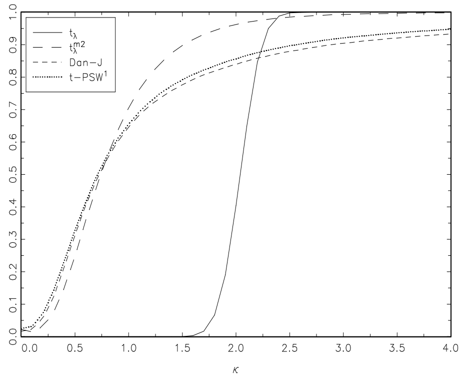
Figure 6. Asymptotic size and local power of s_β: $c = 10$



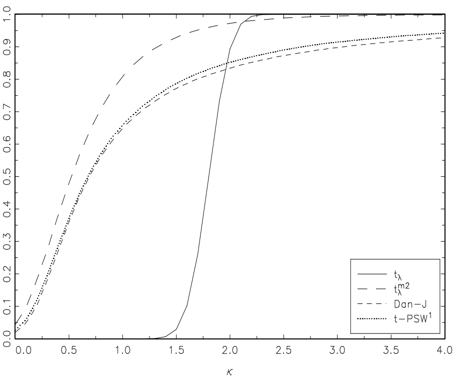
(a) $\alpha = 0$



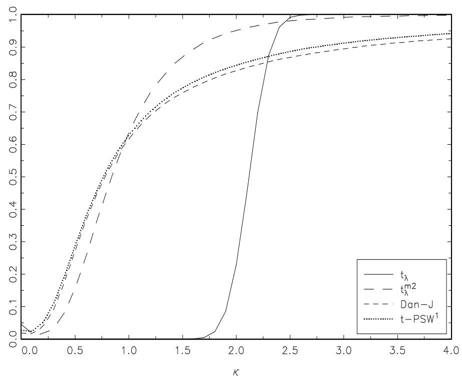
(b) $\alpha = -0.5$



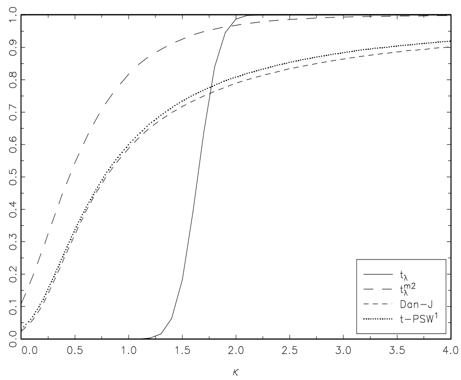
(c) $\alpha = 0.5$



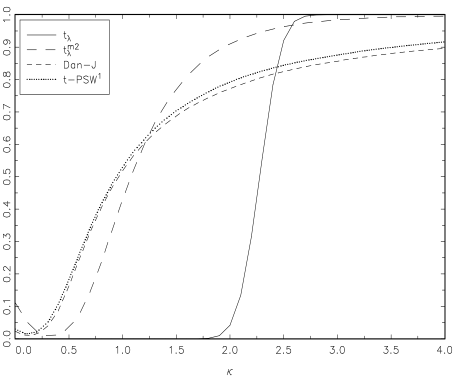
(d) $\alpha = -1$



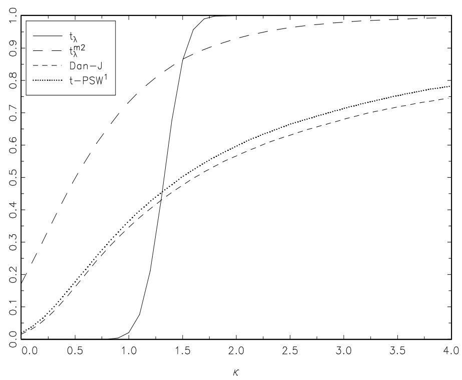
(e) $\alpha = 1$



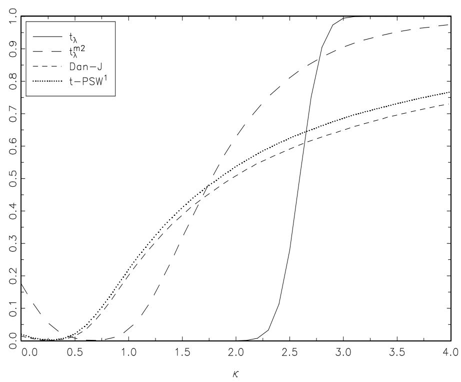
(f) $\alpha = -2$



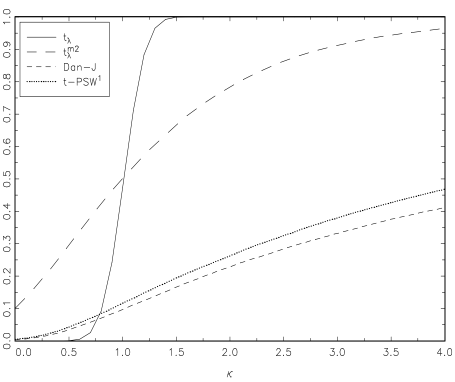
(g) $\alpha = 2$



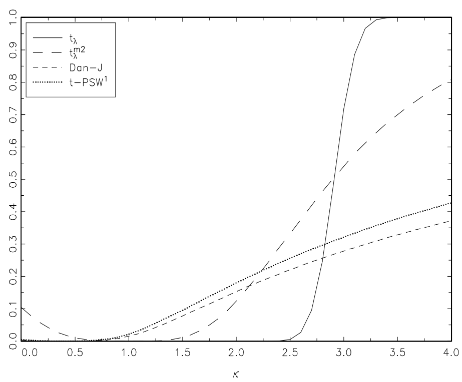
(h) $\alpha = -4$



(i) $\alpha = 4$

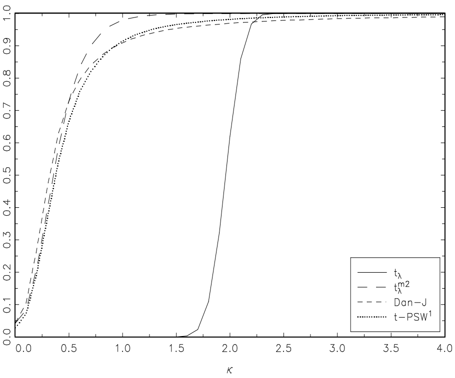


(j) $\alpha = -6$

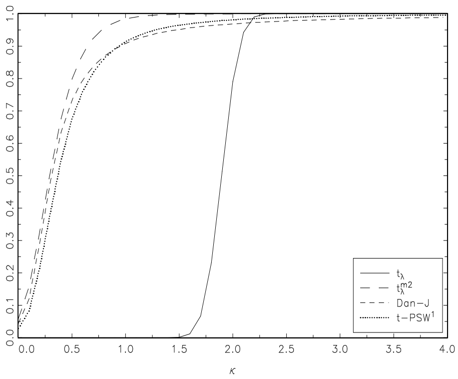


(k) $\alpha = 6$

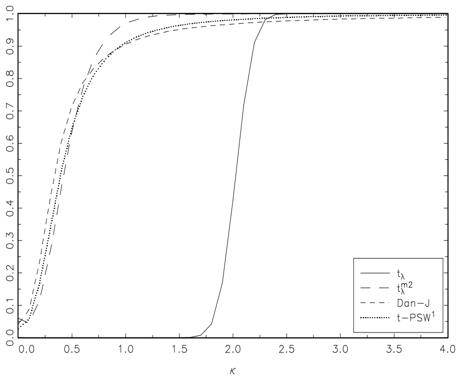
Figure 7. Asymptotic size and local power s_β: $c = 20$



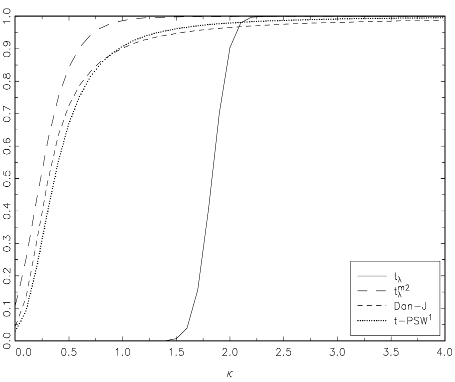
(a) $\alpha = 0$



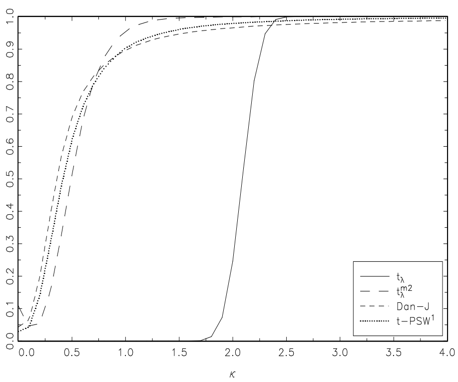
(b) $\alpha = -0.5$



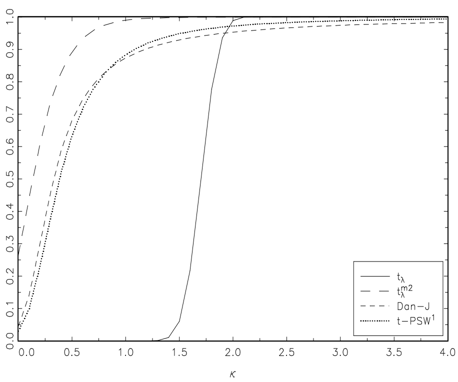
(c) $\alpha = 0.5$



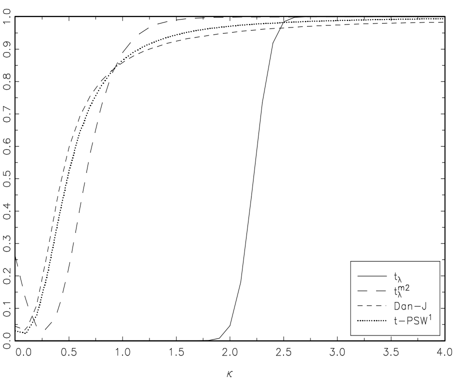
(d) $\alpha = -1$



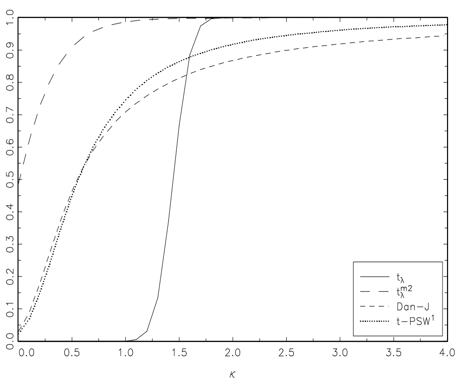
(e) $\alpha = 1$



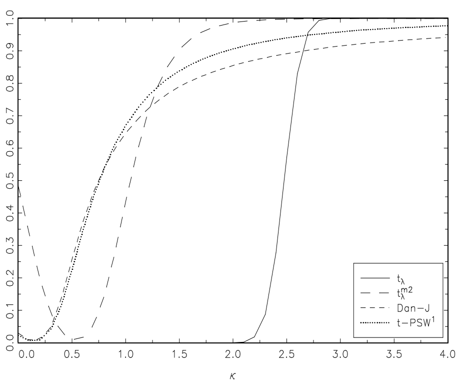
(f) $\alpha = -2$



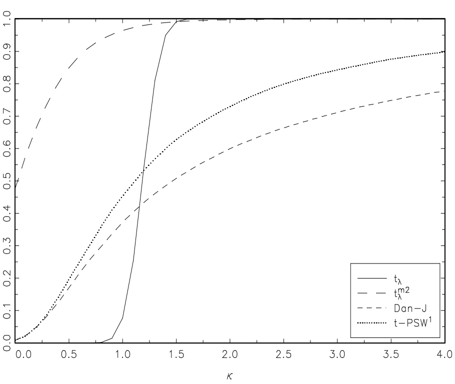
(g) $\alpha = 2$



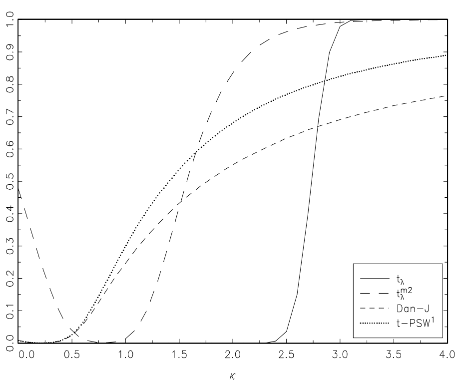
(h) $\alpha = -4$



(i) $\alpha = 4$



(j) $\alpha = -6$



(k) $\alpha = 6$

Figure 8. Asymptotic size and local power of s_β: $c = 30$

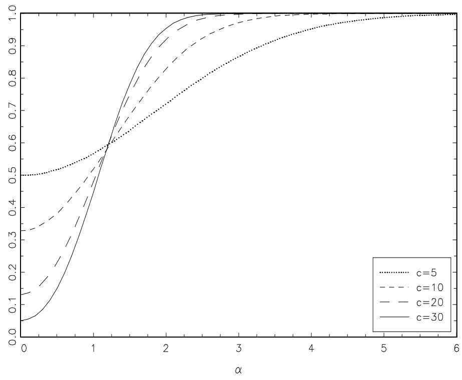
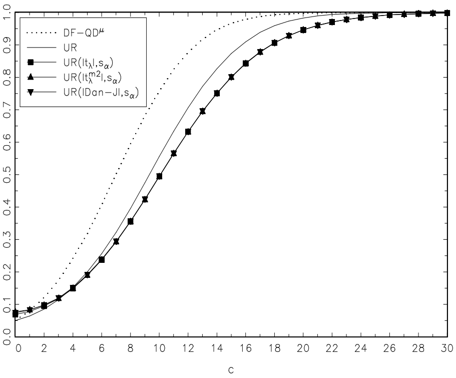
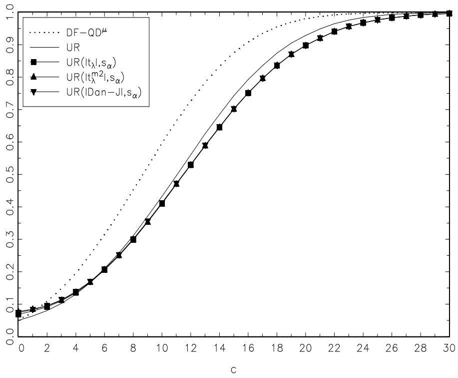


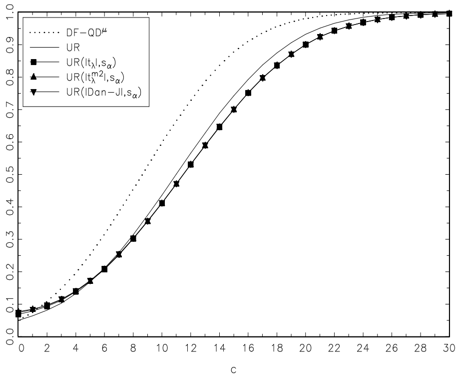
Figure 9. Asymptotic size and local power of s_α



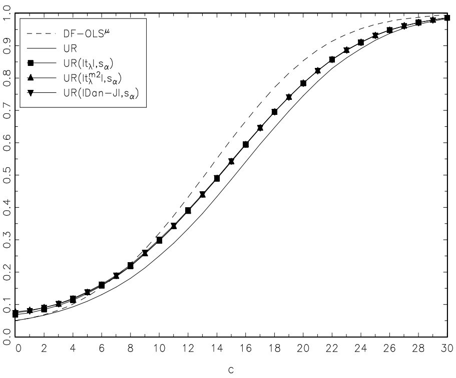
(a) $\alpha = 0$



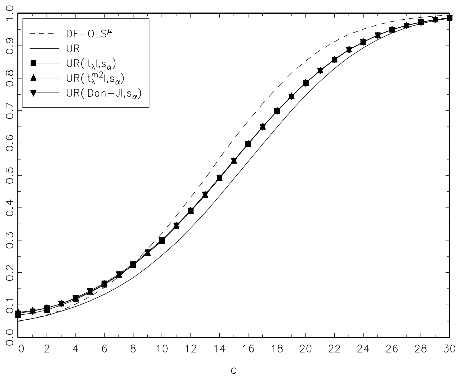
(b) $\alpha = -0.5$



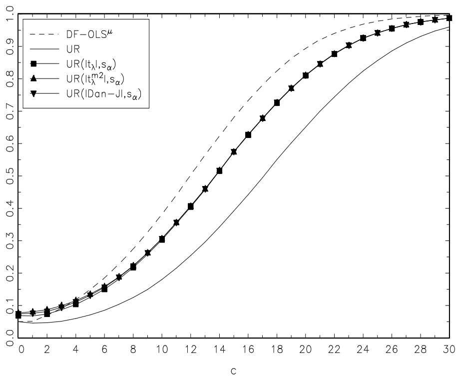
(c) $\alpha = 0.5$



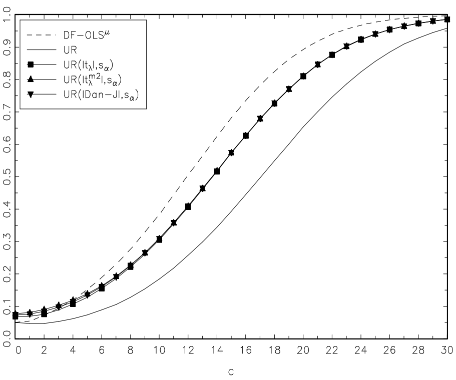
(d) $\alpha = -1$



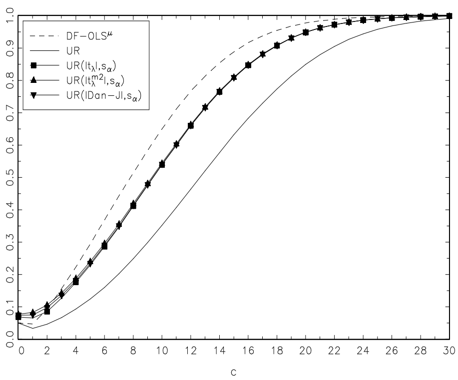
(e) $\alpha = 1$



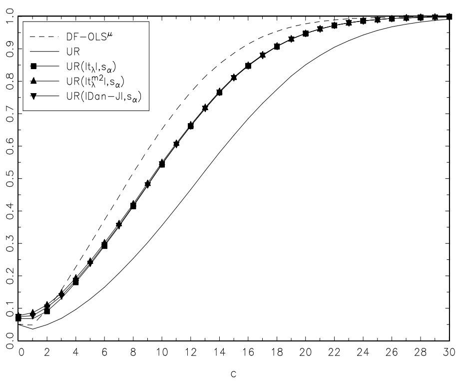
(f) $\alpha = -2$



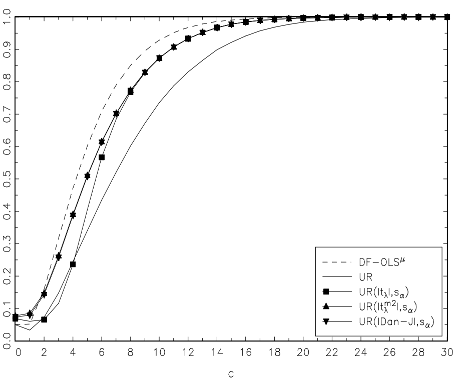
(g) $\alpha = 2$



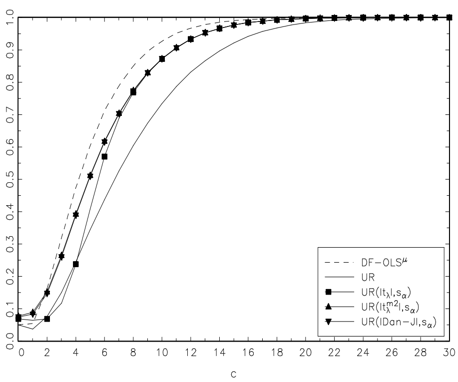
(h) $\alpha = -4$



(i) $\alpha = 4$

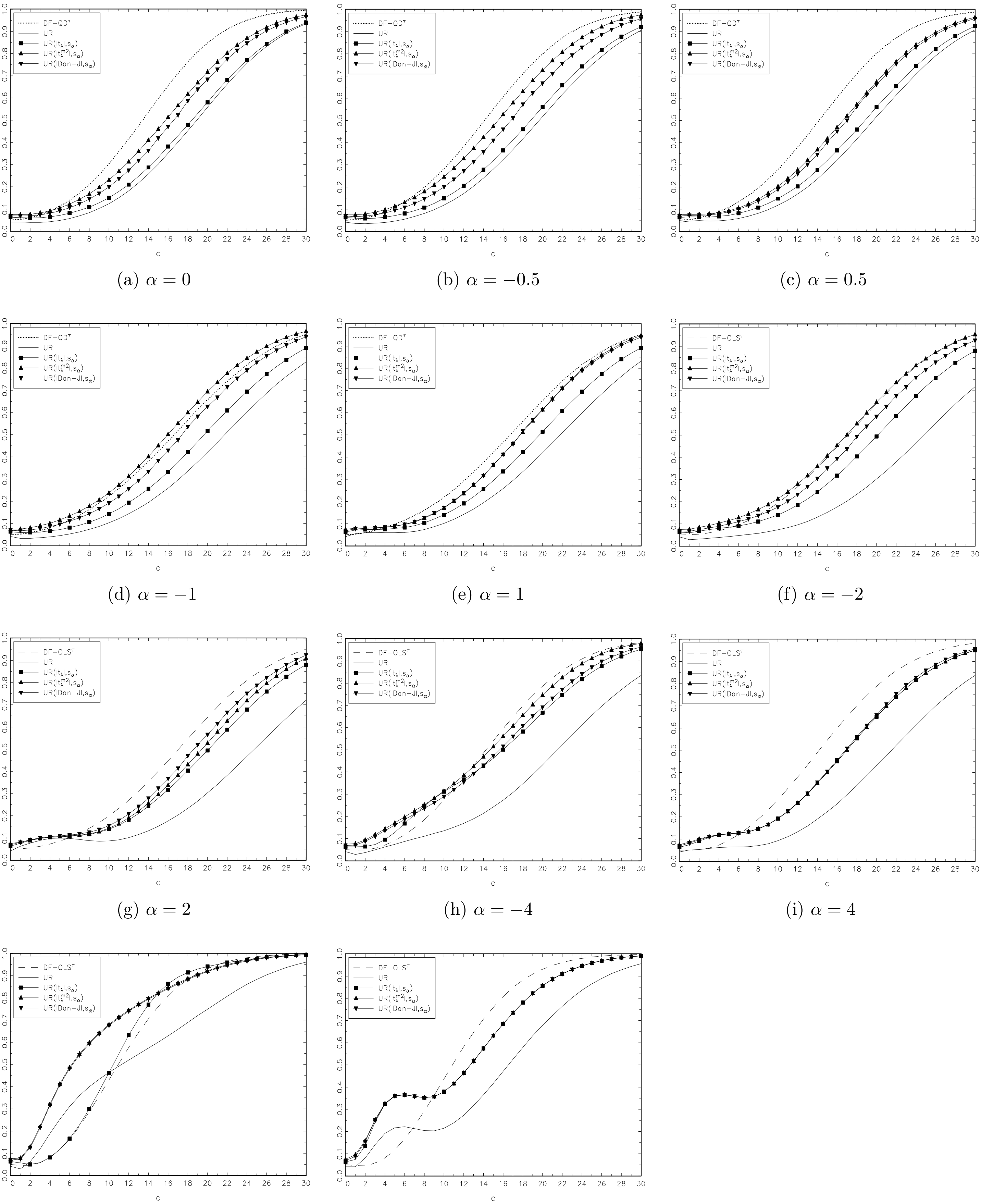


(j) $\alpha = -6$



(k) $\alpha = 6$

Figure 10. Asymptotic size and local power of unit root tests: $\kappa = 0$



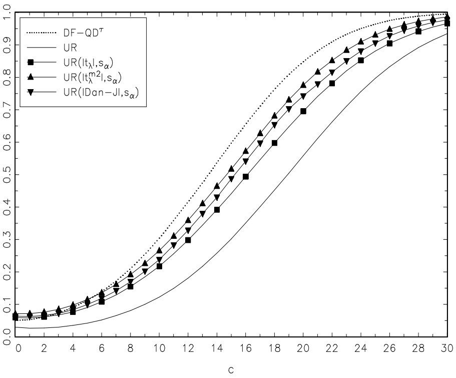
(a) $\alpha = 0$ (b) $\alpha = -0.5$ (c) $\alpha = 0.5$

(d) $\alpha = -1$ (e) $\alpha = 1$ (f) $\alpha = -2$

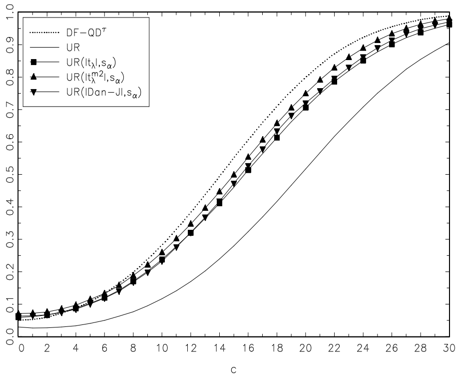
(g) $\alpha = 2$ (h) $\alpha = -4$ (i) $\alpha = 4$

(j) $\alpha = -6$ (k) $\alpha = 6$

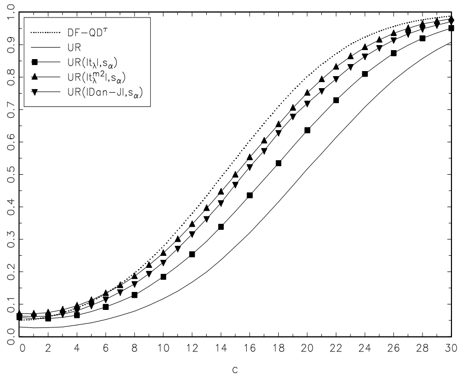
Figure 11. Asymptotic size and local power of unit root tests: $\kappa = 1$



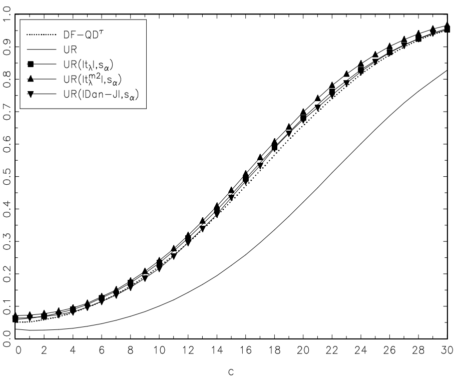
(a) $\alpha = 0$



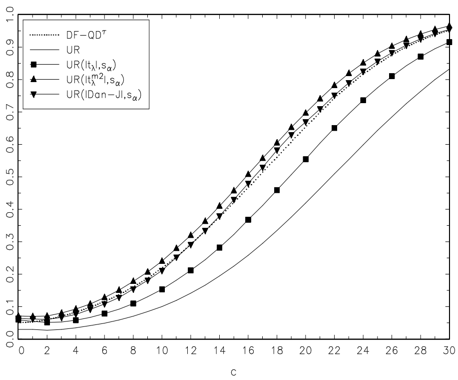
(b) $\alpha = -0.5$



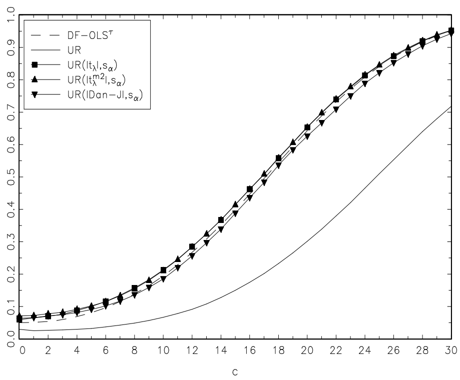
(c) $\alpha = 0.5$



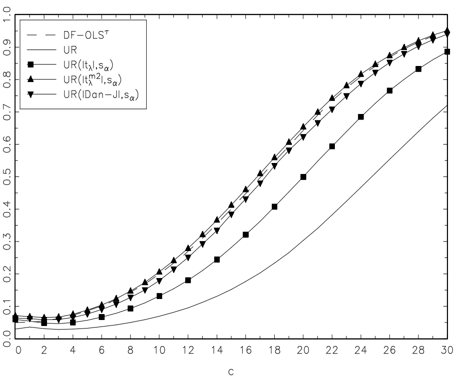
(d) $\alpha = -1$



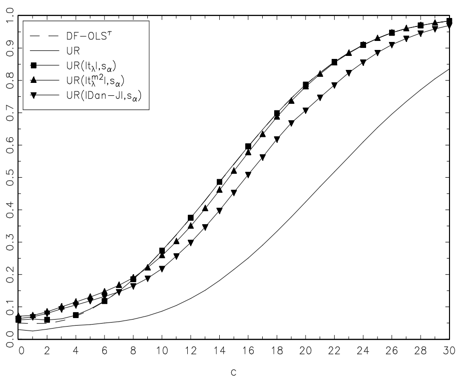
(e) $\alpha = 1$



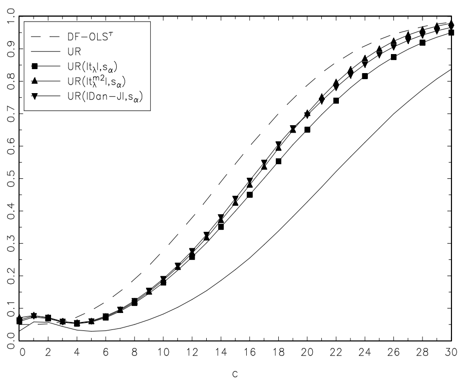
(f) $\alpha = -2$



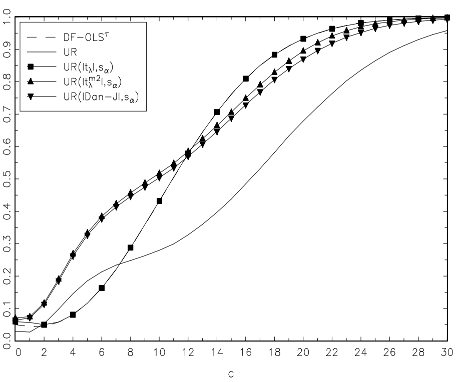
(g) $\alpha = 2$



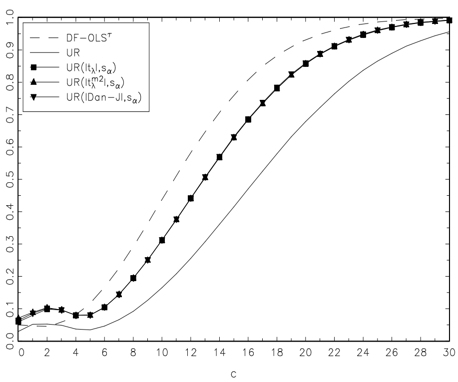
(h) $\alpha = -4$



(i) $\alpha = 4$

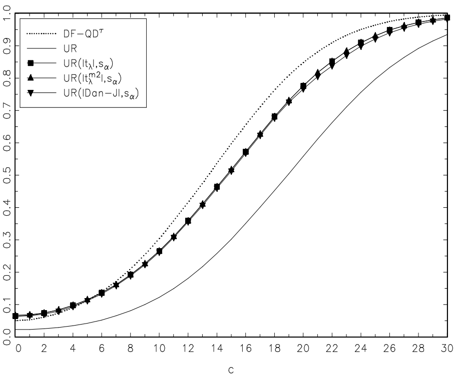


(j) $\alpha = -6$

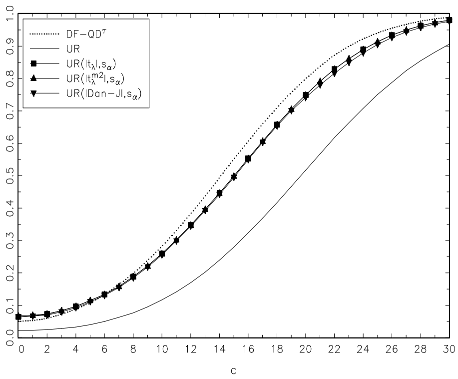


(k) $\alpha = 6$

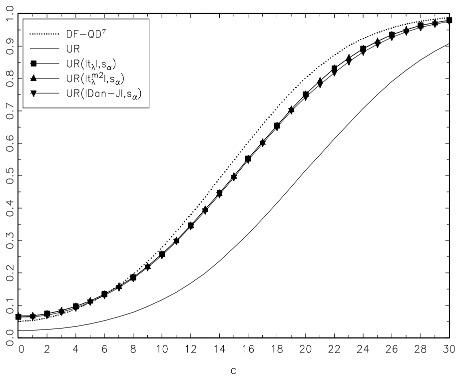
Figure 12. Asymptotic size and local power of unit root tests: $\kappa = 2$



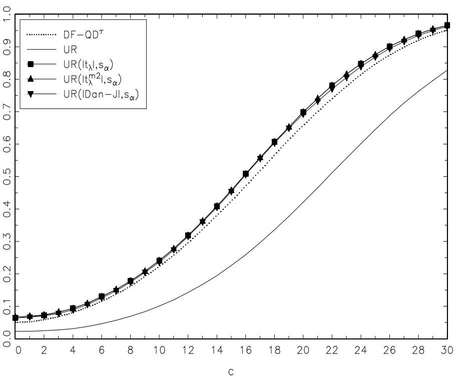
(a) $\alpha = 0$



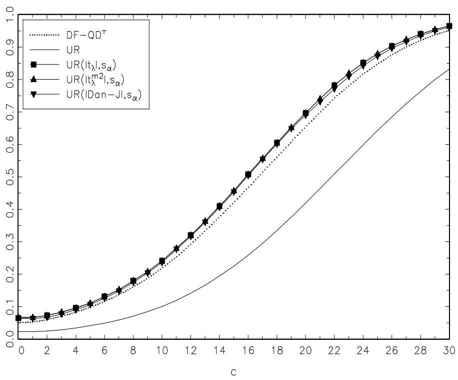
(b) $\alpha = -0.5$



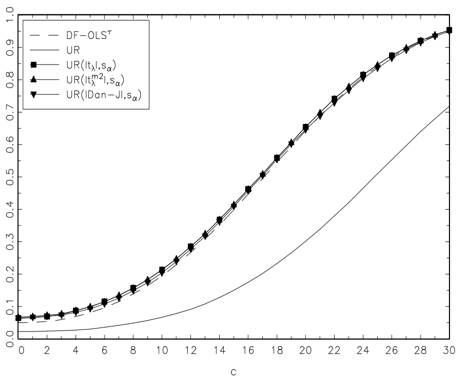
(c) $\alpha = 0.5$



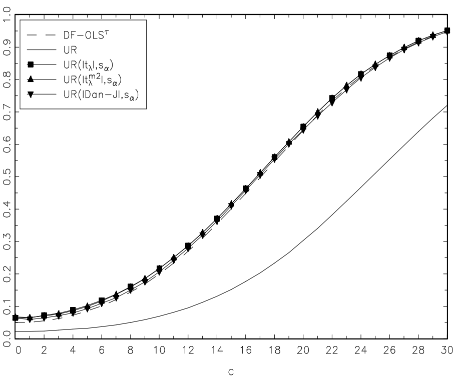
(d) $\alpha = -1$



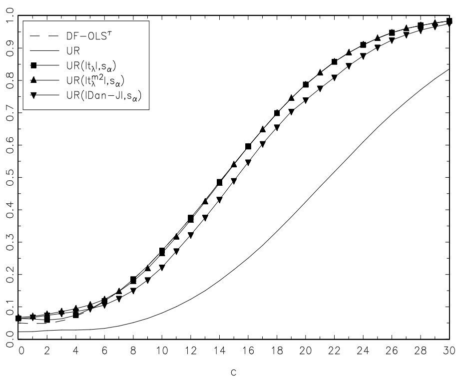
(e) $\alpha = 1$



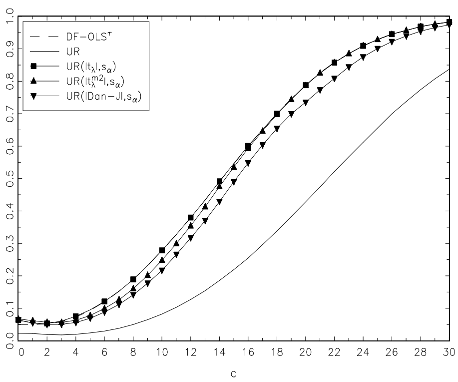
(f) $\alpha = -2$



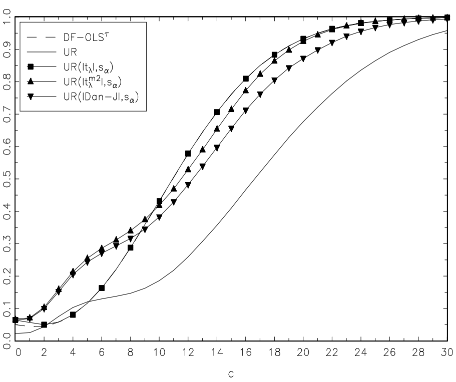
(g) $\alpha = 2$



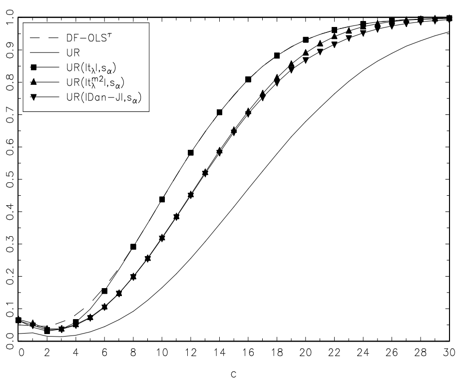
(h) $\alpha = -4$



(i) $\alpha = 4$

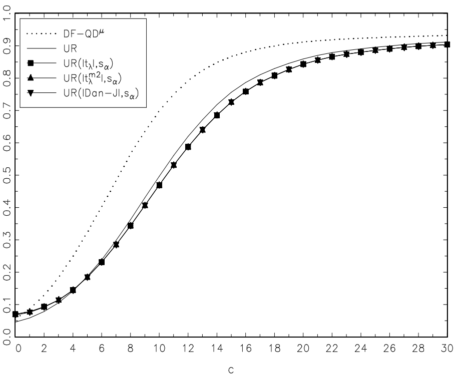


(j) $\alpha = -6$

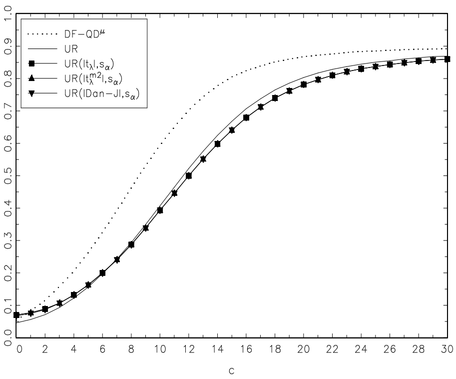


(k) $\alpha = 6$

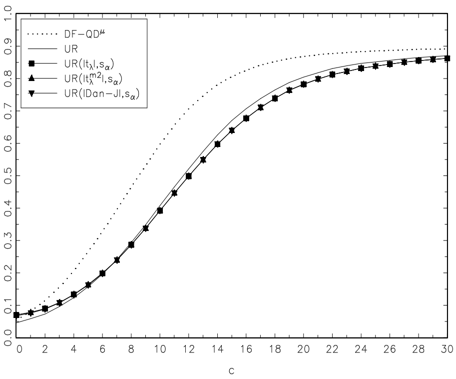
Figure 13. Asymptotic size and local power of unit root tests: $\kappa = 4$



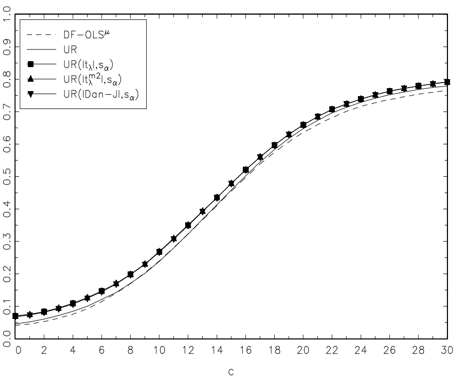
(a) $\alpha = 0$



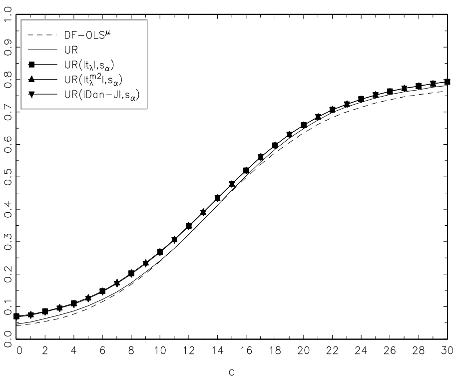
(b) $\alpha = -0.5$



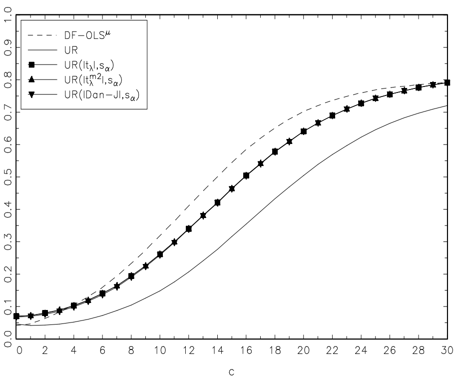
(c) $\alpha = 0.5$



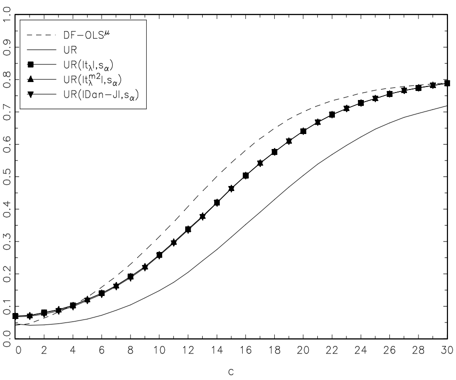
(d) $\alpha = -1$



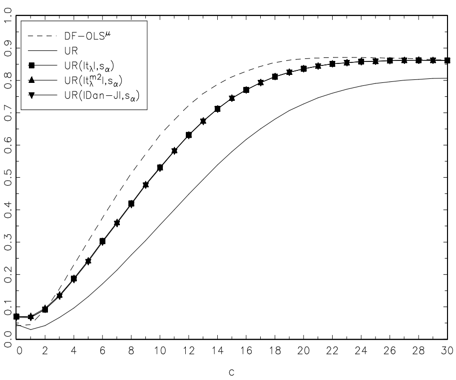
(e) $\alpha = 1$



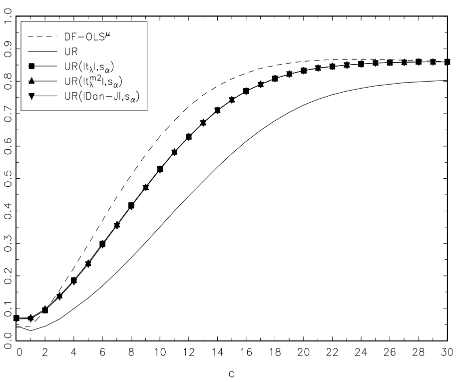
(f) $\alpha = -2$



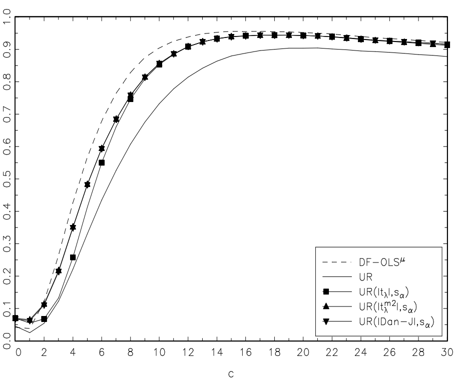
(g) $\alpha = 2$



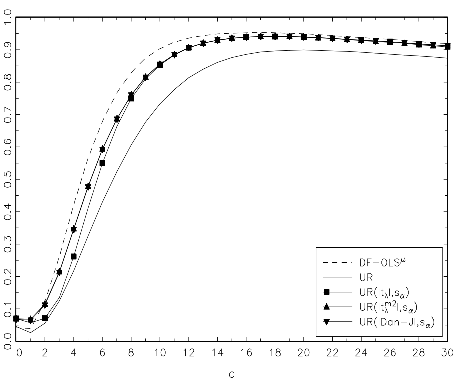
(h) $\alpha = -4$



(i) $\alpha = 4$

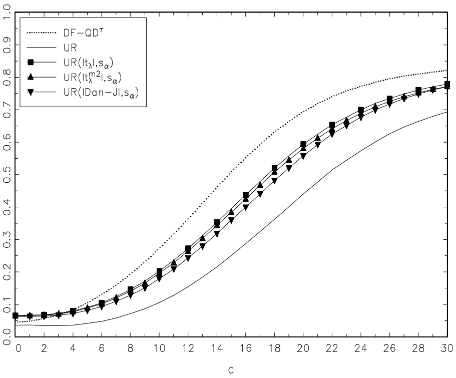


(j) $\alpha = -6$

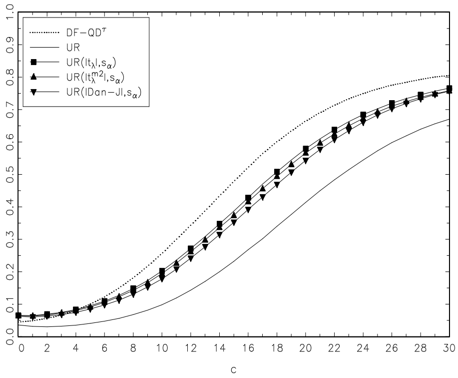


(k) $\alpha = 6$

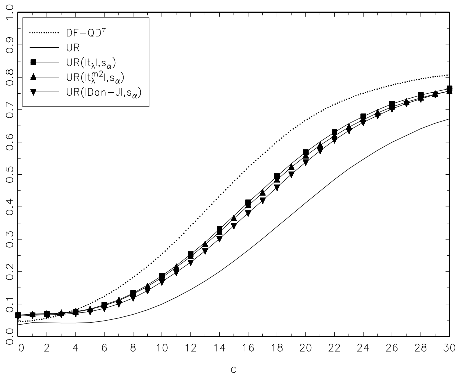
Figure 14. Finite sample size and power of unit root tests: $T = 150$, $\kappa = 0$



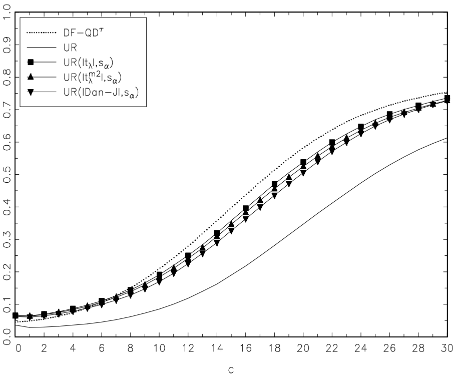
(a) $\alpha = 0$



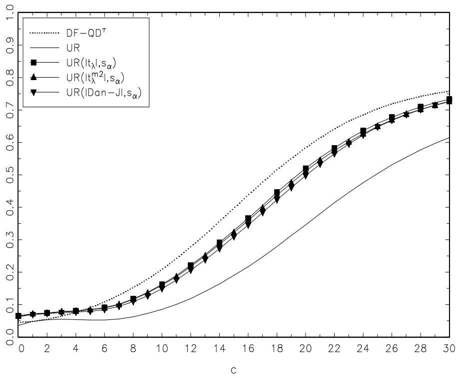
(b) $\alpha = -0.5$



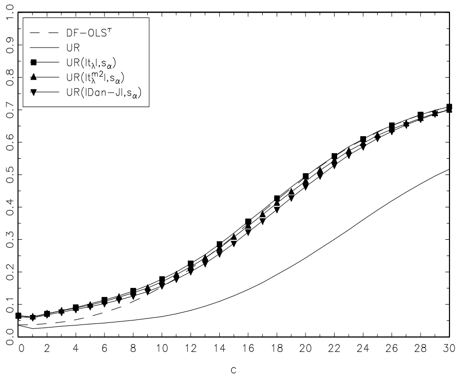
(c) $\alpha = 0.5$



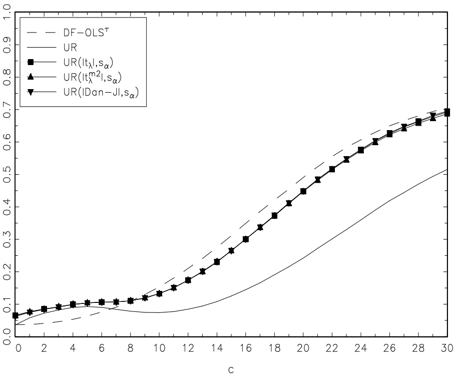
(d) $\alpha = -1$



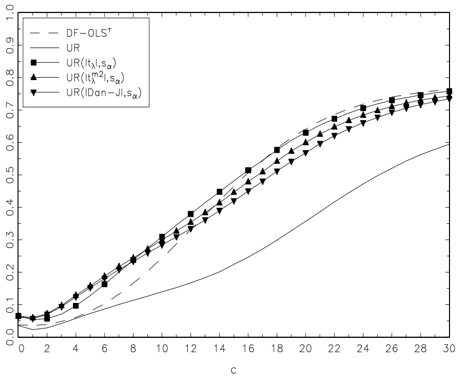
(e) $\alpha = 1$



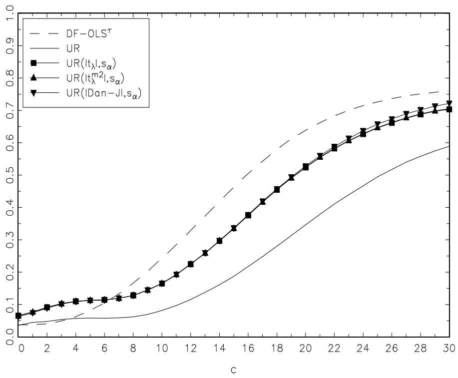
(f) $\alpha = -2$



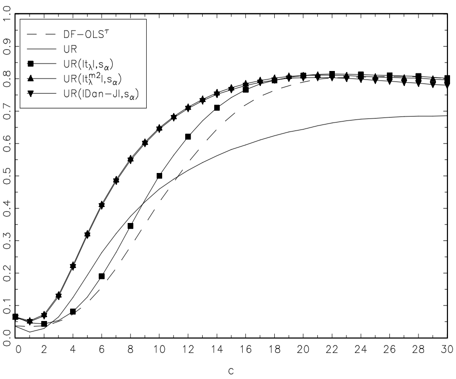
(g) $\alpha = 2$



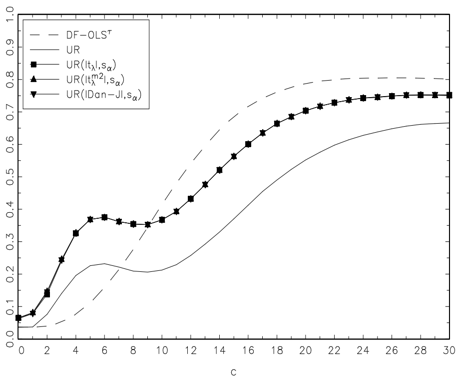
(h) $\alpha = -4$



(i) $\alpha = 4$

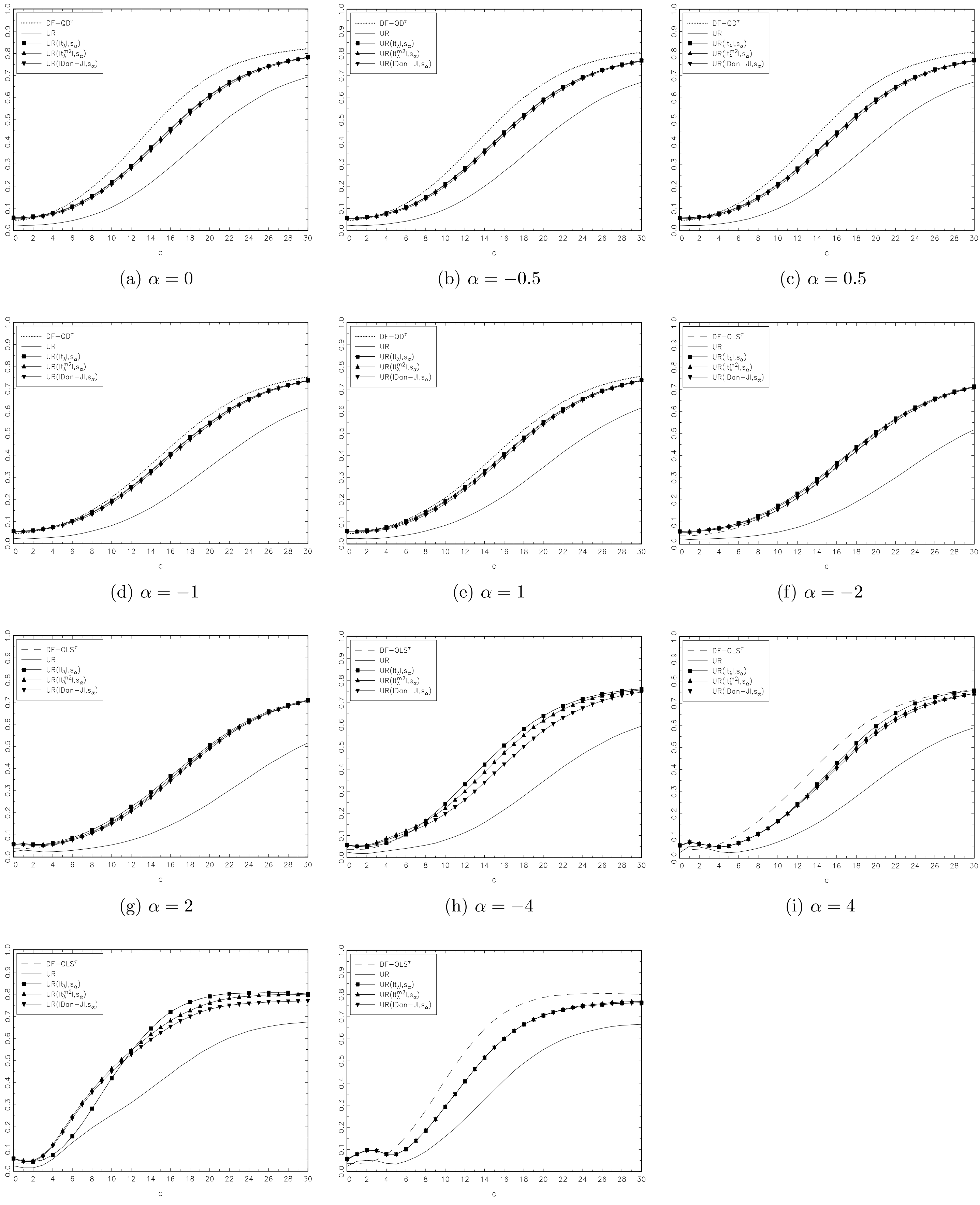


(j) $\alpha = -6$



(k) $\alpha = 6$

Figure 15. Finite sample size and power of unit root tests: $T = 150$, $\kappa = 1$



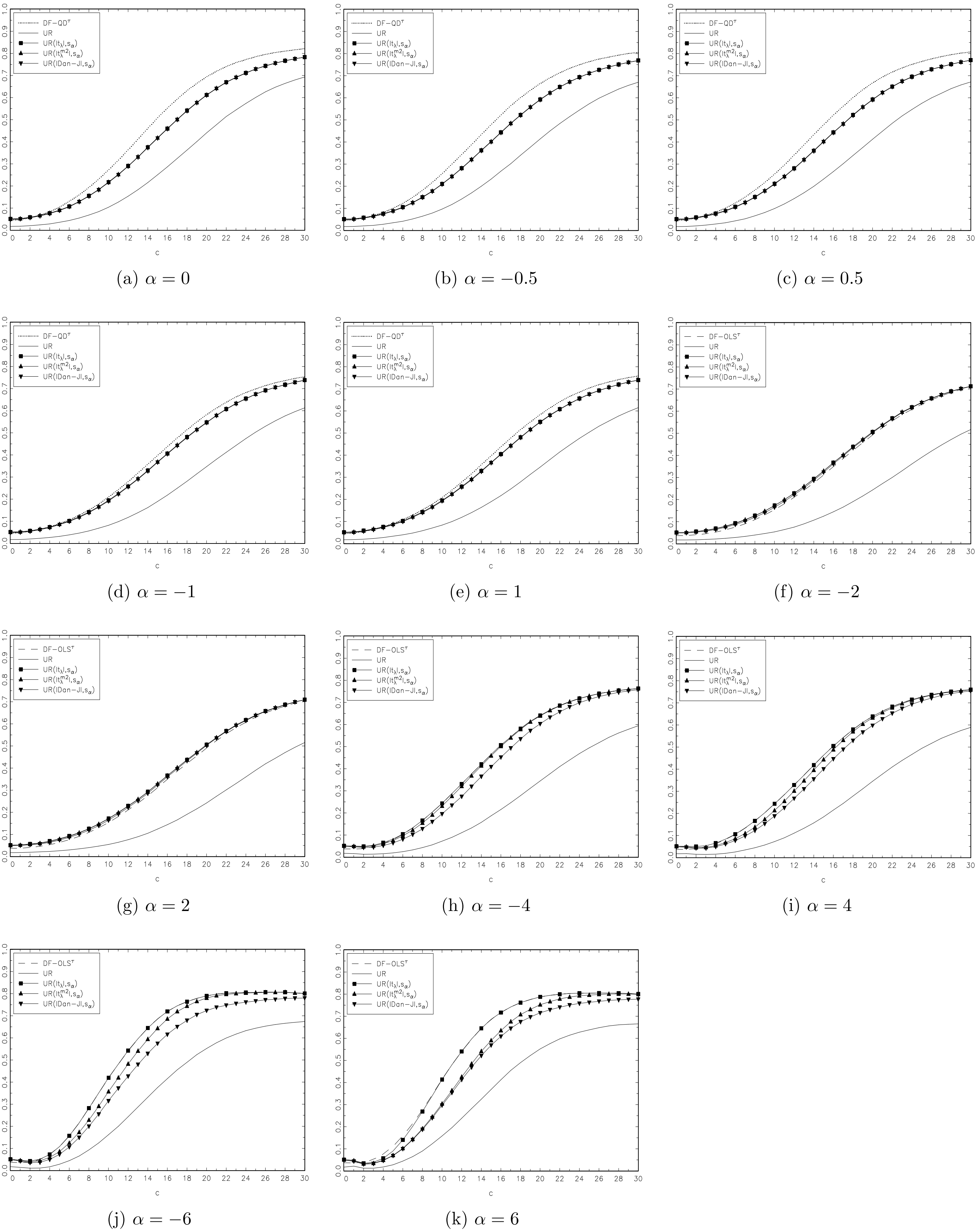
(a) $\alpha = 0$ (b) $\alpha = -0.5$ (c) $\alpha = 0.5$

(d) $\alpha = -1$ (e) $\alpha = 1$ (f) $\alpha = -2$

(g) $\alpha = 2$ (h) $\alpha = -4$ (i) $\alpha = 4$

(j) $\alpha = -6$ (k) $\alpha = 6$

Figure 16. Finite sample size and power of unit root tests: $T = 150$, $\kappa = 2$



(a) $\alpha = 0$ (b) $\alpha = -0.5$ (c) $\alpha = 0.5$

(d) $\alpha = -1$ (e) $\alpha = 1$ (f) $\alpha = -2$

(g) $\alpha = 2$ (h) $\alpha = -4$ (i) $\alpha = 4$

(j) $\alpha = -6$ (k) $\alpha = 6$

Figure 17. Finite sample size and power of unit root tests: $T = 150$, $\kappa = 4$