UNIVERSITY OF NOTTINGHAM

Discussion Papers in Economics

Discussion Paper
No. 01/02

UNIT ROOT TESTS BASED ON INEQUALITY-RESTRICTED ESTIMATORS

by Tae-Hwan Kim and Paul Newbold

February 2001

DP 01/02

ISSN 1360-2438

UNIVERSITY OF NOTTINGHAM

Discussion Papers in Economics

Discussion Paper
No. 01/02

UNIT ROOT TESTS BASED ON INEQUALITY-RESTRICTED ESTIMATORS

by Tae-Hwan Kim and Paul Newbold

Tae-Hwan Kim is Lecturer and Paul Newbold is Professor, both in the School of Economics, University of Nottingham

February 2001

Unit Root Tests Based on Inequality-Restricted Estimators

Tae-Hwan Kim, Paul Newbold

School of Economics, University of Nottingham
University Park, Nottingham, NG7 2RD, UK

May 2000

Abstract: We consider the possibility of estimating a Dickey-Fuller regression, constraining the autoregressive parameter to be at most one, and imposing prior knowledge of the sign of the drift parameter. In spite of apparently encouraging asymptotic results, it emerges that no feasible test of the unit root null hypothesis with superior finite sample properties follows from such inequality-restricted estimation.

Keywords: Parameter boundary; Prior information

JEL Classification: C21, C22

1 Introduction

In a recent paper, Andrews (1999) derives limiting distributions of inequality-restricted parameter estimators when a parameter is on the boundary of the parameter space, employing as an example a Dickey-Fuller regression. It is well known that the original Dickey-Fuller tests have low power and several extensions designed to improve the power of the tests have been developed. (See Pantula, Gonzalez-Farias and Fuller (1994) for a comparison of those extensions.) In this note we ask whether a Dickey-Fuller type test of improved power can be achieved through inequality- restricted least squares accounting simultaneously for the prior knowledge that the autoregressive parameter is at most one and that any drift is non-negative.

Andrews analyses for the time series Y_t the generating model

$$Y_t = \theta_{10} Y_{t-1} + \theta_{20} t + \theta_{30} + \varepsilon_t \tag{1}$$

where we take ε_t to be $i.i.d(0, \sigma^2)$, an assumption that can be relaxed, and where lagged first differences of the series can be added to the regression. If it can be assumed that any drift is non-negative, it is natural to consider the inequality-restricted estimator imposing the prior information $\{\theta_{10} \leq 1, \theta_{20} \geq 0\}$. The least squares, or quasi-maximum likelihood, estimator is then

$$\hat{\theta}_{RT} \in \arg\max_{\theta} l_T(\theta) \;\; \text{s.t.} \;\; \theta \in \{\theta \in R^3 : \theta_1 \leq 1, \theta_2 \geq 0\} \tag{2}$$

where $l_T(\theta) = -\frac{1}{2}\sum_{t=1}^{T}(Y_t - X_t'\theta)^2; X_t' = (Y_{t-1}, t, 1); \theta' = (\theta_1, \theta_2, \theta_3)$.

Let $\hat{\theta}_{1RT}$ be the first element of $\hat{\theta}_{RT}$ and consider the normalized estimator $T(\hat{\theta}_{1RT} - 1)$, on which a test of the unit root null hypothesis might be based. Under the null hypothesis, two true parameters $(\theta_{10}, \theta_{20})$ are on the boundary of the parameter space. When some true parameters are on the boundary, the classical Taylor expansion to derive the asympotic distribution cannot be used because the objective function is not differentiable at the true boundary parameters (Amemiya, 1985). Nor does the bootstrap provide a solution. Andrews (2000) gives an example in which the bootstrap method is not valid when some true parameters are on the boundary. Stochastically approximating the objective function by a quadratic function, Andrews (1999) proves that under the null hypothesis, where $\theta_{10} = 1, \theta_{20} = 0$, and given the prior of non-negative drift, $\theta_{30} \geq 0$,

$$B_T(\hat{\theta}_{RT} - \theta_0) \Rightarrow \hat{\lambda}$$

where B_T is a scaling matrix and $\hat{\lambda}$ is a solution to some quadratic minimization problem. The distribution of $\hat{\lambda}$ can be easily simulated by approximating some functionals of a standard Brownian motion using the Quadratic

Programming Subroutine of GAUSS. In fact, for any $\theta_{30} > 0$, the lower tail area of the limiting null distribution is the same as the Dickey-Fuller distribution - that is, the same as that of $T(\hat{\theta}_{1T} - 1)$ from ordinary least squares estimation of (1) without any restriction. Hence, there is no difference in the critical values. However, for $\theta_{30} = 0$, a different asymptotic null distribution results. We now explore the case of finite sample sizes T, considering the practicalities of basing a unit root test on the estimator (2).

2 Finite sample properties of the test

Although the limiting null distribution of the Inequality-Restricted DF (IRDF) statistic $T(\hat{\theta}_{1RT} - 1)$ is the same for all $\theta_{30} > 0$, ideally one would like the same results to hold in finite samples for practical test implementation. We simulated series from the generating model (1) with $\theta_{10} = 1, \theta_{20} = 0$, and a range of values of $\theta_{30} \geq 0$, taking ε_t to be standard normal. Figure 1 shows 1%, 2.5%, 5%, and 10% critical values, estimated from 200,000 replications for T = 25, 50, 100, 250. It is noticeable that these critical values do depend on the values of the drift parameter θ_{30}, and that the discontinuity at $\theta_{30} = 0$ in the asymptotic null distribution is not evident in the finite sample counterparts. Note also that, as θ_{30} increases, the critical values quickly approach limits. Those limits are in fact the corresponding finite sample critical values of the standard Dickey-Fuller statistic $T(\hat{\theta}_{1T} - 1)$. It thus appears that any practical implementation of the test employing finite sample critical values would require knowledge of the parameter θ_{30} under the null hypothesis, and presumably therefore of $E(\Delta Y_t)$ more generally. Of course, were such information truly available, it could be used at the estimation stage.

We now assess the power, or rejection probabilities, of five variants of the tests. These are:
(i) The usual Dickey-Fuller test, based on $T(\hat{\theta}_{1T} - 1)$, and using empirical finite sample critical values taken from Fuller (1996), denoted DF_e.
(ii) The test based on $T(\hat{\theta}_{1T} - 1)$, but compared with the asymptotic Dickey-Fuller critical values, denoted DF_a. This is considered for two reasons. First, while the asymptotic critical values are valid for any distribution of ε_t, finite sample null distributions depend on the distribution of ε_t, and strictly speaking the published critical values are only correct when ε_t is normal. Second, in this way we can compare with a feasible test based on the inequality-restricted estimator.
(iii) A test based on $T(\hat{\theta}_{1RT} - 1)$ employing empirical critical values computed by the authors, as in Figure 1, in effect assuming $E(\Delta Y_t)$ - that is, θ_{30} under the null hypothesis - is known. Of course, in practice this test is infeasible, but its evaluation allows us to assess size-adjusted powers of

an approach based on inequality-restricted estimation, abstracting from the issue of unavailability of critical values. We denote this test $IRDF_e$.
(iv) A test, $IRDF_a$, in which $T(\hat{\theta}_{1RT}-1)$ is compared with the same critical values as in (ii), as would be correct asymptotically for a null generating process with positive drift.
(v) A final possibility, denoted $IRDF_c$, is comparison of $T(\hat{\theta}_{1RT}-1)$ with the same critical values as in (i), as would be appropriate under the null as $\theta_{30} \rightarrow \infty$.

Time series were generated from the model

$$Y_t = \mu + \beta t + Z_t; \; Z_t = \theta_{10} Z_{t-1} + \varepsilon_t \tag{3}$$

where ε_t are $i.i.d.N(0,1)$. In terms of model (1) then, $\theta_{20} = \beta(1-\theta_{10})$ and $\theta_{30} = \mu(1-\theta_{10}) + \beta\theta_{10}$. The advantage of the formulation (3) is that the interpretation $\beta = E(\Delta Y_t)$ holds under the null and alternative hypotheses (and corresponds, of course, to θ_{30} under the former). Without loss of generality, we can take μ in (3) to be zero. Figure 2 shows estimated rejection probabilities for the five tests, based on 5,000 replications, for a range of values of θ_{10} and β, and for the same sample sizes as in Figure 1, in each case testing at a nominal 5% significance level.

For small values of β, the $IRDF_e$ test, which by construction is correctly sized, has noticeable power advantages, though compared with DF_e and $IRDF_c$ (which use common critical values) these advantages evaporate as β increases. This is to be expected as prior "knowledge" that β is non-negative will only be of value when data are relatively uninformative on this point. So, there appear to be occasions where the use of inequality-restricted estimation is potentially valuable. Unfortunately, however, as already noted, the $IRDF_e$ test is practically infeasible, as the critical values depend on a parameter whose value will be unknown. Moreover, as is clear from Figure 1, the conditions under which inequality-restricted estimation is most valuable, where the drift parameter is relatively small, are precisely those in which the asymptotic critical values are least appropriate. The $IRDF_e$ and $IRDF_a$ tests are both feasible approaches based on inequality-restricted information. However, the results of Figure 2 fail to reveal any grounds for preferring them in comparison with their standard Dickey-Fuller counterparts.

3 Model with constant and no trend

Consider the simpler variant of the Dickey-Fuller test in which the alternative hypothesis is of stationarity around an unknown fixed mean. One might have prior information that this mean is non-negative. In that case, the model

$$Y_t = \theta_{10} Y_{t-1} + \theta_{20} + \varepsilon_t$$

could be estimated subject to the constraints $\{\theta_{10} \leq 1, \theta_{20} \geq 0\}$, and a test based on the statistic $T(\hat{\theta}_{1RT} - 1)$. Using the same approach as in Andrews (1999), it is straightforward to show that the limiting null distribution of this statistic is free of nuisance parameters. However, in simulation experiments not reported here we have found that, under the null hypothesis where $\theta_{10} = 1, \theta_{20} = 0$, the finite sample distribution of the test statistic based on inequality-restricted estimation depends heavily on the initial value Y_0 of the series. This parallels the results reported in Figure 1 for the estimation of model (1).

4 Summary

Prompted by the results of Andrews (1999), we have investigated the possibility of basing a unit root test on inequality-restricted estimation, which might incorporate for example prior knowledge that drift is non-negative. Unfortunately, the finite sample distribution of a test statistic based on such an estimator depends strongly on the amount of drift in the generating process. Hence, a test based on correct critical values is practically infeasible. Moreover, for feasible tests based on inequality-restricted information, we were unable to find any advantages in comparison with corresponding variants of the simpler Dickey-Fuller test.

References

[1] Amemiya, T. (1985) Advanced Econometrics, Cambridge, Mass.: Harvard University Press.

[2] Andrews, D.W.K. (1999) Estimation when a parameter is on a boundary, Econometrica, 67, 1341-1383.

[3] Andrews, D.W.K. (2000) Inconsistency of the bootstrap when a parameter is on the boundary of the parameter space, Econometrica, forthcoming.

[4] Fuller, W.A. (1996) Introduction to Statistical Time Series, 2nd ed., New York: Wiley.

[5] Pantula, S.G., G. Gonzalez-Farias, and W.A. Fuller (1994) A comparison of unit-root test criteria, Journal of Business & Economic Statistics, 12, 449-459.

Figure 1 (a). Empirically estimated critical values

for the IRDF statistic: T = 25

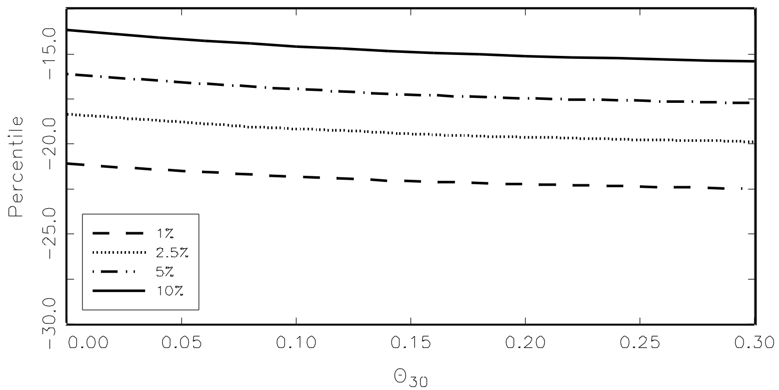

Figure 1 (b). Empirically estimated critical values

for the IRDF statistic: T = 50

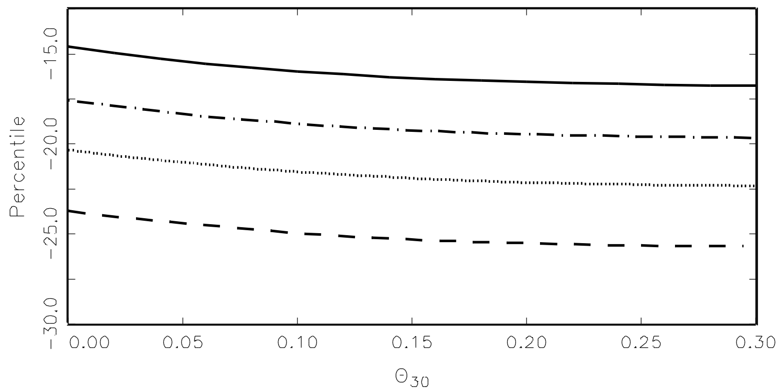

Figure 1 (c). Empirically estimated critical values

for the IRDF statistic: T = 100

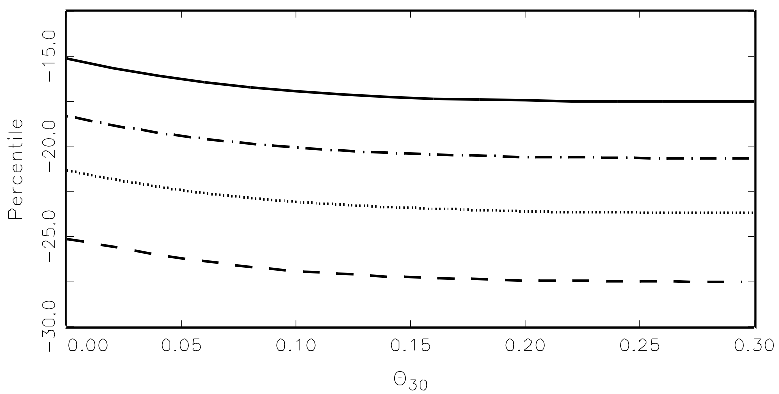

Figure 1 (d). Empirically estimated critical values

for the IRDF statistic: T = 250

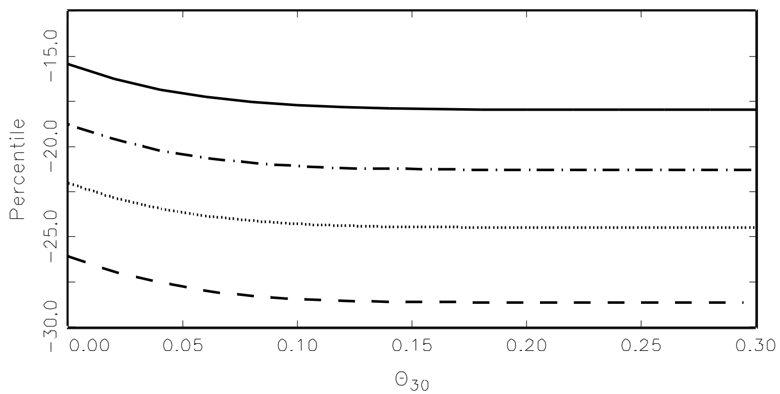

Figure 2 (a). Estimated rejection probabilities at nominal

5%-level for five tests: T = 25, β = 0.01

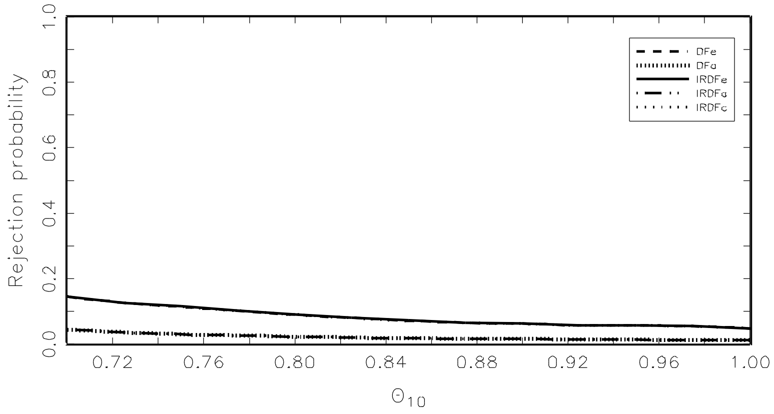

Figure 2 (b). Estimated rejection probabilities at nominal

5%-level for five tests: T = 50, β = 0.01

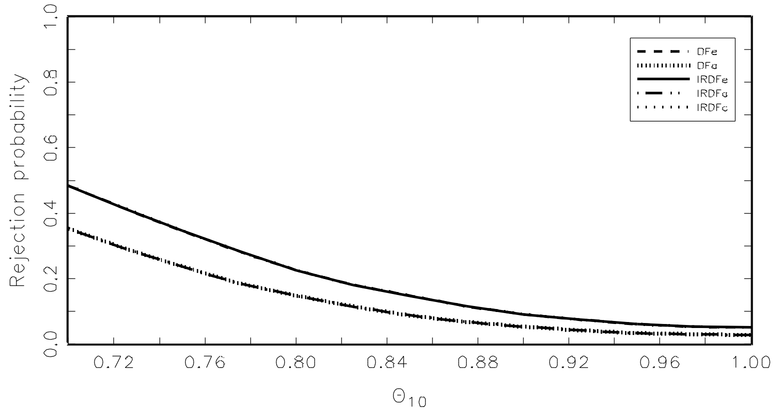

Figure 2 (c). Estimated rejection probabilities at nominal

5%-level for five tests: T = 100, β = 0.01

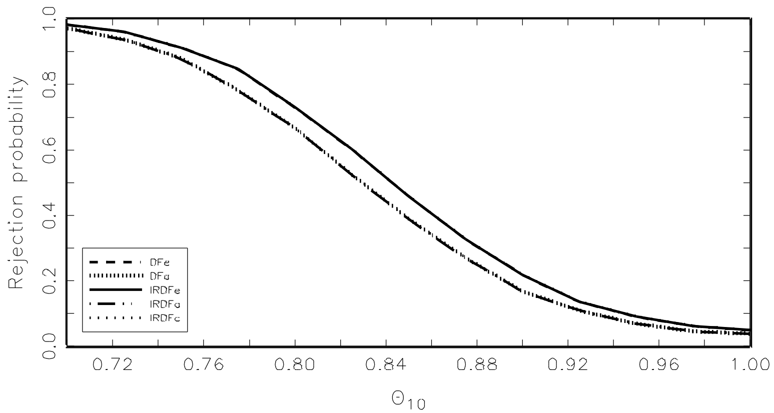

Figure 2 (d). Estimated rejection probabilities at nominal

5%-level for five tests: T = 250, β = 0.01

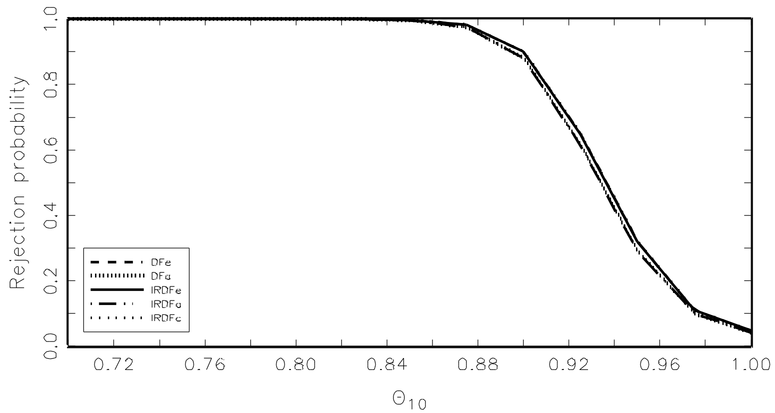

Figure 2 (e). Estimated rejection probabilities at nominal

5%-level for five tests: T = 25, β = 0.1

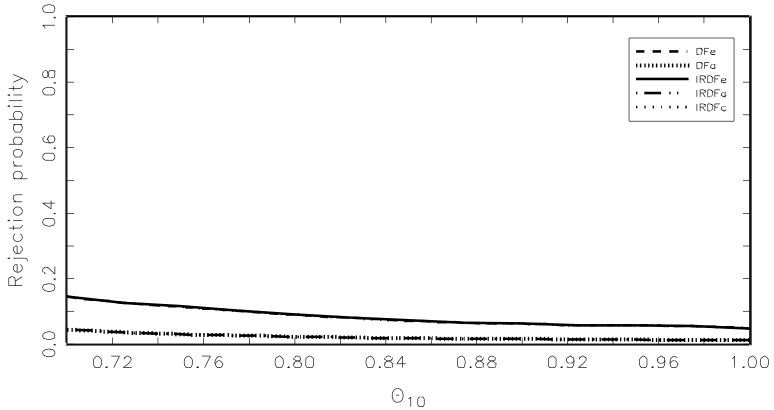

Figure 2 (f). Estimated rejection probabilities at nominal

5%-level for five tests: T = 50, β = 0.1

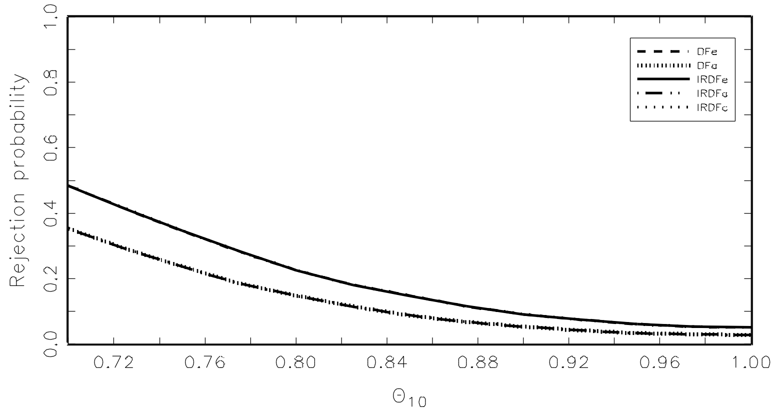

Figure 2 (g). Estimated rejection probabilities at nominal

5%-level for five tests: T = 100, β = 0.1

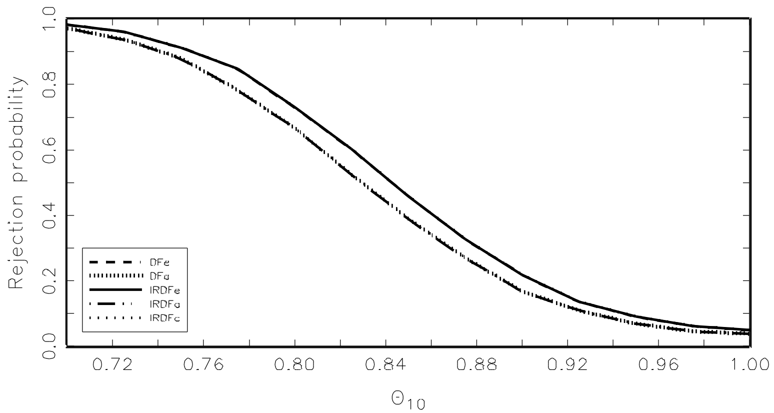

Figure 2 (h). Estimated rejection probabilities at nominal

5%-level for five tests: T = 250, β = 0.1

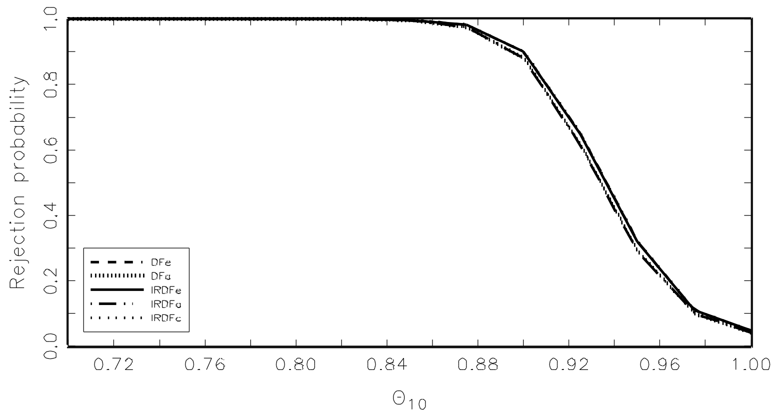

Figure 2 (i). Estimated rejection probabilities at nominal

5%-level for five tests: T = 25, β = 0.5

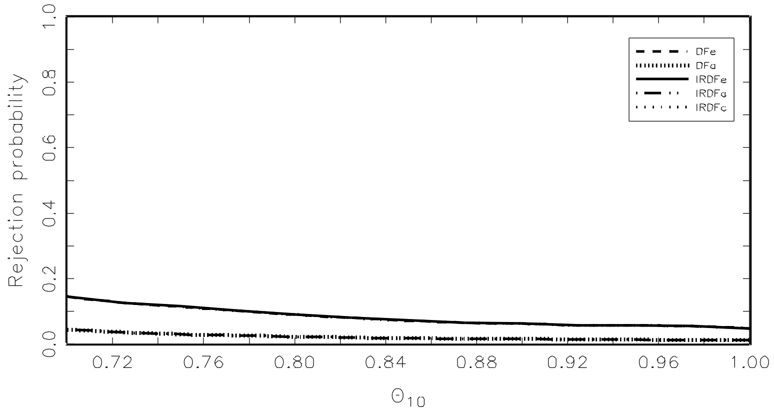

Figure 2 (j). Estimated rejection probabilities at nominal

5%-level for five tests: T = 50, β = 0.5

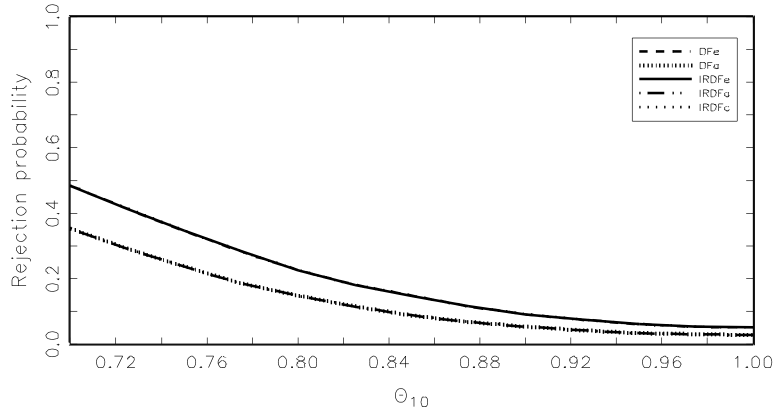

Figure 2 (k). Estimated rejection probabilities at nominal

5%-level for five tests: T = 100, β = 0.5

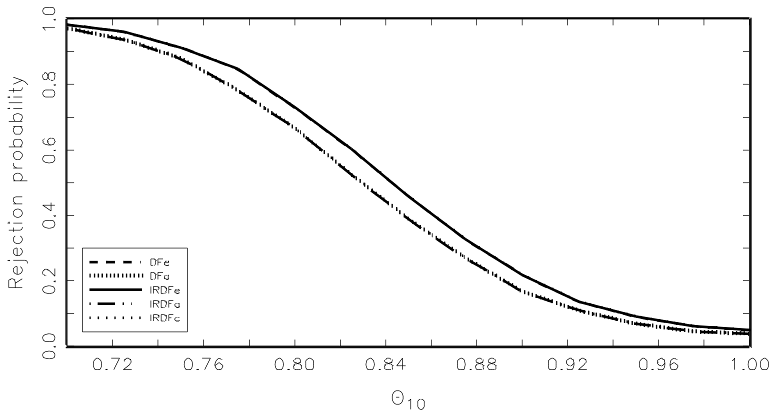

Figure 2 (l). Estimated rejection probabilities at nominal

5%-level for five tests: T = 250, β = 0.5