IMES DISCUSSION PAPER SERIES

Estimating Deterministic Trends with an Integrated or Stationary Noise Component

Pierre Perron and Tomoyoshi Yabu

Discussion Paper No. 2006-E-1

IMES

INSTITUTE FOR MONETARY AND ECONOMIC STUDIES

BANK OF JAPAN

C.P.O BOX 203 TOKYO

100-8630 JAPAN

IMES Discussion Paper Series 2006-E-1
January 2006

Estimating Deterministic Trends with an Integrated or Stationary Noise Component

Pierre Perron* and Tomoyoshi Yabu**

Abstract

We propose a test for the slope of a trend function when it is a priori unknown whether the series is trend-stationary or contains an autoregressive unit root. Let α be the sum of the autoregressive coefficients in the autoregressive representation of the series. The procedure is based on a Feasible Quasi Generalized Least Squares method from an AR(1) specification with parameter α. The estimate of α is the OLS estimate obtained from an autoregression applied to detrended data and is truncated to take a value 1 whenever the estimate is in a $T^{-\delta}$ neighborhood of 1. This makes the estimate "super-efficient" when $\alpha=1$ and implies that inference on the slope parameter can be performed using the standard Normal distribution whether $\alpha=1$ or $|\alpha|<1$. Theoretical arguments and simulation evidence show that $\delta=1/2$ is the appropriate choice. Simulations show that our procedure has good size properties and greater power than the tests proposed by Vogelsang (1998). Applications to inference about the growth rates of GNP for many countries show the usefulness of the method.

Keywords: Linear Trend, Unit Root, Median-Unbiased Estimates, GLS Procedure, Super Efficient Estimates

JEL classification: C22

*Department of Economics, Boston University (E-mail: perron@bu.edu)
**Institute for Monetary and Economic Studies, Bank of Japan (E-mail: tomoyoshi.yabu@boj.or.jp)

This paper is drawn from chapter 1 of Tomoyoshi Yabu's Ph.D. Dissertation at Boston University (See Yabu (2005)). We wish to thank Eiji Kurozumi for useful comments. The views expressed in this paper are those of the authors and do not necessarily reflect the official views of the Bank of Japan or IMES.

1 Introduction

Many time series are well captured by deterministic linear trend. With a logarithmic transformation, the slope of the trend function represents the average growth rate of the time series, a quantity of substantial interest. To be more precise, consider the following model for a time series process $\{y_t\}$:

$$y_t = \mu + \beta t + u_t \tag{1}$$

where u_t is the deviations of the series from the trend. The parameter β is then of primary interest. Hypothesis testing on the slope of the trend function is important for many reasons. First, assessing whether a trend is present is of direct interest. One application, among many, pertains to global warming. Second, Perron (1988) showed that the correct specification of the trend function is important in the context of testing for a unit root in the noise component u_t. He shows that unit root tests that omit a trend when one is present leads to an inconsistent test. On the other hand, including one when not needed leads to a substantial loss of power (e.g., DeJong et al., 1992). Third, it is often of considerable interest to form confidence intervals on the rate of growth of series such as real GNP or other index of real aggregate production. For example, this allows cross-country comparisons or sub-period comparisons to determine if structural changes are present.

There is a large literature on issues pertaining to inference about the slope of a linear trend function, most of it related to the case where the noise component is stationary, i.e. integrated of order zero, $I(0)$. A classic result due to Grenander and Rosenblatt (1957) states that the estimate of β obtained from a simple least-squares regression of the form (1) is asymptotically as efficient as that obtained from a Generalized Least Squares (GLS) regression when the process for u_t is correctly specified. However, it is now recognized that many economic time series of interest are potentially characterized as having a noise component u_t with an autoregressive root that is unity, i.e. integrated of order one, $I(1)$ (e.g. Nelson and Plosser, 1982), or a root that is close to one. In the former case, the least square estimate of β obtained from (1) is no longer asymptotically efficient, but the sample mean of the first-differenced series is. In the case of a root close to one, the standard Grenander and Rosenblatt (1957) result still holds but the limiting Normal distribution is a poor approximation in sample sizes of interest. In most practical applications, the noise component is either $I(0)$ or $I(1)$ and in general no a priori knowledge about which holds is available. The limiting distribution of test statistics depends on the $I(0)$ or $I(1)$ dichotomy so that methods of inference that are robust to both possibilities are needed. Consider the standard t-statistic based on the OLS regression. The t-statistic for testing β converges in distribution to a $N(0,1)$ random variable when u_t is $I(0)$. On the other hand, as shown by Phillips and Durlauf (1988), the t-statistic normalized by $T^{-1/2}$, has a non-degenerate non-Normal limiting distribution when the errors are $I(1)$. This dichotomy is one of the major source of problems that makes inference about the slope of the trend function a difficult issue.

Several papers have tackled the issue of constructing tests and confidence intervals about the parameter β when it is not known a priori whether the noise component u_t is $I(1)$ or not. Sun and Pantula (1999) proposed a pre-test method which first applies a test of the unit root hypothesis and then chooses the critical value to be used for the t-statistic according to the outcome of the test. Using this method, however, the probability of using the critical values from the $I(0)$ case does not converges to zero when the errors are $I(1)$, and the simulations reported accordingly show substantial size distortions. Canjels and Watson (1997) consider various Feasible GLS methods. Their analysis is, however, restricted to the cases where u_t is either $I(1)$ or the autoregressive root is local to one (i.e., $u_t = \alpha_T u_{t-1} + e_t$ with $\alpha_T = 1 + c/T$

and e_t being $I(0)$). Hence, they do not allow $I(0)$ processes for the noise. Also, even with $I(1)$ or near $I(1)$ processes, their method yields confidence intervals that are substantially conservative with common sample sizes. Roy et al. (2004) consider a test based on a one-step Gauss Newton regression that uses a truncated weighted symmetric least-squares estimate of the autoregressive parameter (as suggested by Roy and Fuller, 2001). The limit distribution of their test is not the same in the $I(1)$ and $I(0)$ cases, but their simulations show that the size is similar in finite samples. Vogelsang (1998) is the only one, to our knowledge, who proposes a test valid with either $I(1)$ or $I(0)$ errors (see section 2.4) but its power is very low in the $I(1)$ case. He then advocates the use of an additional statistic which have complementary properties with the usual drawback induced by the use of multiple tests.

We propose a new robust test statistic that is valid with either $I(0)$ or $I(1)$ noise in the sense that the limit distribution is the same in both cases. It applies a Feasible GLS procedure with an estimate of the sum of the autoregressive coefficients that is truncated to take a value of 1 when the usual estimate is in a neighborhood of 1. Hence, it is super-efficient when the true value is 1. As a result, the limiting distribution of our statistic does not depend on whether the noise is I(0) or I(1) and is the usual standard Normal. Theoretical arguments are provided to show that the size of the neighborhood where truncation applies should be of order $T^{1/2}$. Also, to improve the finite sample performance, we advocate the use of a median unbiased estimate of the sum of the autoregressive coefficients. The resulting statistic is easy to implement, its size is close to the nominal level in finite samples and its power exceeds that of previously suggested procedures.

The outline of the paper is as follows. In section 2, we analyze the simple AR(1) case and lay out the basic framework, the test, its large sample properties and simulations about its finite sample size and power. In section 3, we consider extensions to a general class of processes for u_t. Section 4 considers the usefulness of our test in the context of the proper specification of the trend component when conducting unit root tests. Section 5 contains applications of our procedure to the data analyzed by Vogelsang (1998) and Canjels and Watson (1997). Section 6 contains brief concluding comment and an appendix some technical derivations.

2 The AR(1) Case

We begin by considering the leading case with $AR(1)$ errors so that the univariate time series $\{y_t; t = 1, ..., T\}$, is generated by

$$y_t = \mu + \beta t + u_t; \quad u_t = \alpha u_{t-1} + e_t \tag{2}$$

where $e_t \sim i.i.d.(0, \sigma^2)$ and $E(e_t^4) < \infty$. For simplicity we let u_0 be some finite constant. Here $-1 < \alpha \leq 1$ so that both stationary and integrated errors are allowed. The analysis extends easily to more general trend functions of the form $\sum_{i=0}^{n} \beta_i t^i + \sum_{j=1}^{m} \sum_{i=0}^{q} \theta_i (t - T_{B,j})^i 1(t > T_{B,j})$, where $1(\cdot)$ is the indicator function. If $m \neq 0$, m breaks are present occuring at dates $T_{B,j}$ $(j = 1, ..., m)$. When the break dates are known, all our main theoretical results remain valid, though some equations may have a different form. For simplicity of exposition, we consider throughout the simpler first-order linear trend, which is the leading case of interest in practice. The GLS estimator considered is the one that applies Ordinary Least Squares (OLS) to the regression

$$y_t - \alpha y_{t-1} = (1 - \alpha)\mu + \beta[t - \alpha(t - 1)] + e_t \tag{3}$$

for $t = 2, ..., T$, together with $y_1 = \mu + \beta + u_1$. Consider the infeasible GLS estimate of β that assumes a known value of α. It is well known that the t-statistic for testing the null hypothesis that $\beta = \beta_0$, t_β^G, is then asymptotically distributed as a $N(0, 1)$ random variable for any values of α in the permissible range.

Canjels and Watson (1997) considers the case where the initial value u_0 has a variance that can depend on T. Two methods are advocated. First, the estimator described above, which is the GLS estimator assuming $u_0 = 0$ and works best when the variance of u_0 is small. The second is the Prais-Winsten (1954) estimator with the first observation specified by $(1-\alpha^2)^{1/2}y_1 = (1-\alpha^2)^{1/2}\mu + (1-\alpha^2)^{1/2}\beta + (1-\alpha^2)^{1/2}u_1$. This estimator is superior when the variance of u_0 is not small. We shall not be concerned about the effect of the initial condition in this paper and, hence, we have assumed a fixed value for simplicity and will use the specification defined after (3). Nevertheless, all our results remain valid using one or the other GLS estimate.

2.1 The Feasible GLS estimate

Consider the Cochrane-Orcutt (1949) Feasible GLS (FGLS) estimate of β that uses the estimate $\hat{\alpha} = \sum_{t=2}^{T} \hat{u}_t\hat{u}_{t-1} / \sum_{t=2}^{T} \hat{u}_{t-1}^2$where $\{\hat{u}_t\}$ are the OLS residuals from a regression of y_t on $\{1, t\}$. When $|\alpha| < 1$, $T^{1/2}(\hat{\alpha} - \alpha) \rightarrow^d N(0, 1-\alpha^2)$ and from the Grenander-Rosenblatt (1957) result, the OLS and FGLS estimates of β are asymptotically equivalent and the t-statistic from the FGLS procedure, denoted t_β^F, still has a $N(0,1)$ limit distribution. Things are different when $\alpha = 1$. From standard results,

$$T(\hat{\alpha} - 1) \Rightarrow \int_0^1 W^*(r)dW(r) / \int_0^1 W^*(r)^2 dr \ \equiv \kappa \tag{4}$$

with $W^*(r)$, $0 \leq r \leq 1$, the residual process from a continuous time regression of a unit Wiener process $W(r)$ on $\{1, r\}$ and where $\Rightarrow$ denotes weak convergence in distribution under the Skorohod topology. It is shown in the appendix that, provided $T(\hat{\alpha} - 1) = O_p(1)$, the t-statistic from the Feasible GLS procedure is such that

$$\begin{aligned} t_\beta^F \ &= \ \Bigg[[T^{-1/2}\sum_{t=2}^{T} e_t - T(\hat{\alpha}-1)T^{-3/2}\sum_{t=2}^{T} u_{t-1}] - T(\hat{\alpha}-1)[T^{-3/2}\sum_{t=2}^{T} te_t \\ &\quad -T(\hat{\alpha}-1)T^{-5/2}\sum_{t=2}^{T} tu_{t-1}] \Bigg] / \left[\sigma^2(1 - T(\hat{\alpha}-1) + T^2(\hat{\alpha}-1)^2/3)\right]^{1/2} + o_p(1) \end{aligned} \tag{5}$$

Given (4), $T^{-1/2}\sum_{t=2}^{T} e_t \Rightarrow \sigma W(1)$, $T^{-3/2}\sum_{t=2}^{T} u_{t-1} \Rightarrow \sigma \int_0^1 W(r)dr$, $T^{-5/2}\sum_{t=2}^{T} tu_{t-1} \Rightarrow \sigma \int_0^1 rW(r)dr$ and $T^{-3/2}\sum_{t=2}^{T} te_t \Rightarrow \sigma \int_0^1 rdW(r)$, the limit can be expressed as

$$t_\beta^F \Rightarrow \frac{W(1) - \kappa \int_0^1 W(r)dr - \kappa[\int_0^1 rdW(r) - \kappa \int_0^1 rW(r)dr]}{(1 - \kappa + \kappa^2/3)^{1/2}} \tag{6}$$

which is, indeed, different from a Normal distribution. Now suppose that κ, the limit of $T(\hat{\alpha} - 1)$, is zero. Then $t_\beta^F \Rightarrow W(1) =^d N(0,1)$, we would recover in the $I(1)$ case the same limit distribution as in the $I(0)$ case and no discontinuity would be present. Our main idea exploits this feature. The proposed estimate of α is a super-efficient estimate when $\alpha = 1$, defined by

$$\hat{\alpha}_S = \begin{cases} \hat{\alpha} & \text{if} \quad |\hat{\alpha} - 1| > dT^{-\delta} \\ 1 & \text{if} \quad |\hat{\alpha} - 1| \leq dT^{-\delta} \end{cases} \tag{7}$$

for some $\delta \in (0, 1)$ and $d > 0$. Hence, when $\hat{\alpha}$ is in a $T^{-\delta}$ neighborhood of 1 it is assigned a value of 1. When the GLS procedure is applied with $\hat{\alpha}_S$, we denote the resulting t-statistic by t_β^{FS}. The following Theorem, proved in the appendix, shows that this indeed changes nothing in the stationary case and deliver a t-statistic with a Normal limit distribution when $\alpha = 1$.

Theorem 1 *a)* $T^{1/2}(\hat{\alpha}_S - \alpha) \to^d N(0, 1 - \alpha^2)$ *when* $|\alpha| < 1$*; and b)* $T(\hat{\alpha}_S - 1) \to_p 0$ *when* $\alpha = 1$*. Hence* $t_\beta^{FS} \to^d N(0,1)$ *for both* $|\alpha| < 1$ *and* $\alpha = 1$*.*

Constructing the GLS regression using the truncated estimate effectively bridges the gap between the $I(0)$ and $I(1)$ cases, and the Normal limit distribution applies.

2.2 The case with α local to unity

The result obtained in Theorem 1 is pointwise in α with $0 < \alpha \leq 1$ and may not hold uniformly, in particular in a local neighborhood of 1. Hence, it is of interest to see what happens when the true value of α is close to but not equal to one. Adopting the standard local to unity approach, we have the following result proved in the Appendix.

Theorem 2 *Suppose the data are generated by (2) with* $\alpha = \alpha_T = 1 + c/T$ *for some* $c \leq 0$*, then* $T(\hat{\alpha}_S - 1) \to_p 0$ *and* $t_\beta^{FS} \to^d N(0, (\exp(2c) - 1)/2c)$.

The results are fairly intuitive. Since the true value of α is in a T^{-1} neighborhood of 1, and $\hat{\alpha}_S$ truncates values of $\hat{\alpha}$ in a $T^{-\delta}$ neighborhood of 1 for some $0 < \delta < 1$ (i.e., a larger neighborhood), in large enough samples $\hat{\alpha}_S = 1$. Then, the estimate of β is the first-difference estimator, $\hat{\beta}_{FD} = T^{-1} \sum_{t=1}^T \Delta y_t$ and the t-statistic satisfies:

$$t_{FD} = \frac{\hat{\beta}_{FD} - \beta}{std(\hat{\beta}_{FD})} = \frac{T^{-1} \sum_1^T \Delta u_t}{\sqrt{(T^{-1} \sum_1^T \Delta u_t^2) T^{-1}}} = \frac{T^{-1/2} u_T - T^{-1/2} u_0}{\sqrt{T^{-1} \sum_1^T \Delta u_t^2}} \Rightarrow J_c(1)$$

since $T^{-1/2} u_T \Rightarrow \sigma J_c(1)$, $T^{-1/2} u_0 = o_p(1)$, and $T^{-1} \sum_1^T \Delta u_t^2 \to_p \sigma^2$. Here $J_c(1) \equiv \int_0^1 \exp(c(1-s)) dW(s) \sim N(0, (\exp(2c) - 1)/2c)$.

Note that when $c = 0$, we recover the result of Theorem 1 for the $I(1)$ case. However, when $c < 0$, the variance of the limiting distribution is different from 1. In fact, the variance is lower than 1, so that, without modifications, a conservative test may be expected for values of α close to 1, relative to the sample size. Also, one can note that the limit of the variance as $c \to -\infty$ is 0, not 1, and we do not recover the same result that applies to the $I(0)$ case. As noted by Phillips and Lee (1996), the local to unity asymptotic framework with $c \to -\infty$ involves a doubly infinite triangular array such that the limit of the statistic depends on the relative approach to infinity of c and T. The next Theorem shows that, indeed, $t_\beta^{FS} \to^d N(0, 1)$ as $c \to -\infty$. What is especially interesting is that to obtain this result, a condition on δ needs to be imposed. It will turn out that this condition is very useful to guide us to a value that should provide the least size distortions when α is in a neighborhood of 1. The following result is proved in the appendix.

Theorem 3 *Let the data be generated by (2) with* $\alpha = 1 + c/T$*, and* $\hat{\alpha}_S$ *constructed with* $\delta \geq 1/2$*. Let* $c = c_0\sqrt{T}$ *with* $c_0 < 0$*, then: 1)* $T^{1/2}(\hat{\alpha}_S - 1) = c_0 + o_p(1)$*, where* *the* $o_p(1)$ *term does not depend on* c_0*; 2)* $t_\beta^{FS} \to^d N(0,1)$ *as* $c_0 \to -\infty$ *and* $T \to \infty$*.*

This result is important for the following reasons. In order to bridge the gap between the $I(0)$ and $I(1)$ cases and ensure that for values of the autoregressive parameter local to one the test t_β^{FS} has the least possible size distortions, we need $\delta \geq 1/2$. This permits the limit $N(0,1)$ distribution to apply as $c \to -\infty$. Otherwise, from Theorem 2, a conservative test is to be expected. This in fact restricts the neighborhood where truncation applies. On the other hand, increasing δ beyond 1/2 would imply that in moderate samples the truncation applies less and less and that $\hat{\alpha}_S$ would basically be equivalent to the OLS estimate $\hat{\alpha}$. These considerations suggest that $\delta = 1/2$ should be the preferred choice. Indeed, simulations reported below will show that this value leads to a procedure which works best in small samples.

2.3 Useful Modifications for Improved Finite Sample Properties

In finite samples, it is well known that the OLS estimate is biased downward, especially when the data are linearly detrended. A solution is to replace the OLS estimate by one with less bias. We consider two such modifications: 1) the median unbiased estimate as described by Andrews (1993), 2) a modified version of the weighted symmetric least-squares estimate as described by Roy and Fuller (2001).

When using the median unbiased estimate, we replace the OLS estimate $\hat{\alpha}$ by a value that would produce $\hat{\alpha}$ as the median of the distribution assuming Normal errors. The specific correction depends on the nature of the deterministic components and Andrews (1993) provide tabulated values for the linear trend case. More precisely the estimate is defined as follows. Let $m(\alpha)$ be the median function of $\hat{\alpha}$, then $\hat{\alpha}_M = 1$ if $\hat{\alpha} > m(1)$, $\hat{\alpha}_M = m^{-1}(\hat{\alpha})$ if $m(-1) < \hat{\alpha} < m(1)$, and $\hat{\alpha}_M = -1$ if $\hat{\alpha} < m(-1)$, where $m(-1) = \lim_{\alpha \to -1} m(\alpha)$, and $m^{-1} : (m(-1), m(1)] \to (-1, 1]$ is the inverse function of $m(\cdot)$ that satisfies $m^{-1}(m(\alpha)) = \alpha$ for $\alpha \in (-1, 1]$. Note that the median unbiased estimate also applies a truncation device given that the parameter space is restricted to $|\alpha| \leq 1$. This occurs since for values of $\hat{\alpha}$ above some threshold the assigned value is 1. Now this threshold depends on the sample size T and shrinks as T increases. It is easy to verify that this implies a truncation of the form specified by (7) with $\delta = 1$. Since our truncation imposes $0 < \delta < 1$, we can substitute the OLS estimate by the median-unbiased estimate without changing any of the stated large sample results.

The weighted symmetric least-squares (WSLS) estimate of α is defined by $\hat{\alpha}_W = [\sum_{t=2}^{T-1} \hat{u}_t^2 + \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2]^{-1} \sum_{t=2}^{T} \hat{u}_t \hat{u}_{t-1}$. An estimate of its variance is given by $\hat{\sigma}_W^2 = [\sum_{t=2}^{T-1} \hat{u}_t^2 + \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2]^{-1}(T-3)^{-1} \sum_{t=2}^{T}(\hat{u}_t - \hat{\alpha}_W \hat{u}_{t-1})^2$ and the associated t-ratio for testing $\alpha = 1$ is $\hat{\tau}_W = (\hat{\alpha}_W - 1)/\hat{\sigma}_W$. The modification proposed by Roy and Fuller (2001) is similar in spirit to our superefficient estimate in that it also replaces estimates in a $T^{1/2}$ neighborhood of one. It is, however, discontinuous and depends on the value of the t-ratio $\hat{\tau}_W$, making the procedure explicitly dependent on some pre-test. The modified value is given by $\hat{\alpha}_{TW} = \hat{\alpha}_W + C(\hat{\tau}_W)\hat{\sigma}_W$, where

$$C(\hat{\tau}_W) = \begin{cases} -\hat{\tau}_W & \text{if} \quad \hat{\tau}_W > \tau_{pct} \\ T^{-1}\hat{\tau}_W - 3[\hat{\tau}_W + k(\hat{\tau}_W + 5)]^{-1} & \text{if} \quad -5 < \hat{\tau}_W \leq \tau_{pct} \\ T^{-1}\hat{\tau}_W - 3[\hat{\tau}_W]^{-1} & \text{if} \quad -(3T)^{1/2} < \hat{\tau}_W \leq -5 \\ 0 & \text{if} \quad \hat{\tau}_W \leq -(3T)^{1/2} \end{cases}$$

with $k = [3T - \tau_{pct}^2(I_p + T)][\tau_{pct}(5 + \tau_{pct})(I_p + T)]^{-1}$, where with an $AR(p)$ structure for u_t, I_p is the integer part of $(p+1)/2$, and τ_{pct} is the percentile of the limiting distribution of $\hat{\tau}_W$ when $\alpha = 1$. The form of the WSLS estimator for an $AR(p)$ process is given in Fuller (1996, p. 572). If $\tau_{pct} = -1.96$, $\hat{\alpha}_{TW}$ is nearly unbiased when $\alpha < 1$ and has a median of 1 when $\alpha = 1$.

The procedure recommended is the following: 1) Detrend the data by OLS to obtain residuals $\hat{u}_t$; 2) Estimate an $AR(1)$ for $\hat{u}_t$ yielding the estimate $\hat{\alpha}$ (or construct the WSLS estimate $\hat{\alpha}_W$) ; 3) Get the corresponding median-unbiased estimate $\hat{\alpha}_M$

(or the modified version of $\hat{\alpha}_W$, $\hat{\alpha}_{TW}$); 4) Construct $\hat{\alpha}_{MS}$ by replacing $\hat{\alpha}_M$ by 1 if $|\hat{\alpha}_M - 1| \leq dT^{-\delta}$ (or $\hat{\alpha}_{TW}$ by 1 if $|\hat{\alpha}_{TW} - 1| \leq dT^{-\delta}$); 5) Apply a GLS procedure with $\hat{\alpha}_{MS}$ to obtain an estimate β and construct the t-statistic denoted by t_{β}^{FMS}.

2.4 Simulation evidence

We conducted simple Monte Carlo experiments to illustrate the size and power of our test. The data are generated by $y_t = \beta t + u_t$ where $u_t = \alpha u_{t-1} + e_t$ with $e_t \sim i.i.d.$ $N(0,1)$ and $u_0 = 0$. In all cases, $50,000$ replications are used, and the nominal size is 5%. Under the null hypothesis H_0: $\beta = 0$ while under the alternative hypothesis H_1: $\beta > 0$. We first discuss the properties of the procedure when the median unbiased estimate $\hat{\alpha}_{MS}$ is used and return later with a comparision with the method that uses the truncated weighted symmetric least-squares (TWSLS) estimate. Unreported simulations showed that the use of a bias correction is essential to bring the size of the test to an acceptable level. We henceforth only consider the t-statistic t_{β}^{FMS} when the OLS estimate of α is replaced by the median unbiased estimate. The null rejection probabilities were simulated for values of α in the range $[0,1]$ with increments of 0.05 in the range $[0.0, 0.90]$ and in increments of 0.01 in the range $[0.90, 1.0]$. The sample sizes used are $T = 100$, 250, and 500. We consider four cases for the value of δ, namely $\delta = 0.3$, 0.4, 0.5, and 0.6, and d is set to 1.

The results are presented in Figure 1. As expected, when $\delta < 1/2$, the test is conservative for α near 1. When $\delta > 1/2$, the test shows substantial liberal size distortions. What is important is that with $\delta = 1/2$ the test shows basically no size distortion even with $T = 100$. The simulations also show that the choice of $d = 1$ is appropriate if the goal is to reduce size distortions to a minimum. For example, when $T = 100$, setting $d = 1/2$ would yield results similar to the case $\delta = 0.6$ and setting $d = 2$ would be almost the same as using $\delta = 0.3$. Hence, we shall continue to use $d = 1$. Further simulations showed the procedure to have similar properties when using two-sided 10% tests. With two-sided 5% tests, size distortions are somewhat higher when $T = 100$ and $\alpha = 1$, but decrease rapidly as T increases.

The only available alternative procedure that offers a test with the same critical values under both the $I(0)$ and $I(1)$ cases is that of Vogelsang (1998). He considers model (1) in partial sums form so that the scaled Wald test has a non-degenerate limit distribution under both the $I(0)$ and $I(1)$ cases, though different. The novelty is that he weights the statistic by a unit root test scaled by some parameter, which for any given significance level can be chosen so that the asymptotic critical values are the same. The resulting test statistic is denoted $t\text{-}PS_T$. His simulations show, however, the test to have little power in the $I(1)$ case so that he resorts to advocating the joint use of that test and a normalized Wald test from regression (1) estimated by OLS, labelled $T^{-1/2}t\text{-}W_T$, that has good properties in the $I(1)$ case but has otherwise very little power in the $I(0)$ case. This leads to problems related to the use of multiple tests where the size of each needs to be modified.

Consider now the power of the tests. Given the theoretical and simulation results, we henceforth use $\delta = 1/2$ and tests constructed with a median unbiased adjustment, t_β^{FMS}. The power of our test is compared to three other tests: the t-test based on the infeasible GLS estimate which uses the true value of α, as well as the $T^{-1/2}t\text{-}W_T$ and the $t\text{-}PS_T$ tests of Vogelsang (1998) for which we use a 5% nominal size and, hence, the proper comparisons should be made assuming they are applied independently. The power curves are plotted for $\alpha = 1.0$, $0.95, 0.90$ and 0.80 for a range of values of $\beta > 0$. The number of replications is again $50,000$. Results for $T = 100, 250$ and 500 were obtained but only those for $T = 100$ are explicitely reported in Figure 2. Consider first the case $\alpha = 1$. The test t_β^{FMS} has basically the same power as the infeasible GLS test (slightly higher when $T = 100$ because of a small liberal size distortion). The test is more powerful than Vogelsang's (1998) $T^{-1/2}t\text{-}W_T$ test which is the preferred one for this value of α. Consistent with the results in Vogelsang (1998), the $t\text{-}PS_T$ test has substantially lower power. When $\alpha = 0.95$ or 0.90, the power of the test t_β^{FMS} is not close to that of the infeasible GLS test but it is higher than either of Vogelsang's tests. As T increases, the power of our test gets closer to that of the infeasible GLS test. For instance, when $T = 500$ and $\alpha = 0.9$, the power of our test is as good as the infeasible test. When $\alpha = 0.8$, for any T the t_β^{FMS} has again power close to the infeasible test and still higher than the test $t\text{-}PS_T$ (the test $T^{-1/2}t\text{-}W_T$ is so conservative that power is non zero only for high values of β).

Roy et al. (2004) consider the same problem with an $AR(p)$ noise component that is either stationary or has one unit root. Their procedure is based on a one step Gauss Newton regression using the TWSLS estimate. Consider a first-order expansion of the regression (3) around initial estimates $(\widetilde{\mu}_0, \widetilde{\beta}_0, \widetilde{\alpha}_0)$

$$\widetilde{e}_t = \mu_\Delta(1 - \widetilde{\alpha}_0) + \beta_\Delta[t - \widetilde{\alpha}_0(t-1)] + \alpha_\Delta \widetilde{y}_{t-1} + \omega_t \tag{8}$$

where $\widetilde{y}_t = y_t - \widetilde{\mu}_0 - \widetilde{\beta}_0 t$, $\widetilde{e}_t = \widetilde{y}_t - \widetilde{\alpha}_0 \widetilde{y}_{t-1}$, and $\{\omega_t\}$ are the errors. Estimating (8) by OLS yields the estimates $(\widetilde{\mu}_\Delta, \widetilde{\beta}_\Delta, \widetilde{\alpha}_\Delta)$ and the one-step Gauss-Newton estimates are $(\hat{\mu}_{GN}, \hat{\beta}_{GN}, \hat{\alpha}_{GN}) = (\widetilde{\mu}_0, \widetilde{\beta}_0, \widetilde{\alpha}_0) + (\widetilde{\mu}_\Delta, \widetilde{\beta}_\Delta, \widetilde{\alpha}_\Delta)$. The test statistic is $t_{GN} = (\hat{\beta}_{GN} - \beta_0)/se(\hat{\beta}_{GN})$, where $se(\hat{\beta}_{GN})$ is essentially the same as the standard error from the regression (8), see Roy et al. (2004) for details. They suggest as the initial estimate of α the TWSLS estimate with $\tau_{pct} = -2.85$. The initial values of (μ, β) are obtained from the Feasible GLS regression using the TWSLS estimate with $\tau_{pct} = -1.96$. The limit distribution of t_{GN} is standard Normal when $|\alpha| < 1$ but not so when $\alpha = 1$. Nevertheless, the simulations of Roy et al. (2004) show that even in small samples, there are only small departures from the nominal size in the $I(1)$ case when two-sided 5% tests are used. We now address the following two issues: a) evaluate whether there is any benefit in using the TWSLS estimate instead of the median unbiased estimate; b) how our procedure performs relative to theirs in finite samples.

We conducted simulation experiments for the size and power of the tests. Table 1 presents the size of one-sided 5% tests for the Gauss Newton statistics t_{GN}, the t_{β}^{FMS} test using a median unbiased estimate, labelled $t_{\beta}^{FMS}(MU)$, and the t_{β}^{FMS} test using the TWSLS estimate with $\tau_{pct} = -2.85$, labelled, $t_{\beta}^{FMS}(TW)$. The exact sizes of the tests t_{GN} and $t_{\beta}^{FMS}(MU)$ are comparable, except when $T = 100$ and α is close to or equal to one. The test $t_{\beta}^{FMS}(MU)$ has slightly higher liberal size distortions when $\alpha = 1$. On the other hand, the test t_{GN} is more conservative for values of α between 0.8 and 1.0. When comparing the tests $t_{\beta}^{FMS}(MU)$ and $t_{\beta}^{FMS}(TW)$, one sees a similar trade-off. We again generated power results for $T = 100$, 250 and 500 but we explicitely report only those for $T = 100$ in Figure 3. It is seen that the test $t_{\beta}^{FMS}(MU)$ dominates $t_{\beta}^{FMS}(TW)$. Hence, in the sequel, we use the median unbiased estimate. Comparing the tests $t_{\beta}^{FMS}(MU)$ and t_{GN}, the results show that power is higher (and close to that of the infeasible GLS procedure) with our test for all values of α when $T = 100$. When $T = 250$ or 500 the power of our test is, in general, higher unless $\alpha = .95$. In particular, our test is noticeably better when $\alpha = 1$.

3 Generalization of the model

We now consider an extension of the analysis to the case where the error term u_t is allowed to have a more general structure than the simple $AR(1)$ process with *i.i.d.* shocks assumed so far. The data generating process is now assumed to be

$$y_t = \mu + \beta t + u_t; \quad u_t = \alpha u_{t-1} + v_t \tag{9}$$

with $v_t = d(L)e_t$, $d(L) = \sum_{i=0}^{\infty} d_i L^i$, $\sum_{i=0}^{\infty} i|d_i| < \infty$, $d(1) \neq 0$, and $\{e_t\}$ a martingale difference sequence with respect to a filtration $\mathcal{F}_t$ to which it is adapted. Also,

$E[e_t^2|\mathcal{F}_t] = \sigma^2$ and $\sup_t E[e_t^4] < \infty$. Again, we assume for simplicity that u_0 is some constant. These conditions imply that the following functional central limit theorem hold for the partial sums of v_t, $T^{-1/2}\sum_{t=1}^{[rT]} v_t \Rightarrow \sigma d(1)W(r)$. Under the stated conditions, u_t has an autoregressive representation, say $A(L)u_t = e_t$ where $A(L) = 1-\sum_{i=1}^{\infty} a_i L^i$. In the representation (9), we want α to represent the sum of the autoregressive coefficients. Hence, we use the representation $u_t = \alpha u_{t-1} + A^*(L)\Delta u_{t-1} + e_t$, with $A^*(L) = \sum_{i=1}^{\infty} a_i^* L^i$ where $a_i^* = -\sum_{j=i+1}^{\infty} a_j$. We cannot use the estimate $\hat{\alpha}$ based on an autoregression of order one since it is inconsistent for α when the errors u_t are a general $I(0)$ process. Instead, we base our estimate on a truncated autoregression of order k. Let $\hat{u}_t$ be the residuals from a regression of y_t on $\{1, t\}$, then the estimate of α considered is the OLS estimate $\widetilde{\alpha}$ obtained from the regression

$$\hat{u}_t = \alpha\hat{u}_{t-1} + \sum_{i=1}^{k} \zeta_i \Delta\hat{u}_{t-i} + e_{tk} \tag{10}$$

The estimate $\widetilde{\alpha}$ has the following properties provided $k \to \infty$ and $k^3/T \to 0$ as $T \to \infty$, see Berk (1974) and Said and Dickey (1984). When u_t is $I(0)$, $T^{1/2}(\widetilde{\alpha}-\alpha) = O_p(1)$. On the other hand, if $\alpha = 1 + c/T$, $T(\widetilde{\alpha} - 1) \Rightarrow c + d(1)\int_0^1 J_c^*(r)dW(r)/\int_0^1 J_c^*(r)^2 dr$ where $J_c^*(r)$ is the residual function from the regression of $J_c(r) \equiv \int_0^r \exp(c(r-s))dW(s)$ on $\{1, r\}$. The truncated estimate $\widetilde{\alpha}_S$ defined by (7) with $\widetilde{\alpha}$ instead of $\hat{\alpha}$ is therefore still superefficient under a local unit root, i.e. $T(\widetilde{\alpha}_S - 1) \to_p 0$.

As in the $AR(1)$ case, to improve the finite sample performance, we consider an approximately median unbiased estimate of α as described in Andrews and Chen (1994). Since the distribution of $\widetilde{\alpha}$ depends on (α, ζ_1,...,ζ_k), the method is based on an iterative procedure that jointly estimates α and the nuisance parameters (ζ_1,...,ζ_k) based on the regression (10) and then treats the estimate ($\widetilde{\zeta}_1$,..., $\widetilde{\zeta}_k$) as though they were the true values of (ζ_1,..., ζ_k) and computes the median unbiased estimate $\widetilde{\alpha}_M$. One then treats $\widetilde{\alpha}_M$ as given and re-computes the OLS estimates of (ζ_1,..., ζ_k). The procedure is repeated until convergence and $\widetilde{\alpha}_{MS}$ is constructed by replacing $\widetilde{\alpha}_M$ by 1 if $|\widetilde{\alpha}_M - 1| \leq dT^{-\delta}$. To estimate β we use a quasi-FGLS procedure assuming $AR(1)$ errors, i.e. the OLS estimate in the transformed regression:

$$y_t - \widetilde{\alpha}_{MS} y_{t-1} = (1 - \widetilde{\alpha}_{MS})\mu + \beta[t - \widetilde{\alpha}_{MS}(t-1)] + (u_t - \widetilde{\alpha}_{MS} u_{t-1}) \tag{11}$$

for $t = 2, ..., T$ together with $y_1 = \mu + \beta + u_1$. Denote the resulting estimate of β by $\widetilde{\beta}$. Since $v_t \equiv u_t - \alpha u_{t-1}$exhibits serial correlation in general, we need to correct for this. Hence, the statistic considered is (where the subscript RQF stands for Robust Quasi Feasible (GLS)): $t_\beta^{RQF} = (\widetilde{\beta} - \beta_0)/\sqrt{\hat{h}_v (X'X)_{22}^{-1}}$ where $(X'X)_{22}^{-1}$ is the (2,2) element of the matrix $(X'X)^{-1}$ with $X = [x_1, ..., x_T]'$, $x_t' = [(1 - \widetilde{\alpha}_{MS}), t - \widetilde{\alpha}_{MS}(t-1)]$ for $t = 2, ..., T$ and $x_1' = (1, 1)$. Also, $\hat{h}_v$ is a consistent estimate of (2π) times the spectral density function of v_t at frequency 0. Hence, the usual estimate of the variance of the residuals is replaced by $\hat{h}_v$. We consider two classes of estimates $\hat{h}_v$. One is based on a weighted sum of the autocovariance function given by:

$$\hat{h}_v = T^{-1} \sum_{t=1}^{T} \hat{v}_t^2 + T^{-1} \sum_{j=1}^{T-1} \lambda(j, m) \sum_{t=j+1}^{T} \hat{v}_t \hat{v}_{t-j} \tag{12}$$

for some weight function $\lambda(j, m)$ and bandwidth m, with $\hat{v}_t$ the OLS residuals from the regression (11). The quadratic spectral window is used with the bandwidth m selected according to the "plug-in method" of Andrews (1991) using an AR(1) approximation. We also consider autoregressive spectral density estimates. Note that the residuals from the regression (11) are $(1 - \widetilde{\alpha}_{MS} L)u_t$ and, hence, in large samples, they are equivalent to $v_t = (1 - \alpha L)u_t$. Consider first the case where $|\alpha| < 1$. $A(L)$ is then invertible so that $u_t = A(L)^{-1} e_t$ and $v_t = (1 - \alpha L)A(L)^{-1} e_t$. Since $A(1) = 1 - \alpha$, the spectral density at frequency zero of v_t is simply σ^2. To obtain a consistent estimate, we use the following approximate regression

$$y_t - \widetilde{\alpha}_{MS} y_{t-1} = \mu^* + \beta^* t + \sum_{i=1}^{k} \rho_i \Delta y_{t-i} + e_{tk}$$

with $\hat{e}_{tk}$ the corresponding OLS residuals. The estimate of σ^2 is then $\hat{\sigma}^2 = \hat{h}_v = (T-k)^{-1}\sum_{t=k+1}^{T}\hat{e}_{tk}^2$, which is consistent provided $k \to \infty$ and $k^3/T \to 0$ as $T \to \infty$. Consider now the case $\alpha = 1$. Then, in large samples, $(1-\widetilde{\alpha}_{MS}L)u_t$ is equivalent to $v_t = \Delta u_t$ and an autoregressive spectral density estimate at frequency zero can be obtained from the regression $\hat{v}_t = \sum_{i=1}^{k}\phi_i\hat{v}_{t-i} + \eta_{tk}$, where $\hat{v}_t$ are the OLS residuals from regression (11). Denote the estimate by $(\hat{\phi}_1, ..., \hat{\phi}_k)$ and $\hat{\sigma}_{\eta k}^2 = (T-k)^{-1}\sum_{t=k+1}^{T}\hat{\eta}_{tk}^2$, then $\hat{h}_v = \hat{\sigma}_{\eta k}^2/(1-\sum_{i=1}^{k}\hat{\phi}_i)^2$. The decision rule to select the estimate for the case $\alpha = 1$ or $|\alpha| < 1$ is based on whether the truncated value $\widetilde{\alpha}_{MS}$ is 1 or not. Asymptotically, this results in a correct classification and a consistent estimate for both cases. Hence, the procedure recommended is: 1) Detrend the data by OLS to obtain residuals $\hat{u}_t$; 2) Estimate (10) by OLS with k selected using an information criterion (we recommend using the BIC with k allowed to be in the range $[0, 12(T/100)^{1/4}]$). If the order selected is $k = 0$, the procedure of Section 2 applies, otherwise, the next steps are applied; 3) Use the method of Andrews and Chen (1994) to obtain the approximately median unbiased estimate $\widetilde{\alpha}_M$ and truncate its value to 1 if $|\widetilde{\alpha}_M - 1| \leq T^{-1/2}$ to get $\widetilde{\alpha}_{MS}$; 4) Apply the quasi GLS procedure with $\widetilde{\alpha}_{MS}$ to obtain the estimate of β and construct the t-statistic t_{β}^{RQF} using one of the estimate $\hat{h}_v$ suggested to construct the spectral density function at frequency zero of v_t.

3.1 Finite sample simulations

This section presents results about the finite sample size and power of the test with an AR(2) error component. The DGP is assumed to be

$$y_t = \beta t + u_t; \quad u_t = \alpha u_{t-1} + \psi(u_{t-1} - u_{t-2}) + e_t$$

where $e_t \sim i.i.d.\ N(0,1)$ and $u_0 = u_{-1} = 0$. Our hypotheses are H_0: $\beta = 0$ and H_1: $\beta > 0$. Given the high computation costs, 1000 replications are used. We first consider the size properties of our test at the nominal 5% level. These are obtained for $\alpha = 1, 0.95, 0.90, 0.80$, $\psi = 0.0, 0.3, 0.5, 0.7$, and $T = 100, 250$. We consider positive AR coefficients since this is the most relevant case in practice. We use four variations of our test pertaining to the choice of the estimate $\hat{h}_v$. In the first case, the estimate (12) is used and is referred to as "NP" (for Non Parametric) and $\widetilde{\alpha}$ is obtained from (10) with $k = 1$ (i.e., we impose the true value). The other three are autoregresive based estimates with k chosen as follows: 1) fixed at $k = 1$ (to assess the effect of estimating k); 2) with the Akaike (1969) information criterion, AIC; 3) with the Bayesian Information Criterion, BIC, see Schwarz (1978).

The results, presented in Table 2, show some features of interest. When the test is based on the weighted sum of autocovariances to construct $\hat{h}_v$, it is very conservative for $\psi > 0$ and $\alpha < 1$, and the distortions become more serious as ψ increases. When $\alpha = 1$, the test is liberal with the distortions reducing only slightly as T increases from 100 to 250. Using an autoregressive spectral density estimate, the size distortions are very small even with $T = 100$ unless $\alpha = 1$ and ψ is small (in which case the test is slightly liberal) or $\alpha = 0.95$ and ψ is large (in which case the test is conservative). However, these size distortions diminish when T is increased to 250. Also, there is no noticeable difference between the cases with the lag length assumed known or estimated. Hence, our recommendation is to use the autoregressive spectral estimate with an information criterion to select the lag length.

We now consider the power of our test relative to the t-test based on the infeasible GLS using the true value of α, and the $T^{-1/2}t\text{-}W_T$ and $t\text{-}PS_T$ tests of Vogelsang (1998), with a 5% size so that their properties pertains to the case when both tests are applied independently. The power curves for $T = 100$ are plotted for $\alpha = 1.0, 0.95, 0.90$, 0.80 and a range of values of $\beta > 0$. Results were obtained for $\psi = 0.3$, 0.5 and 0.7, and those for $\psi = 0.5$ are presented in Figure 4. Consider first the case $\alpha = 1$. Our test is as powerful as the infeasible GLS test for any ψ. When $\alpha = 0.95$ or 0.90, the power of our test is not close to that of the infeasible GLS but is still competitive with the most powerful of the $T^{-1/2}t\text{-}W_T$ and $t\text{-}PS_T$ tests. When $\alpha = 0.8$, the t_{β}^{RQF} has power close to the infeasible GLS test. Note that all tests lose power as ψ increases.

4 Specifying the trend function for unit root tests

As discussed in the introduction, the correct specification of a trend function is important when testing for unit roots. Ayat and Burridge (2000) used Vogelsang's tests to select the specification of the trend function as part of a sequential unit root testing procedure and concluded that there is no advantage in doing so due to their lack of power. We use simulations to see if there is any advantage in using our test as part of a sequential unit root testing procedure. The data are generated by $y_t = \beta t + u_t$, where $u_t = \alpha u_{t-1} + e_t$, $e_t \sim i.i.d.\ N(0,1)$ and $u_0 = 0$. We consider $\alpha = 1.0, 0.95, 0.90, 0.80$ and a range of values of $\beta \geq 0$. The sample size is $T = 100$ and $10,000$ replications are used. We consider unit root tests with only a constant, a constant and a time trend, as well as sequential unit root tests performed as follows: 1) test the null hypothesis of no trend, $\beta = 0$, using Vogelsang's tests or our proposed test with a 5% nominal size (note that Vogelsang's procedure involves the combine use of the statistics of $T^{-1/2}t\text{-}W_T$ and $t\text{-}PS_T$, hence 2.5% critical values are used for each statistic to have an overall test with nominal size no larger than 5%); 2) if the null hypothesis is not rejected, we conduct a unit root test with only a constant, otherwise we include a constant and a trend. The nominal size of the unit root test is also 5%.

Figure 5 reports the size and power of the four variants when the Dickey-Fuller (1979) unit root test is used. Consider first the size properties with $\alpha = 1$. When $\beta = 0$ and with only a constant, the test has size close to 5%. As β increases, the misspecification of the trend function becomes more serious so that the test has an increasing bias toward not rejecting the unit root. On the contrary, the unit root test with a constant and a trend always has exact size close to nominal size since there is no misspecification in the trend function for any values of β. The sequential tests have small liberal size distortions when β is near zero but otherwise maintain an exact size close to nominal level. Consider now the power of the tests with $\alpha < 1$. When $\beta = 0$, the unit root test with only a constant is the most powerful test. However, as β increases, the misspecification of the trend function becomes more serious and the power of the test decreases to zero. On the contrary, the unit root test including a trend has non-trivial power for any values of β since no misspecification occurs. However, when β is small the power is substantially lower than with a constant only. The sequential unit root test with Vogelsang's trend tests works well when β is very small or very large. However, when β is moderate size, the most relevant case in practice, the procedure evidently lacks power. In contrast, the sequential unit root test procedure with our trend test works very well for any values of β due to the higher power of the trend test applied in the first step. Of special interest is the fact that the power function of the unit root test with our trend test is higher than the best of the "constant only" or "constant and trend" specifications for any values of β. This shows that our trend test adapts well to the particular sample in selecting the appropriate specification of the trend. We obtained similar results with the GLS-detrended unit root tests of Elliott et al. (1996).

5 Empirical applications

We consider two empirical applications related to the data sets analyzed in Vogelsang (1998) and in Canjels and Watson (1997). In all cases, the estimate $\hat{h}_v$ is the autoregressive spectral estimate using the BIC to select the lag length. Vogelsang (1998) estimated postwar real GNP quarterly growth rates for the G7 countries. The exact sample period for each country is indicated in Table 3. He considered the possibility of a structural change in the slope of the trend function in 1973 (see, e.g., Perron, 1989). Hence, the interest is in comparing the estimates and confidence intervals for the pre-1973 and post-1973 periods. Table 3 reports the estimates of real GNP quarterly growth rates for both periods and the full sample with the 95% confidence intervals. There are several features of interest. First, the confidence intervals for our test statistic are much tighter than those of Vogelsang for 19 of 21 cases, as expected from our simulation results. Second, in all cases, the point estimates of β are higher in the pre-1973 sample compared to the post-1973 sample. For France, Germany, Italy, and Japan, the pre-1973 and post-1973 confidence intervals using our test statistic do not overlap, indicating a statistically significant decline in the rate of growth for these countries. This contrasts with Vogeslang's procedure for which non-overlapping 95% confidence intervals occur only for France and Japan, and only using the $T^{-1/2}t\text{-}W_T$ statistic. With three exceptions, (pre-1973 for France and the U.K, and post-1973 for Japan), the truncated median unbiased estimate of α is 1. The poor properties of the test $t\text{-}PS_T$ when α is close to one is reflected in its wide confidence intervals (in many cases, the limits exceed ± 99). When no truncation applies, the test $t\text{-}PS_T$ gives tighter confidence intervals than $T^{-1/2}t\text{-}W_T$ in 2 out of 3 cases. However, these are between 3 and 41 times as wide as those provided by our test.

Consider now the data set analyzed by Canjels and Watson (1997) which consists of annual real GDP per capita series over the post-war period for 128 countries (from the Penn World Table database). We here summarize the results, which can be found in the working paper version. The main feature of interest is that our method provides tighter confidence intervals for 117 countries out of the 128 considered, compared with the method of Canjels and Watson (1997). Improvements are very large when $\widetilde{\alpha}_{MS}$ is not close to one, although there are still some improvements when the truncation is effective and $\widetilde{\alpha}_{MS}$ is one. When no truncation is applied (i.e., $\widetilde{\alpha}_{MS} < 1$, occuring for 24 countries), the ratio of Canjels and Watson's confidence interval relative to ours varies from 1.11 (Burundi) to 12.09 (Luxembourg) with the average being 5.36. When the truncation is in effect (i.e., $\widetilde{\alpha}_{MS} = 1$, occuring for 104 countries), this ratio varies from 0.35 (USSR) to 1.51 (Argentina) with the average being 1.09.

6 Conclusions

We proposed a new procedure to carry inference about the slope of a trend function valid without prior knowledge about whether a series is $I(0)$ or $I(1)$. The test is based on a Feasible quasi GLS method with a superefficient estimate of the sum of the autoregressive coefficients α when $\alpha = 1$. Simulations and empirical applications have shown its usefulness and that it provides a clear improvement over existing methods. Its power is basically equivalent to that based on the infeasible GLS estimate when α is assumed known, unless α is near but not equal to one. It may appear that there is room for improvements in this case but as argued by Roy et al. (2004, p. 1088), this seems unlikely if we require a test that controls size in both the $I(1)$ and $I(0)$ cases.

7 Appendix: Technical Derivations

Proof of equation (5): With data generated by (2), using straightforward algebra we can express the t-statistic as (see also, Canjels and Watson, 1997):

$$t_{\beta}^{F} = \frac{q_{11}T^{-1/2}r_2 - T^{-1/2}q_{12}r_1}{[\hat{\sigma}^2 q_{11}(q_{11}T^{-1}q_{22} - T^{-1}q_{12}^2)]^{1/2}} \tag{A.1}$$

where $q_{11} = 1 + (T-1)(1-\hat{\alpha})^2 = 1 + o_p(1)$, $r_1 = u_1 + (1-\hat{\alpha})\sum_{t=2}^{T} u_t^* = u_1 + o_p(1)$

$$T^{-1/2}q_{12} = T^{-1/2}[1 + (T-1)\hat{\alpha}(1-\hat{\alpha}) + (1-\hat{\alpha})^2 \sum_{t=2}^{T} t] = o_p(1)$$

$$
\begin{aligned}
T^{-1}q_{22} &= T^{-1}[1+(T-1)\hat{\alpha}^2+2\hat{\alpha}(1-\hat{\alpha})\sum_{t=2}^{T}t+(1-\hat{\alpha})^2\sum_{t=2}^{T}t^2] \\
&= 1+T(1-\hat{\alpha})+T^2(1-\hat{\alpha})^2/3+o_p(1) \\
T^{-1/2}r_2 &= T^{-1/2}[u_1+\hat{\alpha}\sum_{t=2}^{T}u_t^*+(1-\hat{\alpha})\sum_{t=2}^{T}tu_t^*] \\
&= T^{-1/2}\sum_{t=2}^{T}u_t^*+T(1-\hat{\alpha})T^{-3/2}\sum_{t=2}^{T}tu_t^*+o_p(1)
\end{aligned}
$$

with $u_t^* = u_t - \hat{\alpha}u_{t-1}$, $\hat{\sigma}^2 = T^{-1}\sum_{t=1}^{T}\hat{e}_t^2$ and $\hat{e}_t$ the residuals from a regression of $y_t^* = y_t - \hat{\alpha}y_{t-1}$ on $x_t^* = [1-\hat{\alpha},\, t-\hat{\alpha}(t-1)]'$ for $t = 2, ..., T$ and $x_1^* = [1,1]'$, $y_1^* = y_1$. It is easy to show that $\hat{\sigma}^2 = \sigma^2 + o_p(1)$. Equation (5) follows substituting in (A.1).

Proof of Theorem 1: Let $A = \{T^{\delta}|\hat{\alpha}-1| > d\}$ and let $\bar{A} = \{T^{\delta}|\hat{\alpha}-1| < d\}$. For part (a), it suffices to show that $T^{1/2}(\hat{\alpha}_S-\alpha) - T^{1/2}(\hat{\alpha}-\alpha) \to_p 0$. We have

$$
\begin{aligned}
\lim_{T\to\infty}\Pr(|T^{1/2}(\hat{\alpha}_S-\hat{\alpha})| &> \varepsilon) = \\
\lim_{T\to\infty}\Pr(|T^{1/2}(\hat{\alpha}_S-\hat{\alpha})| &> \varepsilon|A)\Pr(A) + \lim_{T\to\infty}\Pr(|T^{1/2}(\hat{\alpha}_S-\hat{\alpha})| > \varepsilon|\bar{A})\Pr(\bar{A})
\end{aligned}
$$

The first term is zero given that, if A is true, we have $\hat{\alpha}_S = \hat{\alpha}$ so that $\Pr(|T^{1/2}(\hat{\alpha}_S - \hat{\alpha})| > \varepsilon|A) = 0$. The second term is zero since $\hat{\alpha} \to_p \alpha \neq 1$ implies $\lim_{T\to\infty}\Pr(\bar{A}) \to 0$ as $T \to \infty$. Therefore, $\Pr(|T^{1/2}(\hat{\alpha}_S-\hat{\alpha})| > \varepsilon) \to 0$ as $T \to \infty$. For part (b), we have

$$
\begin{aligned}
\lim_{T\to\infty}\Pr(|T(\hat{\alpha}_S-1)| &> \varepsilon) = \\
\lim_{T\to\infty}\Pr(|T(\hat{\alpha}_S-1)| &> \varepsilon|A)\Pr(A) + \lim_{T\to\infty}\Pr(|T(\hat{\alpha}_S-1)| > \varepsilon|\bar{A})\Pr(\bar{A})
\end{aligned}
$$

Now $T(\hat{\alpha}-1) = O_p(1)$ and $\delta \in (0,1)$ imply that $\lim_{T\to\infty}\Pr(A) = \Pr(T^{\delta-1}|T(\hat{\alpha}-1)| > d) \to 0$ as $T \to \infty$, so that the first term is zero. For the second term, if $\bar{A}$ is true, $\hat{\alpha}_S = 1$ so that $\Pr(|T(\hat{\alpha}_S-1)| > \varepsilon|\bar{A}) = 0$. Thus, $\Pr(|T(\hat{\alpha}_S-1)| > \varepsilon) \to 0$ as $T \to \infty$.

Proof of Theorem 2: To prove part a) first note that with $\alpha_T = 1 + c/T$, $T(\hat{\alpha} - 1) \Rightarrow c + \int_0^1 J_c^*(r)dW(r)/\int_0^1 J_c^*(r)^2dr$, with $J_c^*(r)$, $0 \le r \le 1$, the residual process from a continuous time regression of an Ornstein-Uhlenbeck $J_c(r)$ on $\{1, r\}$, where $J_c(r) = \int_0^r \exp(c(r-s))dW(s)$. Since the true value of α is in a T^{-1} neighborhood of 1, and $\hat{\alpha}_S$ truncates values of $\hat{\alpha}$ in a $T^{-\delta}$ neighborhood of 1 for some $0 < \delta < 1$ (i.e., a larger neighborhood), the results of Theorem 1 continue to apply in the local to unity case and $T(\hat{\alpha}_S - 1) \to_p 0$. Now the t-statistic can be written a

$$\begin{aligned} t_\beta^F &= \Bigg[T^{-1/2}\sum_{t=2}^{T}(u_t - u_{t-1}) - T(\hat{\alpha}_S - 1)T^{-3/2}\sum_{t=2}^{T} u_{t-1} \\ &\quad -T(\hat{\alpha}_S - 1)[T^{-3/2}\sum_{t=2}^{T} t(u_t - u_{t-1}) - T(\hat{\alpha}_S - 1)T^{-5/2}\sum_{t=2}^{T} tu_{t-1}]\Bigg] \\ &\quad / \left[\sigma^2(1 - T(\hat{\alpha}_S - 1) + T^2(\hat{\alpha}_S - 1)^2/3)\right]^{1/2} + o_p(1) \end{aligned} \tag{A.2}$$

We have $T^{-1/2}\sum_{t=2}^{T}(u_t - u_{t-1}) = T^{-1/2}\sum_{t=2}^{T} e_t + cT^{-3/2}\sum_{t=2}^{T} u_{t-1} \to \sigma[W(1) + c\int_0^1 J_c(r)dr] = \sigma J_c(1) \sim N(0, \sigma^2(\exp(2c) - 1)/2c)$. The result of part (a) follows using the fact that $T(\hat{\alpha}_S - 1) \to_p 0$, and that the following quantities are $O_p(1)$, $T^{-3/2}\sum_{t=2}^{T} u_{t-1}$, $T^{-3/2}\sum_{t=2}^{T} t(u_t - u_{t-1})$, and $T^{-5/2}\sum_{t=2}^{T} tu_{t-1}$.

Proof of Theorem 3: We start with the following Lemmas.

Lemma A.1 *Let $u_t = \alpha_T u_{t-1} + e_t$ with e_t as defined by (2), and $\alpha_T = \exp(b\frac{1}{T^{1-h}}) \approx 1 + b/T^{1-h}$ for some $b < 0$ and some $0 < h < 1$; then, a) $u_{[Tr]} = \sum_{j=0}^{[Tr]-1} \alpha_T^j e_{[Tr]-j} + \alpha_T^{[Tr]} u_0 = O_p(T^{\frac{1-h}{2}})$; b) $T^{-2+h}\sum_{t=1}^{T} u_{t-1}^2 \to_p -\sigma^2/(2b)$; c) $T^{-1}\sum_{t=1}^{T} u_{t-1}e_t \to_p 0$.*

Proof: For part (a), $u_{[Tr]}$ has a mean that is $o(1)$ and

$$\begin{aligned} var(T^{-\frac{1-h}{2}}u_{[Tr]}) &= \sigma^2 T^{-(1-h)}(1 + \alpha_T^2 + ... + \alpha_T^{2([Tr]-1)}) = \frac{\sigma^2(1 - \alpha_T^{2[Tr]})}{T^{1-h}(1 - \alpha_T^2)} \\ &= \frac{\sigma^2(1 - (1 + b/T^{1-h})^{2[Tr]})}{T^{1-h}(1 - (1 + b/T^{1-h})^2)} \to \frac{\sigma^2}{-2b} \end{aligned}$$

Part (b) follows using similar derivations. To prove part (c), start by squaring both sides of the equation $u_t = \alpha_T u_{t-1} + e_t$, substract u_{t-1} from both sides, take summations, and then after rearrangements

$$
\begin{aligned}
T^{-1}\sum_{t=1}^{T} u_{t-1}e_t &= \frac{1}{2\alpha_T}\left[T^{-1}u_T^2 - T^{-1}u_0^2 - T^{1-h}(\alpha_T^2-1)T^{-2+h}\sum_{t=1}^{T}u_{t-1}^2 - T^{-1}\sum_{t=1}^{T}e_t^2\right] \\
&\to {}_p\frac{1}{2}[0-0-(2b)(-1/2b)\sigma^2-\sigma^2] = 0
\end{aligned}
$$

using part (a), (b), $T^{-1}u_0^2 \to_p 0$ and $T^{-1}\sum_{t=1}^{T} e_t^2 \to_p \sigma^2$.

Lemma A.2 *Let u_t be generated as specified in Lemma A.1, and let $\hat{\alpha}$ be the OLS estimate in a regression of u_t on u_{t-1}, then $T(\hat{\alpha}-1) = b + O_p(1)$*

We have, using Lemma A.1 and the fact that $T^{1-h}(\alpha_T - 1) = b$,

$$
\hat{\alpha} = \frac{\sum_{t=2}^{T} u_t u_{t-1}}{\sum_{t=2}^{T} u_{t-1}^2} = \alpha_T + \frac{\sum_{t=2}^{T} u_{t-1}e_t}{\sum_{t=2}^{T} u_{t-1}^2}
$$

$$
T^{1-h}(\hat{\alpha}-1) = T^{1-h}(\alpha_T - 1) + \frac{T^{-1}\sum_{t=2}^{T} u_{t-1}e_t}{T^{-2+h}\sum_{t=2}^{T} u_{t-1}^2} = b + o_p(1)
$$

Note that the results of Lemmas A.1 and A.2 extend easily to the case with u_t replaced by the OLS residuals $\hat{u}_t$, though the derivations are more tedious.

To continue with the proof of Theorem 3, set $b = c_0$ and $h = 1/2$, so that $T^{1/2}(\hat{\alpha}-1) = c_0 + o_p(1)$. Now consider the truncated estimate $\hat{\alpha}_S$. For c_0 large and in large samples $\hat{\alpha}_S = \hat{\alpha}$ provide $\delta \geq 1/2$, and there is no truncation. Hence, for large c_0, $T^{1/2}(\hat{\alpha}_S - 1) = c_0 + o_p(1)$ and we can treat the value of c as known. We can then apply the results of Phillips and Lee (1996) who cover limit results of the estimate of β when the quasi differencing parameter $\alpha_T = 1 + c/T$ is known. From their results on p. 307, we deduce that $t_\beta^{FS} \to^d N(0,1)$ as $c_0 \to -\infty$.

References

Akaike, H. (1969), "Fitting Autoregressions for Predictions," *Annals of the Institute of Statistical Mathematics*, 21, 243-247.

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

Andrews, D.W.K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

Andrews, D.W.K. (1993), "Exactly Median-Unbiased Estimation of First Order Autoregressive/Unit Root Models," *Econometrica*, 61, 139-165.

Andrews, D.W.K., and Chen, H.-Y. (1994): "Approximately Median-Unbiased Estimation of Autoregressive Models," *Journal of Business and Economic Statistics*, 12, 187-204.

Berk, K.N. (1974), "Consistent Autoregressive Spectral Estimates," *Annals of Statistics*, *2*, 489-502.

Canjels, E., and Watson, M.W. (1997), "Estimating Deterministic Trends in the Presence of Serially Correlated Errors," *Review of Economics and Statistics*, 79, 184-200.

Cochrane, D., and Orcutt, G.H. (1949), "Applications of Least Squares Regressions to Relationships Containing Autocorrelated Error Terms," *Journal of the American Statistical Association*, 44, 32-61.

DeJong, D.N., Nankervis, J.C., Savin, N.E., and Whiteman, C.H. (1992), "Integration Versus Trend Stationary in Time Series," *Econometrica*, 60, 423-433.

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, 427-431.

Elliott, G., Rothenberg, T.J., and Stock, J.H. (1996), "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813-836.

Fuller, W.A. (1996), *Introduction to Statistical Time Series* (2nd ed.), Wiley (New York, NY).

Grenander, U., and Rosenblatt, M. (1957), *Statistical Analysis of Stationary Time Series*, New York: John Wiley.

Nelson, C.R., and Plosser, C. (1982), "Trends and Random Walks in Macroeconomic Time Series: Some Evidence and Implications," *Journal of Monetary Economics*, 10, 139-162.

Perron, P. (1988), "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach," *Journal of Economic Dynamics and Control*, 12, 297-332.

Perron, P. (1989), "The Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Econometrica*, 57, 1361-1401.

Phillips, P.C.B., and Durlauf, S. (1988), "Trends Versus Random Walks in Time Series Analysis," *Econometrica*, 56, 1333-1354.

Phillips, P.C.B., and Lee, C.C. (1996), "Efficiency Gains from Quasi-differencing under Nonstationarity," in P.M. Robinson and M. Rosenblatt (eds), *Athens Conference on Applied Probability and Time Series, Volume II: Time Series Analysis in Memory of E.J. Hannan*, Springer-Verlag (New York, NY), 300-313.

Prais, S.J., and Winsten, C.B. (1954), "Trend Estimators and Serial Correlation," Cowles Foundation Discussion Paper 383.

Roy, A., Falk, B., and Fuller, W.A. (2004), "Testing for Trend in the Presence of Autoregressive Errors," *Journal of the American Statistical Association* 99, 1082-1091.

Roy, A., and Fuller, W.A. (2001), "Estimation for Autoregressive Processes with a Root Near One," *Journal of Business and Economic Statistics* 19, 482-493.

Said, S.E., and Dickey, D.A. (1984), "Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order," *Biometrika*, 71, 599-608.

Schwarz, G. (1978), "Estimating the Dimension of a Model," *The Annals of Statistics*, 6, 461-464.

Sun, H., and Pantula, S.G. (1999), "Testing for Trends in Correlated Data," *Statistics and Probability Letters*, 41, 87-95.

Vogelsang, T.J. (1998), "Trend Function Hypothesis Testing in the Presence of Serial Correlation," *Econometrica*, 66, 123-148.

Table 1: Exact Sizes of One-sided 5% Nominal Size Tests

	t_{GN}			$t_{\beta}^{FMS}(MU)$			$t_{\beta}^{FMS}(TW)$		
α	T=100	T=250	T=500	T=100	T=250	T=500	T=100	T=250	T=500
1.00	0.056	0.053	0.052	0.100	0.067	0.056	0.069	0.057	0.054
0.98	0.025	0.020	0.024	0.049	0.024	0.016	0.019	0.012	0.011
0.97	0.023	0.023	0.032	0.049	0.028	0.027	0.019	0.016	0.020
0.95	0.025	0.031	0.041	0.051	0.040	0.048	0.022	0.028	0.041
0.90	0.032	0.043	0.048	0.058	0.058	0.056	0.034	0.051	0.055
0.80	0.043	0.048	0.048	0.063	0.055	0.052	0.052	0.054	0.052
0.70	0.050	0.049	0.049	0.062	0.053	0.051	0.057	0.052	0.050
0.60	0.052	0.052	0.049	0.059	0.054	0.052	0.056	0.053	0.051
0.40	0.050	0.052	0.049	0.054	0.053	0.050	0.051	0.052	0.050
0.00	0.052	0.049	0.051	0.047	0.045	0.049	0.052	0.049	0.051

Table 2: Finite Sample Null Rejection Probabilities with 5% Nominal Size

Model: $y_t = \beta t + u_t$, $u_t = \alpha u_{t-1} + \psi(u_{t-1} - u_{t-2}) + e_t$; H_0: $\beta = 0$

		T=100				T=250			
		NP	AR			NP	AR		
α	φ	k=1	k=1	AIC	BIC	k=1	k=1	AIC	BIC
1.00	0.0	0.138	0.103	0.110	0.120	0.083	0.070	0.056	0.079
	0.3	0.103	0.087	0.105	0.092	0.074	0.066	0.060	0.067
	0.5	0.077	0.062	0.082	0.059	0.065	0.060	0.062	0.050
	0.7	0.098	0.074	0.090	0.073	0.093	0.041	0.053	0.058
0.95	0.0	0.092	0.064	0.063	0.053	0.069	0.044	0.046	0.040
	0.3	0.025	0.036	0.044	0.036	0.016	0.040	0.038	0.038
	0.5	0.008	0.023	0.026	0.020	0.005	0.034	0.035	0.038
	0.7	0.004	0.011	0.017	0.013	0.002	0.017	0.024	0.012
0.90	0.0	0.098	0.069	0.055	0.056	0.083	0.066	0.066	0.051
	0.3	0.036	0.054	0.061	0.049	0.023	0.057	0.060	0.065
	0.5	0.021	0.056	0.057	0.058	0.007	0.079	0.070	0.067
	0.7	0.002	0.055	0.054	0.049	0.004	0.070	0.067	0.060
0.80	0.0	0.098	0.072	0.058	0.065	0.070	0.060	0.064	0.044
	0.3	0.039	0.074	0.072	0.058	0.019	0.060	0.065	0.075
	0.5	0.013	0.059	0.095	0.075	0.016	0.075	0.060	0.055
	0.7	0.011	0.074	0.080	0.074	0.007	0.077	0.059	0.064

Note: NP stands for the nonparametric method based on the weighted sum of autocovariances and AR refers to the autoregressive spectral density estimates.

Table 3: 95 % Confidence Interval for Quarterly Growth Rates of Post War Real GNP

		Vogelsang (1998)														
		$T^{-1/2}t-W_T$				$t-PS_T$				t_{β}^{RQF}						
Country	Period	$\beta_{\min}<$	$\hat{\beta}$ <	$\beta_{\max}$	width	$\beta_{\min}<$	β^* <	$\beta_{\max}$	width	$\tilde{\alpha}$	$\tilde{\alpha}_{MS}$	$\hat{k}$	$\beta_{\min}<$	$\hat{\beta}$	$<\beta_{\max}$	width
Canada	48:1-89:2	0.83	1.11	1.40	0.57	-15.36	1.15	17.66	33.02	0.97	1.00	1	0.89	1.10	1.31	0.42
	48:1-73:4	0.96	1.20	1.45	0.49	0.94	1.19	1.44	0.50	0.87	1.00	0	0.99	1.25	1.51	0.52
	74:1-89:2	0.41	0.80	1.19	0.78	-10.48	0.77	12.02	22.50	0.90	1.00	1	0.48	0.85	1.21	0.73
France	63:1-89:2	0.27	0.82	1.36	1.09	<-99	0.88	>99	>99	0.96	1.00	1	0.70	0.89	1.08	0.39
	63:1-73:4	0.98	1.31	1.64	0.66	1.20	1.31	1.43	0.23	0.06	0.12	0	1.27	1.31	1.35	0.08
	74:1-89:2	0.30	0.53	0.76	0.45	-35.40	0.53	36.47	71.86	0.92	1.00	0	0.41	0.56	0.70	0.29
Germany	50:1-89:2	0.26	0.98	1.70	1.45	<-99	1.07	>99	>99	0.97	1.00	0	0.85	1.09	1.34	0.49
	50:1-73:4	0.75	1.37	2.00	1.24	<-99	1.44	>99	>99	0.93	1.00	0	1.12	1.46	1.80	0.68
	74:1-89:2	0.18	0.50	0.82	0.65	-9.78	0.49	10.76	20.54	0.85	1.00	0	0.23	0.52	0.81	0.58
Italy	52:1-82:4	0.66	1.14	1.63	0.98	<-99	1.22	>99	>99	1.01	1.00	0	0.82	1.06	1.30	0.48
	52:1-73:4	1.10	1.34	1.58	0.48	0.55	1.36	2.17	1.62	0.83	1.00	0	1.09	1.35	1.61	0.51
	74:1-82:4	-0.06	0.62	1.30	1.36	<-99	0.68	>99	>99	0.72	1.00	2	-0.15	0.35	0.86	1.02
Japan	52:1-89:2	0.81	1.74	2.67	1.86	<-99	1.88	>99	>99	0.99	1.00	0	1.39	1.68	1.96	0.57
	52:1-73:4	1.84	2.30	2.75	0.91	-3.09	2.29	7.67	10.76	0.87	1.00	0	1.75	2.18	2.61	0.86
	74:1-89:2	0.87	1.04	1.22	0.35	-0.80	1.05	2.89	3.69	0.78	0.87	0	1.00	1.04	1.08	0.09
U.K.	60:1-89:2	0.29	0.56	0.83	0.54	-2.54	0.57	3.68	6.22	0.89	1.00	0	0.34	0.60	0.87	0.52
	60:1-73:4	0.44	0.73	1.02	0.57	0.52	0.72	0.92	0.40	0.54	0.61	0	0.67	0.72	0.77	0.10
	74:1-89:2	0.02	0.49	0.96	0.94	<-99	0.42	>99	>99	0.95	1.00	1	0.29	0.53	0.78	0.49
U.S.	47:1-89:2	0.53	0.78	1.03	0.49	-1.95	0.80	3.56	5.51	0.95	1.00	1	0.57	0.81	1.05	0.48
	47:1-73:4	0.58	0.88	1.18	0.61	-1.37	0.87	3.11	4.48	0.92	1.00	1	0.60	0.90	1.20	0.60
	74:1-89:2	0.25	0.67	1.10	0.85	-60.05	0.64	61.32	>99	0.89	1.00	1	0.29	0.67	1.05	0.76

Note: The columns with headings "$\beta_{\min}$ and $\beta_{\max}$" refer to the lower and upper bound of the confidence interval and the column with the heading "width" refers to its width.

Figure 1: Finite Sample Size of t_{β}^{FMS}

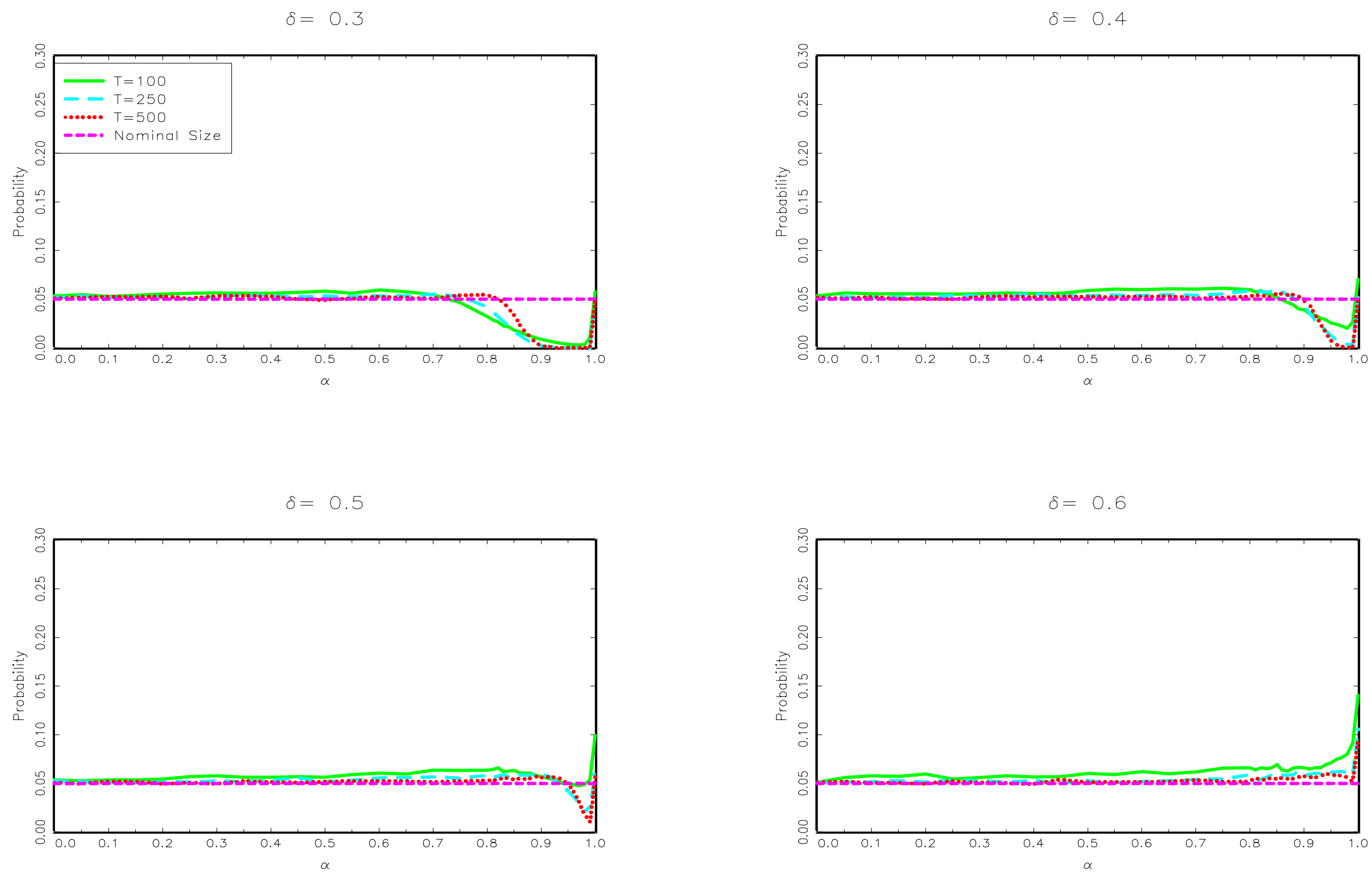

Figure 2: Finite Sample Power Comparisons With Vogelsang's (1998) Tests: T=100

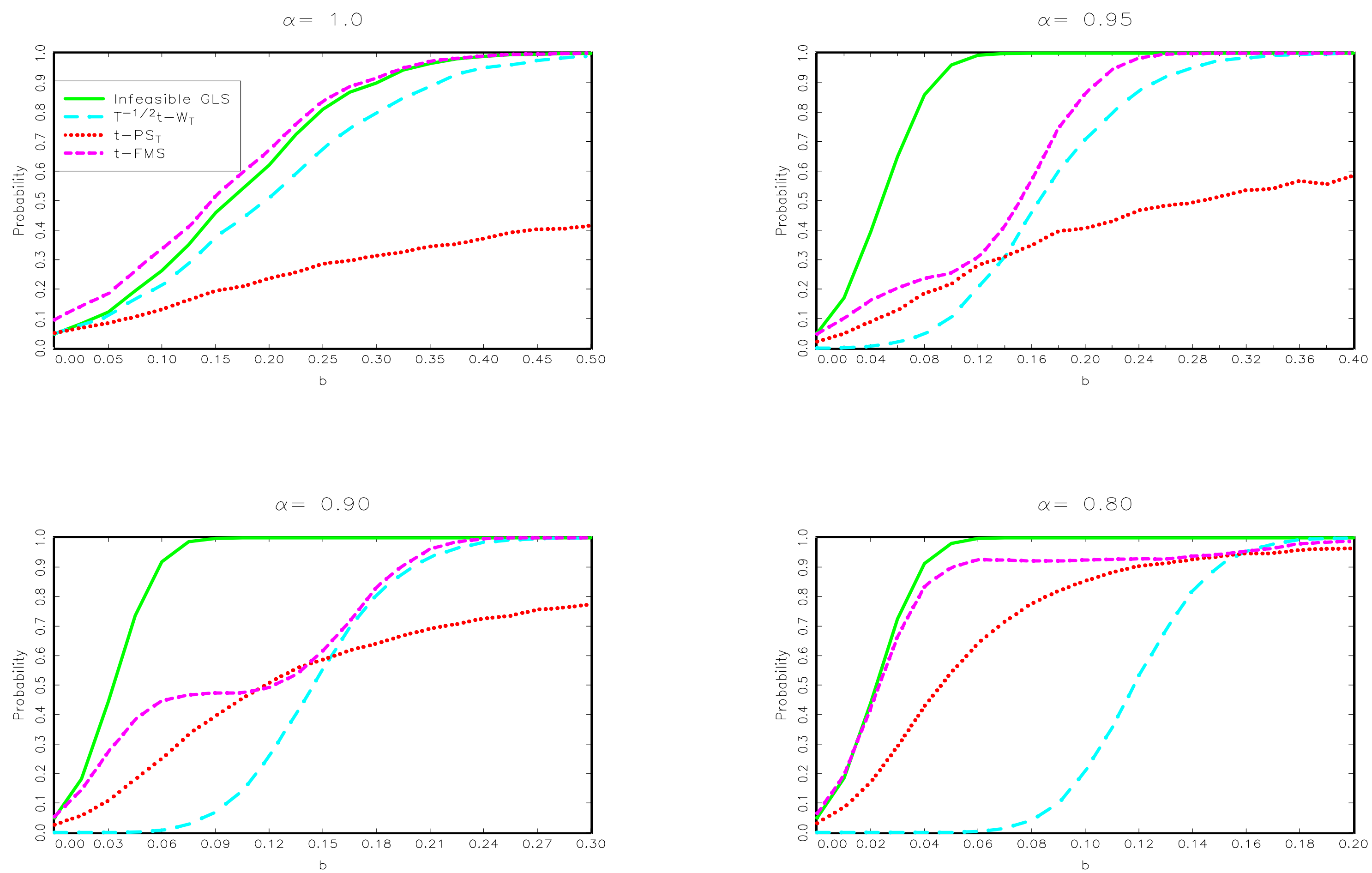

Figure 3: Finite Sample Power Comparisons
With Roy et al. (2004) Test: T=100

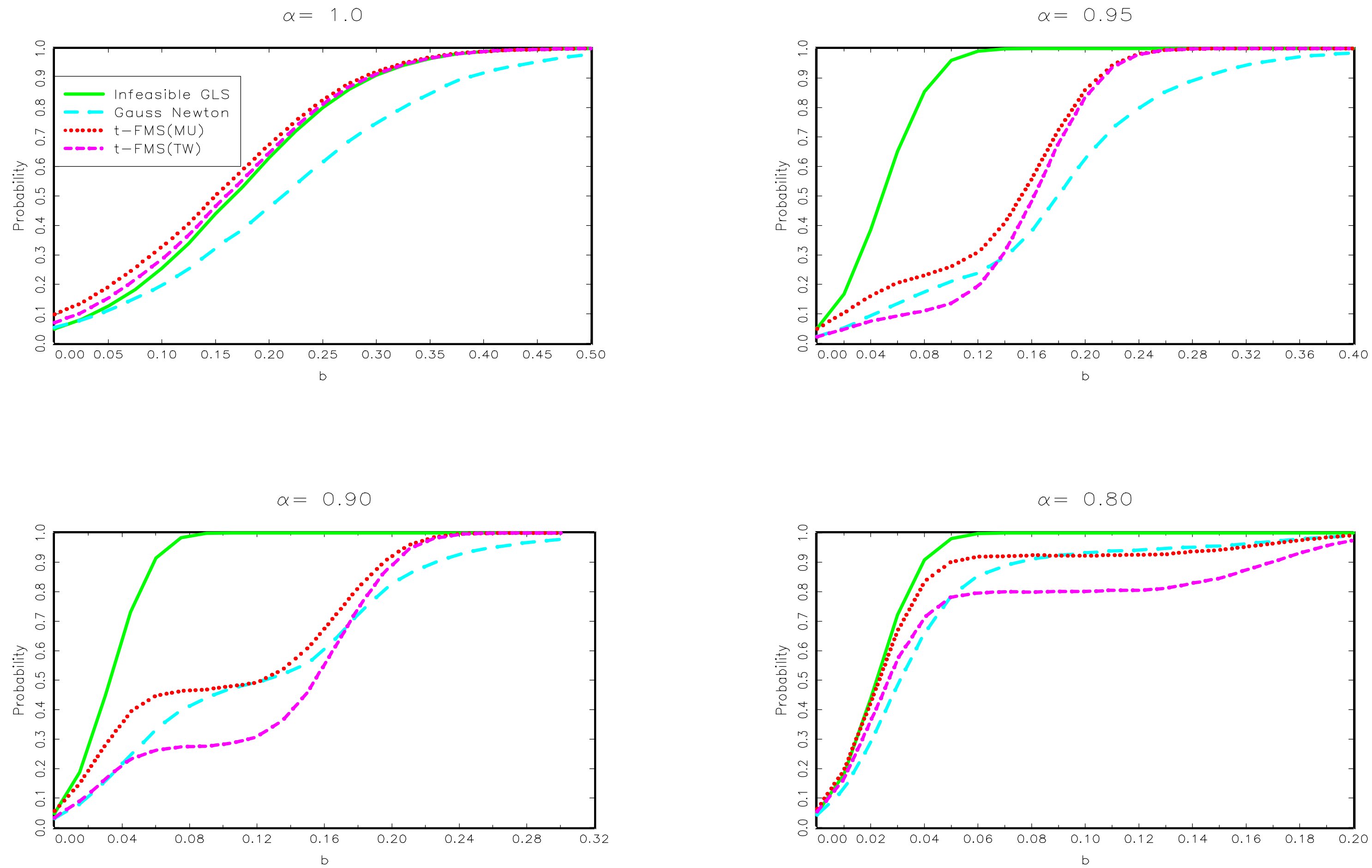

Figure 4: Finite Sample Power, AR(2) Case:

$T=100, \ \varphi=0.5$

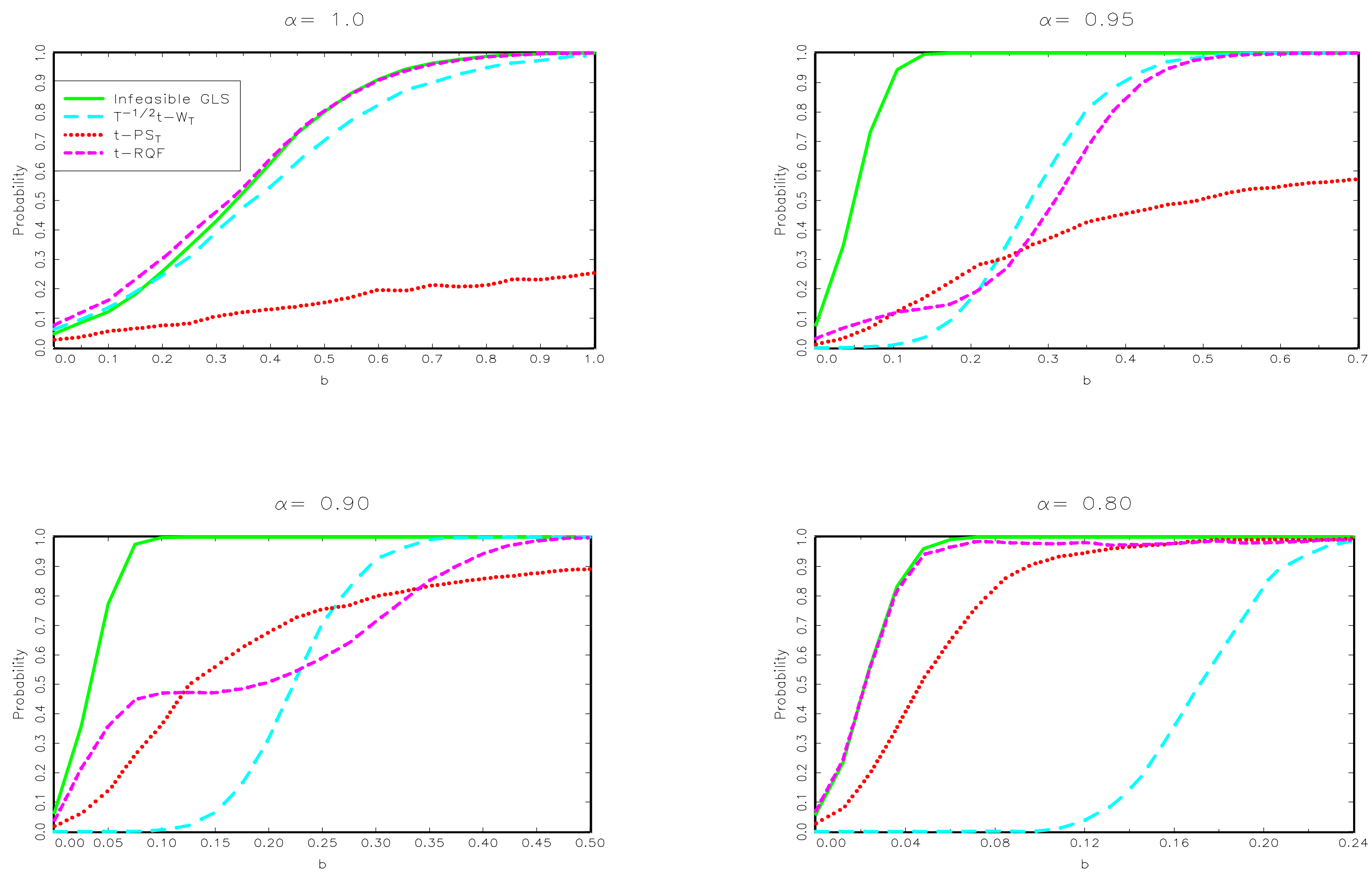

Figure 5: Sequential ADF unit root tests