IMES DISCUSSION PAPER SERIES

Testing for Shifts in Trend with an Integrated or Stationary Noise Component

Pierre Perron and Tomoyoshi Yabu

Discussion Paper No. 2006-E-2

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-2
January 2006

Testing for Shifts in Trend with an Integrated or Stationary Noise Component

Pierre Perron* and Tomoyoshi Yabu**

Abstract

This paper considers the problem of testing for structural changes in the trend function of a univariate time series without any prior knowledge as to whether the noise component is stationary or contains an autoregressive unit root. We propose a new approach that builds on the work of Perron and Yabu (2005), based on a Feasible Quasi Generalized Least Squares procedure that uses a superefficient estimate of the sum of the autoregressive parameters α when $\alpha=1$. In the case of a known break date, the resulting Wald test has a chi-square limit distribution in both the I(0) and I(1) cases. When the break date is unknown, the Exp functional of Andrews and Ploberger (1994) yields a test with nearly identical limit distributions in the two cases so that a testing procedure with nearly the same size in the I(0) and I(1) cases can be obtained. To improve the finite sample properties of the tests, we use the bias corrected version of the OLS estimate of α proposed by Roy and Fuller (2001). We show our procedure to be substantially more powerful than currently available alternatives and also to have a power function that is close to that attainable if we knew the true value of α in many cases. The extension to the case of multiple breaks is also discussed.

Keywords: Structural Change, 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 2 of Tomoyoshi Yabu's Ph.D. Dissertation at Boston University (See Yabu (2005)). We wish to thank Eiji Kurozumi and Kenji Miyazaki 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

This paper considers the problem of testing for structural changes in the trend function of a univariate time series without any prior knowledge as to whether the noise component is stationary or contains an autoregressive unit root. This problem is of practical importance for several reasons. For example, one is often interested in testing whether the rate of growth of some macroeconomic variables, such as Real GDP, exhibits a structural change. With data in logarithmic form, the coefficient on the trend component represents this average growth rate. In many cases, however, one has no prior knowledge about whether the noise component can be perceived as being stationary, or $I(0)$, or as having an autoregressive unit root, or $I(1)$. Doing a structural change test using first-differenced data or growth rates, as is commonly done, is tantamount to assuming an $I(1)$ noise component and leads to tests with very poor properties when the series has an $I(0)$ noise component. But doing a structural change test on the slope of a linear trend function for the level of the data entails different limit distributions in both cases.

On the other hand to get information about whether the noise component is $I(0)$ or $I(1)$, it is useful to have information about whether a structural change is present or not, at least in the unknown break date. If the break date is known, as considered in Perron (1989), one can have unit root tests with no knowledge about the presence or absence of a break since it is possible to have tests that are invariant to the parameters of the trend function with the inclusion of the appropriate dummy variables. In the case where the break date is unknown, things are rather different and the usual tests, suggested by Zivot and Andrews (1992) and others, are no longer invariant to the magnitude of change if one is present (see Vogelsang and Perron, 1998, and Perron, 2005, for an extensive discussion). In the presence of a change in trend both the size and the power of commonly used tests can be affected by a change in slope or intercept. To devise unit root tests with good power, information about the presence or absence of a change is needed (see, Kim and Perron, 2005).

The way to break this circular problem is to have tests for structural changes in level and/or intercept that are valid whether the noise component is $I(0)$ or $I(1)$, the object of our study. Papers related to the problem tackled here include the following. First, Perron (1991) considers extensions of Gardner's (1969) test for a change in the coefficients of a polynomial trend function. The limit is shown to be different in the $I(0)$ and $I(1)$ cases but the normalization to obtain a non-degenerate limit distribution is the same when accounting for correlation in the noise component by using a parametric autoregressive approximation.

One strategy is then to use the set of critical values that are largest across the $I(0)$ and $I(1)$ cases. However, the test suffers from important problems of non-monotonicity in power, i.e., the power function decreases as the magnitude of the change increases. Perron (1991) shows that applying the test to first-differenced data can mitigate this problem when testing for a change in slope. Vogelsang (1997) considers a similar strategy but using a Wald type procedure and shows that its power is superior. But in both cases, the critical values in the $I(0)$ case are substantially smaller than those in the $I(1)$ case so that the tests are quite conservative in the $I(0)$ case. The solution is then to devise a procedure that has the same limit distribution in both the $I(0)$ and $I(1)$ cases. The first to provide such a solution in the context of a change in the coefficients of a linear trend function is Vogelsang (2001), building on prior work related to hypothesis testing on the coefficients of a polynomial time trend reported in Vogelsang (1998). He considers again a setup where the correlation is accounted for by a parametric autoregressive approximation so that the Wald test has a non-degenerate limit distribution in both the $I(0)$ and $I(1)$ cases. The novelty is that he weights the statistic by a unit root test scaled by some parameters. For any given significance level, a value of this scaling parameter can be chosen so that the asymptotic critical values will be the same. 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 that has good properties in the $I(1)$ case but has otherwise very little power in the $I(0)$ case.

We propose a new approach that builds on the work of Perron and Yabu (2005) who analyzed the problem of hypothesis testing on the slope coefficient of a linear trend model when no information about the nature, $I(0)$ or $I(1)$, of the noise component is available. The method is based on a Feasible Quasi Generalized Least Squares approach that uses a superefficient estimate of the sum of the autoregressive parameters α when $\alpha = 1$. 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 or Chi-square distribution whether $\alpha = 1$ or $|\alpha| < 1$. Theoretical arguments and simulation evidence show that $\delta = 1/2$ is the appropriate choice.

We extend Perron and Yabu's (2005) analysis to the case of testing for changes in level or slope of the trend function of a univariate time series. When the break date is known, things are relatively straightforward as their asymptotic results directly apply, in particular the standard Wald test from the GLS regression with the truncated estimate of α has a

Chi-square limit distribution irrespective of the nature of the noise components, i.e., with $I(0)$ or $I(1)$ errors. Our analysis shows that the procedure has in this case good finite sample properties and a power function that is close to that attainable with the infeasible GLS procedure that uses the true value of α. When the break dates are unknown, the limit distributions of the Exp, Mean and Sup functionals of the Wald test across all permissible breaks dates (see Andrews and Ploberger, 1994) is no longer the same in the $I(0)$ and $I(1)$ cases. It turns out, however, that the limit distribution is nearly the same when considering the Exp functional. Hence, it is possible to have tests with nearly the same size in both cases. To improve the finite sample properties of the test, we also use a bias-corrected version of the OLS estimate of α (obtained from an autoregression based on the residuals from estimating the parameters of the trend function by OLS) as suggested by Roy and Fuller (2001). This makes possible a testing procedure that has good size and power properties in finite sample.

Since the only other procedure that delivers a testing procedure valid with both $I(0)$ and $I(1)$ errors is that of Vogelsang (2001), we make extensive comparisons of the finite sample power of both set of tests. We show our procedure to be substantially more powerful and also to have a power function that is close to that attainable if we knew the true value of α in many cases.

The paper is organized as follows. Section 2 considers the basic case where the noise component has an AR(1) structure and there is a single break. This permits a presentation of the main ingredients and properties of our suggested procedure. Section 3 presents an evaluation of the finite sample properties of our tests. Section 4 extends the analysis to the case of a general correlation structure for the noise component. The extension to the case of more than one break is discussed in Section 5. Section 6 presents empirical applications related to real GDP series for a variety of counties. Section 7 offers brief concluding remarks and an appendix some technical derivations.

2 The AR(1) Case

We start with the simple case of a single break where the noise component is an $AR(1)$ in order to highlight the main issues involved. Extensions to the general case are presented in Section 4 and 5. The data generating process for a scalar random variable y_t is assumed to be:

$$\begin{aligned} y_t &= x_t'\Psi + u_t \qquad (1) \\ u_t &= \alpha u_{t-1} + e_t \end{aligned}$$

for $t = 1, ..., T$ where $e_t \sim i.i.d.\ (0, \sigma^2)$, x_t is a $(r \times 1)$ vector of deterministic components, and Ψ is a $(r \times 1)$ vector of unknown parameters, which are model specific and defined below. For simplicity, we let u_0 be some finite constant. Here $-1 < \alpha \leq 1$, so that both stationary and integrated errors are allowed.

We are interested in testing the null hypothesis $R\Psi = \gamma$ where R is a $(q \times r)$ full rank matrix and γ is a $(q \times 1)$ vector, q being the number of restrictions. The restrictions to be considered pertain to testing whether a structural change in a trend function occurs. Hence, we shall consider the following three models where a change of intercept and/or slope in a trend function occurs. Throughout, the break date is denoted $T_1 = [\lambda_1 T]$ for some $\lambda_1 \in (0, 1)$, where $[\cdot]$ denotes the largest integer that is less than or equal to the argument. Also, $1(.)$ is the indicator function.

Model I: Structural change in intercept. Here $x_t = (1, DU_t, t)'$ and $\Psi = (\mu_0, \mu_1, \beta_0)'$ where $DU_t = 1(t > T_1)$. This specification allows for one-time change in the intercept. The hypothesis of interest is $\mu_1 = 0$.

Model II: Structural change in slope. Here $x_t = (1, t, DT_t)'$, $\Psi = (\mu_0, \beta_0, \beta_1)'$ where $DT_t = 1(t > T_1)(t - T_1)$. This specification allows for a one-time change in the slope of the trend without a change in level so that the trend function is joined at the time of break. The hypothesis of interest is $\beta_1 = 0$.

Model III: Structural change in both intercept and slope. Here $x_t = (1, DU_t, t, DT_t)'$ and $\Psi = (\mu_0, \mu_1, \beta_0, \beta_1)'$. This specification allows for a simultaneous change in the intercept and slope. The hypothesis of interest is $\mu_1 = \beta_1 = 0$.

2.1 The Feasible GLS Estimate

We first consider the case where the break date T_1 is known and return to the unknown break date case in Section 2.4. Suppose, in addition, that the autoregressive coefficient α is known. The GLS estimate of the parameters is then obtained by applying Ordinary Least Squares (OLS) to the regression

$$(1 - \alpha L)y_t = (1 - \alpha L)x_t'\Psi + (1 - \alpha L)u_t$$

for $t = 2, ..., T$ together with

$$y_1 = x_1'\Psi + u_1.$$

It is well known that, in such a case, the Wald-statistic for testing the null hypothesis, $R\Psi = \gamma$, is asymptotically distributed as a chi-square random variable for any values of α.

In practice, however, α is unknown and using an estimate can make things very different. Consider the Cochrane-Orcutt (1949) FGLS estimate of Ψ that uses the following estimate of α:

$$\hat{\alpha} = \sum_{t=2}^{T} \hat{u}_t \hat{u}_{t-1} / \sum_{t=2}^{T} \hat{u}_{t-1}^2 \tag{2}$$

where $\{\hat{u}_t\}$ are the OLS residuals from the regression of y_t on x_t. When $|\alpha| < 1$, $T^{1/2}(\hat{\alpha} - \alpha) \to^d N(0, 1 - \alpha^2)$ and the Wald statistic obtained from the FGLS procedure still has a $\chi^2(q)$ limit distribution. Things are different when $\alpha = 1$. From standard results,

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

where $W^*(r)$, $0 \leq r \leq 1$, is the continuous time residual function from a projection of a Wiener process $W(r)$ on the continuous-time version of the deterministic components (e.g., $\{1, 1(r > \lambda_1), r\}$ for Model I). The limit of the Wald statistic constructed from the FGLS regression no longer has a chi-square limit distribution. To illustrate the problem, consider the simple case with $x_t = DT_t$ and $\Psi = \beta_1$. The limit distribution of the t-statistic for testing $\beta_1 = 0$ is given by (see the appendix for details):

$$t_F \Longrightarrow \frac{W(1) - W(\lambda_1) - \kappa \int_{\lambda_1}^1 W(r)dr - \kappa[\int_{\lambda_1}^1 (r - \lambda_1) dW(r) - \kappa \int_{\lambda_1}^1 (r - \lambda_1) W(r) dr]}{[(1 - \lambda_1) - \kappa(1 - \lambda_1)^2 + \kappa^2 (1 - \lambda_1)^3 / 3]^{1/2}} \tag{3}$$

which is different from a standard Normal distribution. Accordingly, the limit of the Wald statistic, $W_F = {t_F}^2$, is no longer chi-square distributed.

Suppose now that κ, the limit of $T(\hat{\alpha} - 1)$, was zero, then, $t_F \Rightarrow [W(1) - W(\lambda_1)]/(1 - \lambda_1)^{1/2} =^d N(0, 1)$ and hence $W_F \Rightarrow [W(1) - W(\lambda_1)]^2/(1 - \lambda_1) =^d \chi^2(1)$. We would then recover in the unit root case the same limiting distribution as in the stationary case and no discontinuity would be present. Perron and Yabu (2005) exploit this fact to derive a testing procedure with the same limit distribution in both the $I(0)$ and $I(1)$ cases.

2.2 A Super-Efficient Estimate When $\alpha = 1$

Perron and Yabu (2005) proposed the use of a supper efficient estimate of α when $\alpha = 1$ defined by:

$$\hat{\alpha}_S = \begin{cases} \hat{\alpha} & if \quad T^{\delta} |\hat{\alpha} - 1| > d \\ 1 & if \quad T^{\delta} |\hat{\alpha} - 1| \leq d \end{cases} \tag{4}$$

for $\delta \in (0,1)$ and $d > 0$. Hence, when $\hat{\alpha}$ is in a $T^{-\delta}$ neighborhood of 1 it is assigned a value of 1. Perron and Yabu (2005) proved that (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, κ, the limit of $T(\hat{\alpha}_S - 1)$, is zero when the errors are $I(1)$. The following theorem, proved in the Appendix, shows that tests based on the FGLS procedure using $\hat{\alpha}_S$ have a $\chi^2(q)$ distribution in both the $I(0)$ and $I(1)$ cases.

Theorem 1 *Let $\hat{\Psi}$ be the FGLS estimate of Ψ using $\hat{\alpha}_S$, i.e., the OLS estimate from a regression of $y_t^{\hat{\alpha}_S}$ on $x_t^{\hat{\alpha}_S}$ where $y_t^{\hat{\alpha}_S} = (1 - \hat{\alpha}_S L)y_t$ and $x_t^{\hat{\alpha}_S} = (1 - \hat{\alpha}_S L)x_t$ for $t = 2, ..., T$, $y_1^{\hat{\alpha}_S} = y_1$ and $x_1^{\hat{\alpha}_S} = x_1$. Denote the residuals associated with this regression by $\hat{e}_t$. The Wald statistic for testing the null hypothesis $R\Psi = \gamma$ is:*

$$W_{FS}(\lambda_1) = [R(\hat{\Psi} - \Psi)]'[s^2 R(X'X)^{-1}R']^{-1}[R(\hat{\Psi} - \Psi)] \tag{5}$$

where $X = \{x_t^{\hat{\alpha}_S}\}$ and $s^2 = T^{-1}\sum_{t=1}^{T}\hat{e}_t^2$. Then, if $|\alpha| < 1$,

$$\begin{aligned} W_{FS}(\lambda_1) &\Longrightarrow [R(\int_0^1 F(r,\lambda_1)F(r,\lambda_1)'dr)^{-1}\int_0^1 F(r,\lambda_1)dW(r)]'[R(\int_0^1 F(r,\lambda_1)F(r,\lambda_1)'dr)^{-1}R']^{-1} \\ &\quad \times [R(\int_0^1 F(r,\lambda_1)F(r,\lambda_1)'dr)^{-1}\int_0^1 F(r,\lambda_1)dW(r)] \\ &\equiv G_0(\lambda_1) =^d \chi^2(q) \end{aligned}$$

where $F(r,\lambda_1) = [1, 1(r > \lambda_1), r]'$, $[1, r, 1(r > \lambda_1)(r - \lambda_1))]'$, and $[1, 1(r > \lambda_1), r, 1(r > \lambda_1)(r - \lambda_1)]'$ for Models I, II, and III, respectively. If $\alpha = 1$,

$$\begin{aligned} W_{FS}(\lambda_1) &\Longrightarrow \begin{cases} \lim_{T\to\infty} e^2_{[\lambda_1 T]+1}/\sigma^2 & \text{for Model I} \\ [\lambda_1 W(1) - W(\lambda_1)]^2/[\lambda_1(1-\lambda_1)] & \text{for Model II} \\ \lim_{T\to\infty} e^2_{[\lambda_1 T]+1}/\sigma^2 + [\lambda_1 W(1) - W(\lambda_1)]^2/[\lambda_1(1-\lambda_1)] & \text{for Model III} \end{cases} \\ &\equiv G_1(\lambda_1) \end{aligned}$$

The limit distribution is $\chi^2(q)$ for Model II, and if the errors are Normally distributed, also for Models I and III.

Therefore, constructing the GLS regression with the truncated estimate, $\hat{\alpha}_S$, effectively bridges the gap between the $I(0)$ and $I(1)$ cases, and the chi-square asymptotic distribution is obtained in both cases.

Remark 1 *The limit chi-square distribution requires the assumption of Normality of the errors for Models I and III to ensure that $\lim_{T\to\infty} e^2_{[\lambda_1 T]+1}/\sigma^2 \backsim \chi^2(1)$. This is unavoidable* *since, when using a GLS type procedure when $\alpha = 1$, a change in level becomes a single outlier. For tests of hypotheses that do not involve the change in intercept, the Normality assumption is not required.*

2.3 The case with α local to unity

The result obtained in Theorem 1 is pointwise in α for $-1 < \alpha \leq 1$ and does 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 that $\alpha = 1 + c/T$, then*

$$
\begin{aligned}
W_{FS}(\lambda_1) &\Longrightarrow \begin{cases} \lim_{T\to\infty} e^2_{[\lambda_1 T]+1}/\sigma^2 & \text{for Model I} \\ [\lambda_1 J_c(1) - J_c(\lambda_1)]^2/[\lambda_1(1-\lambda_1)] & \text{for Model II} \\ \lim_{T\to\infty} e^2_{[\lambda_1 T]+1}/\sigma^2 + [\lambda_1 J_c(1) - J_c(\lambda_1)]^2/[\lambda_1(1-\lambda_1)] & \text{for Model III} \end{cases} \\
&\equiv G_c(\lambda_1)
\end{aligned}
$$

where $J_c(r) = \int_0^r \exp(c(r-s))dW(s) \sim N(0, (\exp(2cr)-1)/2c)$.

The results are fairly intuitive. Since the true value of α is in a T^{-1} neighborhood of 1, and $\hat{\alpha}_S$ truncates the 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 $\hat{\alpha}_S = 1$. Hence, the estimator of Ψ is essentially the same as the first-difference estimator. Let us consider the case of Model II. The first-difference estimator of β_1 is:

$$\hat{\beta}_1^{FD} = \frac{\sum_{t=1}^T (DU_t - T^{-1}(T-T_1))\Delta u_t}{\sum_{t=1}^T (DU_t - T^{-1}(T-T_1))^2}$$

Then, the Wald statistic has the following limit distribution:

$$
\begin{aligned}
W_{FD} &= \frac{[\sum_{t=1}^T DU_t \Delta u_t - T^{-1}(T-T_1)\sum_{t=1}^T \Delta u_t]^2}{s^2 \sum_{t=1}^T (DU_t - T^{-1}(T-T_1))^2} \\
&= \frac{[T^{-1/2}(u_T - u_{T_1}) - T^{-1}(T-T_1)T^{-1/2}(u_T - u_0)]^2}{s^2 T^{-1}\sum_{t=1}^T (DU_t - T^{-1}(T-T_1))^2} \\
&\Rightarrow \frac{[\lambda_1 J_c(1) - J_c(\lambda_1)]^2}{\lambda_1(1-\lambda_1)}
\end{aligned}
$$

since $T^{-1}\sum_{t=1}^{T}(DU_t - T^{-1}(T-T_1))^2 \to \lambda_1(1-\lambda_1)$, $T^{-1/2}u_{[rT]} \Rightarrow \sigma J_c(r)$, $T^{-1/2}u_0 = o_p(1)$, and $s^2 = \sigma^2 + o_p(1)$. This limit distribution is the same as that in Theorem 2 for Model II.

Note that when $c = 0$, we recover the result of Theorem 1 for the $I(1)$ case. However, for Models II and III, when $c < 0$, the variance of $\lambda_1 J_c(1) - J_c(\lambda_1)$ is different from $\lambda_1(1-\lambda_1)$ and the upper quantiles of the limit distributions are, accordingly, smaller than those of a $\chi^2(q)$, 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 of the component $[\lambda_1 J_c(1) - J_c(\lambda_1)]$ is 0 as $c \to -\infty$, 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. For the case of tests on the coefficients of a linear trend function, Perron and Yabu (2005) showed that indeed, the t-statistic has a $N(0,1)$ limit distribution as $c \to -\infty$. What is especially interesting is that to obtain this result, a condition on δ needs to be imposed, namely that $\delta \geq 1/2$. Their result extends in a straightforward way to the present setup. This result is important for the following reason. 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 tests have the least possible size distortions, we need $\delta \geq 1/2$. 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 in Perron and Yabu (2005) show that this value leads to a procedure which works best in small samples. We also verified by simulations that $\delta = 1/2$ is the best choice for the Models considered here. Hence, we continue to use this value and will calibrate the appropriate value of d using simulations.

Remark 2 *It is important to remark that in Model I, the limit distribution does not depend on c. Therefore, we can expect the tests in this case to be unaffected by size distortions for values of α close to but not equal to* 1*. Simulation results will show that this is in fact the case.*

2.4 The Case with an Unknown Break Date

In general, the break date is unknown, and the analysis must be extended accordingly. Following Andrews (1993) and Andrews and Ploberger (1994), we consider the following three functionals of the Wald test for different break dates:

$$\begin{aligned} Mean\text{-}W_{FS} &= T^{-1}\sum_{\Lambda} W_{FS}\left(\lambda_1'\right) \\ Exp\text{-}W_{FS} &= \log\left[T^{-1}\sum_{\Lambda}\exp\left(\frac{1}{2}W_{FS}\left(\lambda_1'\right)\right)\right] \\ \sup\text{-}W_{FS} &= \sup_{\Lambda} W_{FS}\left(\lambda_1'\right) \end{aligned} \tag{6}$$

where $\Lambda = \{\lambda_1'; \epsilon \leq \lambda_1' \leq 1-\epsilon\}$ for some $\epsilon > 0$. Here T_1' (resp., λ_1') denotes a generic break date (resp., break fraction) used to construct a particular value of the Wald test, while T_1 (resp., λ_1) will continue to denote the true break date (resp., break fraction). Given that the regressors are trending and the errors may have a unit root, none of these functionals need have any optimal properties in our case. Nevertheless, given that they have some optimality properties in the stationary case, it is worth considering them. Their limit distributions are a simple consequence of Theorem 1, stated in the following Corollary.

Corollary 1 *Let $g(\lambda_1')$ denote $G_0(\lambda_1')$ and $G_1(\lambda_1')$ for the I(0) and I(1) cases, respectively, as defined in Theorem 1. Then,*

$$\begin{aligned} Mean\text{-}W_{FS} &\Longrightarrow \int_{\Lambda} g(\lambda_1')d\lambda_1' \\ Exp\text{-}W_{FS} &\Longrightarrow \log\left[\int_{\Lambda}\exp\left(\frac{1}{2}g(\lambda_1')\right)d\lambda_1'\right] \\ \sup\text{-}W_{FS} &\Longrightarrow \sup_{\Lambda} g(\lambda_1') \end{aligned}$$

Here things are not as simple as in the known break date case. Even though the distributions of $G_0(\lambda_1')$ and $G_1(\lambda_1')$ are both chi-square for any fixed λ_1', the two functions are different. Accordingly, the limit distributions of the three tests are different in the $I(0)$ and $I(1)$ cases. Table 1 presents the asymptotic critical values with a value for the trimming parameter $\epsilon = 0.01$ and 0.15 [1]. What transpires from these results is that, though the limit distributions are different, the relevant quantiles are very similar for the $I(0)$ and $I(1)$ cases when considering the *Exp* functional. For example, for Model II, the 5% critical values with $\epsilon = 0.01$ are 1.97 and 2.02 for the I(0) and I(1) cases, respectively. Therefore, taking the

[1]The critical values were calculated by simulations using *i.i.d.* $N(0,1)$ random variables to approximate the Wiener process. The integrals are approximated by normalized sums with 2000 steps, and the number of replications used is 10,000.

larger critical value, corresponding to 2.02, is expected to bring a powerful robust statistic for both stationary and integrated errors.

Table 2 presents an extended grid of critical values for the *Exp* version of the test, which also includes a wider range of values for the trimming parameter ϵ. From this Table one can deduce what is the asymptotic size of the test in both cases when the larger critical value is used. For example, for Model I, when doing a 5% test, the critical values are largest for the $I(0)$ case and the asymptotic size in the $I(1)$ case varies between 0.045 for $\epsilon = .25$ and 0.035 for $\epsilon = 0.10$. For Model II, the critical values are largest in the $I(1)$ case but the asymptotic size in the $I(0)$ case is between 0.045 and 0.05 for all value of ϵ. For Model III, the relevant critical values correspond to those in the $I(1)$ case and the size in the $I(0)$ case is between 0.035 and 0.045. The discrepancies for other significance levels can be obtained from the Table. Hence, overall, we see that the minimum value of the size for the $I(0)$ and $I(1)$ cases is close to 5% and imply tests that are not very conservative, which is a useful feature to get decent power.

Remark 3 *One may be interested in using Model III and test only whether a shift in slope is present (β_1 with μ_1 unrestricted), i.e., to allow the possibility of both a shift in intercept and slope but only test if the latter is present. In that case, the limit distribution in the $I(1)$ case is the same as in Model II. In the $I(0)$ case, it is different but very close to that in the $I(1)$ case, indeed even closer than for Model II. Hence, one can still use the critical values corresponding to Model II.*

The results also show the Mean and Sup versions of the test to be of little use in our context. The discrepancies between the two sets of critical values being quite large, these tests would implies substantial size distortions in either the $I(0)$ or $I(1)$ case depending on which case has the largest critical value. Some of the results in Table 1 can be explained using theoretical arguments. First, for Model I, the limit of the Mean test in the $I(1)$ case is $(1 - 2\epsilon)$, or 0.98 and 0.7 for $\epsilon = 0.01, 0.15$, respectively. The reason is that in large samples, $W_{FS}(\lambda_1')$ is equivalent to $e^2_{T_1'+1}/\sigma^2$ and hence the Mean test is $T^{-1}\sum_\Lambda W_{FS}(\lambda_1') = (1-2\epsilon)[(\sigma^2(1-2\epsilon)T)^{-1}\sum_\Lambda e^2_{T_1'+1}] + o_p(1) \to^p (1-2\epsilon)$ using the fact that, from a standard Law of Large Numbers $[(1-2\epsilon)T]^{-1}\sum_\Lambda e^2_{T_1'+1} \to^p \sigma^2$. To explain why, for the $I(1)$ case, the Sup version diverges in Models I (the explanation is similar for Model III), we again use the fact that in large samples, $W_{FS}(\lambda_1')$ is equivalent to $e^2_{T_1'+1}/\sigma^2$. Thus, the sup statistic is $\sup_\Lambda e^2_{T_1'+1}/\ \sigma^2 = \max_\Lambda\{e^2_{T_1'+1}/\sigma^2\}$. The probability of sup-$W_{FS} < \eta$ is $Pr(\Pi_\Lambda \cap \{e^2_{T_1'+1}/\sigma^2 < \eta\}) = \Pr(e_t^2/\sigma^2 < \eta)^{(1-2\epsilon)T}$ using the assumption of independence. Therefore, Pr(sup-

$W_{FS} < \eta) \to 0$ as $T \to \infty$ if $\Pr(e_t^2/\sigma^2 < \eta) < 1$, which is the case for any $\eta < \infty$. Hence, sup-W_{FS} diverges to infinite as $T \to \infty$.

2.5 Some Modifications to Improve Finite Sample Properties

It is well known that the OLS estimate of α is biased downward especially when α is near one. Hence, in many cases no truncation may apply when some would be desirable. Perron and Yabu (2005) recommend the use of a median-unbiased estimate. In the current context, however, obtaining an exactly median unbiased estimate is computationally very demanding, especially in the unknown break date case and more so when a more general $AR(p)$ structure for the noise component is entertained (as will be in Section 4). An alternative approach with similar finite sample properties is the bias correction proposed by Roy and Fuller (2001), which we shall adopt here.

Roy and Fuller (2001) proposed a class of bias corrected estimators and we consider here that based on the OLS estimate [2]. It is a function of a unit root test, namely the t-ratio $\hat{\tau} = (\hat{\alpha} - 1)/\hat{\sigma}_\alpha$, where $\hat{\alpha}$ is the OLS estimate and $\hat{\sigma}_\alpha$ its standard deviation. The bias-corrected estimate is given by

$$\hat{\alpha}_M = \hat{\alpha} + C(\hat{\tau})\hat{\sigma}_\alpha, \tag{7}$$

$$C(\hat{\tau}) = \begin{cases} -\hat{\tau} & \text{if} \quad \hat{\tau} > \tau_{pct} \\ I_p T^{-1}\hat{\tau} - (1+r)[\hat{\tau} + c_2(\hat{\tau} + a)]^{-1} & \text{if} \quad -a < \hat{\tau} \leq \tau_{pct} \\ I_p T^{-1}\hat{\tau} - (1+r)\hat{\tau}^{-1} & \text{if} \quad -c_1^{1/2} < \hat{\tau} \leq -a \\ 0 & \text{if} \quad \hat{\tau} \leq -c_1^{1/2} \end{cases}$$

where τ_{pct} is some percentile of the limiting distribution of $\hat{\tau}$ when $\alpha = 1$, $c_1 = (1+r)T$ with r the number of parameters estimated, $c_2 = [(1+r)T - \tau_{pct}^2(I_p+T)][\tau_{pct}(a+\tau_{pct})(I_p+T)]^{-1}$ and a is some constant. Also, $I_p = [(p+1)/2]$ where p is the order of the autoregressive process considered for the noise component (here 1 but different in the generalizations considered in Section 4). The parameters for which specific values need to be selected are τ_{pct} and a. Based on extensive simulation experiments, we selected $a = 10$ since it leads to tests with better properties [3], and for τ_{pct} we use $\tau_{.95}$ for the known break date case and $\tau_{.99}$ for

[2] Roy et al. (2004) and Perron and Yabu (2005) use a similar bias corrected estimate but based on a weighted symmetric least-squares estimate of α instead of the OLS estimate used here. Both lead to tests with similar properties.

[3] Roy and Fuller (2001) select $a = 5$ in the context of a linear trend model. However, with a broken linear trend the tails of the distribution of $\hat{\tau}$ are more important and chosing $a = 5$ would make the middle parts of the modifications redundant.

the unknown break date case (see Yabu, 2005, for details). Also, again based on extensive simulations, we found that the value $d = 1$, for the truncation, leads to the best results in finite samples. Hence, our suggested procedure involves the following steps:

1. For any given break date, detrend the data by OLS to obtain residuals, say $\hat{u}_t$;

2. Estimate an autoregression of order one for $\hat{u}_t$ yielding the estimate $\hat{\alpha}$ and the t-ratio $\hat{\tau}$;

3. Use $\hat{\alpha}$ and $\hat{\tau}$ to get the Roy and Fuller (2001) biase corrected estimates $\hat{\alpha}_M$;

4. Apply the truncation

$$\hat{\alpha}_{MS} = \begin{cases} \hat{\alpha}_M & \text{if} \quad |\hat{\alpha}_M - 1| > T^{-1/2} \\ 1 & \text{if} \quad |\hat{\alpha}_M - 1| \leq T^{-1/2} \end{cases}$$

5. Apply a GLS procedure with $\hat{\alpha}_{MS}$ to obtain the estimates of the coefficients of the trend and the estimate of the variance of the residuals and construct the standard Wald-statistic, which we shall denote by W_{FMS}.

6. When dealing with the case of an unknown break date, repeat the 5 steps above for all permissible break dates and construct the Exp-Wald statistic, denoted by *Exp*-W_{FMS}.

Remark 4 *Using the biased corrected versions $\hat{\alpha}_M$, instead of the OLS estimates, does not change anything to the stated large sample results (Theorems 1 and Corollary 1). All that is needed for these asymptotic results to hold is that $T(\hat{\alpha}_M - 1) = O_p(1)$ when $\alpha = 1$, and $T^{1/2}(\hat{\alpha}_M - \alpha) \to^d N(0, 1 - \alpha^2)$ when $\alpha < 1$. These conditions are satisfied.*

3 Simulation Evidence

The aim in this section is to assess the finite sample properties of the procedure. We first consider the size of the tests. To that effect the data are generated by a simple AR(1) process of the form $y_t = \alpha y_{t-1} + e_t$ with $e_t \sim i.i.d.\ N(0, 1)$ and $y_0 = 0$ (setting the constant and trend parameters to zero is without loss of generality due to the fact that the tests are invariant to them). The nominal size of the tests is 5% throughout. We consider the following sample sizes, $T = 100$, 250, 500 for a known break date and $T = 100$, 250 for an unknown break date. The value of the trimming parameter for the *Exp*-W_{FMS} test in the unknown break date case is set to $\epsilon = 0.01$. The numbers of replications for the known and unknown break date cases are $5,000$ and $2,000$, respectively. Using these specifications, we evaluate the size of the tests for values of α in the range of $[0,1]$ with increments of 0.05.

Figure 1 presents the results. In the case of a known break date, the size of the test is very close to 5% for all values of the autoregressive parameters in all cases considered even when $T = 100$. As expected the test is slightly conservative for Models II and III when α is close to but not equal to one. In the case of an unknown break date, the test shows some liberal size distortions when $T = 100$ and a change in intercept is involved (Models I and III). But these distortions are virtually eliminated when $T = 250$, except when α is close to but not equal to one, where the test is somewhat conservative, as expected.

We now consider the power of the tests. The specifications are the same except that the data are generated by the following processes

- For Model I: $y_t = \eta DU_t + u_t$;
- For Model II: $y_t = \eta DT_t + u_t$;
- For Model III: $y_t = \eta(10DU_t + DT_t) + u_t$;

where $u_t = \alpha u_{t-1} + e_t$ with $e_t \sim i.i.d.\ N(0,1)$ and $u_0 = 0$. The break date is set to occur at mid-sample, i.e., $T_1 = [0.5T]$.

The power of our test is compared to three other tests: the Wald test based on the infeasible GLS estimates which uses the true values of α and T_1, as well as the $T^{-1}W_T$ and the PS_T tests of Vogelsang (2001) and its Exp version in the case of an unknown break. For Vogelsang's tests, 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 $\alpha = 1.0$, $0.95, 0.90, 0.80$ and a range of values of $\eta > 0$.

Consider first Figure 2, which presents the results for Model I in the case of a known break date for $T = 100$. For any values of α, our test is as powerful as that based on the the infeasible GLS regression that uses the true value of α. The test is also substantially more powerful than those of Vogelsang (2001). These results are quite remarkable since our test, as any other, is inconsistent when α is local to 1 (which is the case for the values of α considered) since the intercept shift is then basically only an outlier. But an inconsistent test can still have decent power in finite samples. As expected, however, power does not increase with the sample size but only with the magnitude of the change, accordingly the results for $T = 250$ and $T = 500$ are basically identical and not reported. The results for the case of an unknown break date, presented in Figure 3 for $T = 100$, yield similar conclusions with the exception that the power of our test is somewhat below what can be achieved with the infeasible GLS procedure. Still, power rises rapidly with the magnitude of the change.

Consider now Model II. The power functions for the known break date case are presented in Figures 4.a through 4.c, for $T = 100, 250$ and 500, respectively. When $\alpha = 1$, our test is as powerful as that based on the infeasible GLS estimates for any T. When $\alpha = 0.8, 0.9, 0.95$ and $T = 100$, the power of our test is lower, in large part due to the fact that the size in this case is conservative. Still it offers important power improvements over Vogelsang's tests. As T increases, the power function of our test gets closer to that of the infeasible GLS test. For instance, the power functions are equivalent when $T = 500$ and $\alpha = 0.8, 0.9$. The power functions for the unknown break date case are presented in Figures 5.a through 5.b, for $T = 100$ and 250, respectively. The results are qualitatively similar with the power functions being, as expected, lower. Also, our test no longer globally dominates both the $T^{-1}W$ and PS_T tests of Vogelsang (2001), though individually each of these tests has very poor power for some values of α. Figures 6.a through 6.c for the known break date case and Figures 7.a through 7.b for the unknown break date case show that similar features hold for Model III.

Overall these results are very encouraging and points to the usefulness of our testing procedure. Below we extend them in two directions; first to allow a general serial correlation structure in the noise component, and second to allow more than one break.

4 Generalization of the Error Component

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. The data generating process is now assumed to be given by (1) with u_t specified by

$$\begin{aligned} u_t &= \alpha u_{t-1} + v_t \\ v_t &= d(L) e_t \end{aligned} \tag{8}$$

with

$$d(L) = \sum_{i=0}^{\infty} d_i L^i, \sum_{i=0}^{\infty} i|d_i| < \infty,\ d(1) \neq 0.$$

and $e_t \sim i.i.d.\ (0, \sigma^2)$. Again, we assume for simplicity that u_0 is some constant. These conditions imply that the following functional central limit theorem holds 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 (8), we wish to have the parameter α represent the sum of the autoregressive coefficients. Accordingly, we have 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 $a_i^* = -\sum_{j=i+1}^{\infty} a_j$.

4.1 The estimation of α

Since α represents the sum of the autoregressive coefficients, we cannot use the estimate (2) 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 x_t, then the estimate of α considered is the OLS estimate $\widetilde{\alpha}$ obtained from the regression:

$$\hat{u}_t = \alpha \hat{u}_{t-1} + \sum_{i=1}^{k} \zeta_i \Delta \hat{u}_{t-i} + e_{tk} \tag{9}$$

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 a regression of $J_c(r) \equiv \int_0^r \exp(c(r-s))dW(s)$ on the continuous time version of the deterministic components. We also use the same bias correction for the lest-squares estimate as described in Section 2.5, $\widetilde{\alpha}_M$, with the same specifications, and we apply the truncation (4) with $\widetilde{\alpha}_M$ instead of $\hat{\alpha}$. The truncated version of the bias-corrected estimate, $\widetilde{\alpha}_{MS}$, is still superefficient under a local unit root, i.e. $T(\widetilde{\alpha}_{MS} - 1) \rightarrow_p 0$ when $\alpha = 1 + c/T$. As usual, consistency of $\widetilde{\alpha}$ and the truncated bias-corrected version, $\widetilde{\alpha}_{MS}$, depends on having the autoregressive approximation k increase at a suitable rate as T increases, namely $k \to \infty$ and $k^3/T \to 0$ as $T \to \infty$, see Berk (1974). Below we suggest using an information criterion such as the BIC, which ensures that these properties are satisfied.

4.2 The Test Statistics

The estimate of Ψ considered is a quasi-FGLS estimate assuming $AR(1)$ errors, i.e. the OLS estimate in the transformed regression:

$$(1 - \widetilde{\alpha}_{MS}L)y_t = (1 - \widetilde{\alpha}_{MS}L)x_t'\Psi + (1 - \widetilde{\alpha}_{MS}L)u_t \tag{10}$$

for $t = 2, ..., T$, together with

$$y_1 = x_1'\Psi + u_1$$

Denote the resulting estimates by $\widetilde{\Psi}$. The specific form of the Wald test depends on the nature of the errors, $I(0)$ or $I(1)$, and the model.

Consider first the $I(0)$ case. Then for any of the three models, one simply needs to replace s^2 in (5) by $\hat{h}_v$, an estimate of (2π times) the spectral density function at frequency zero of $v_t = (1 - \alpha L)u_t$. We denote the resulting Wald test by $W_{RQF}(\lambda_1)$ where the subscript RQF stands for Robust Quasi Feasible GLS, more precisely

$$W_{RQF}(\lambda_1) = [R(\widetilde{\Psi} - \Psi)]'[\hat{h}_v R(X'X)^{-1}R']^{-1}[R(\widetilde{\Psi} - \Psi)] \tag{11}$$

where $X = \{x_t^{\widetilde{\alpha}_{MS}}\}$, $x_t^{\widetilde{\alpha}_{MS}} = (1 - \widetilde{\alpha}_{MS}L)x_t$ for $t = 2, ..., T$, and $x_1^{\widetilde{\alpha}_{MS}} = x_1$. We consider two types of estimates $\hat{h}_v$. One is based on a weighted sum of the autocovariance function given by:

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

for some weight function $\omega(j, m)$ and bandwidth m, where $\hat{v}_t$ are the OLS residuals from the regression (10). The quadratic spectral window is used with the bandwidth m selected according to the "plug-in method" advocated by Andrews (1991) using an AR(1) approximation.

The other estimate considered is an autoregressive spectral density estimate (at frequency zero). Note that the residuals from the regression (10) are $(1 - \widetilde{\alpha}_S L)u_t$ and, hence, in large samples, they are equivalent to $v_t = (1 - \alpha L)u_t$. When $|\alpha| < 1$, $A(L)$ is invertible so that $u_t = A(L)^{-1}e_t$ and thus $v_t = (1 - \alpha L)A(L)^{-1}e_t$. 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

$$y_t - \widetilde{\alpha}_S y_{t-1} = x_t'\Psi^* + \sum_{i=1}^{k}\rho_i \Delta y_{t-i} + e_{tk}$$

with $\hat{e}_{tk}$ the corresponding OLS residuals. The estimate of σ^2 is then $\hat{\sigma}^2 = \hat{h}_v = (T - k)^{-1}\sum_{t=k+1}^{T}\hat{e}_{tk}^2$.

Consider now the case when the errors are $I(1)$. For Model II, the form of the statistic is the same as in the stationary case, except for the construction of the autoregressive spectral density estimate. With $\alpha = 1$, 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

$$\hat{v}_t = \sum_{i=1}^{k} \zeta_i \hat{v}_{t-i} + e_{tk} \tag{13}$$

Denote the estimate by $\hat{\zeta}(L) = (1 - \hat{\zeta}_1 L \cdots - \hat{\zeta}_k L^k)$ and $\hat{\sigma}^2_{ek} = (T-k)^{-1} \sum_{t=k+1}^{T} \hat{e}^2_{tk}$. Then, $\hat{h}_v = \hat{\sigma}^2_{ek} / \hat{\zeta}(1)^2$.

With $I(1)$ errors and for Models I and III when a change in intercept is involved, things are somewhat more complex. Consider Model III, the limit distribution of the statistic $W_{RQF}(\lambda_1)$ defined by (11) is

$$\lim_{T \to \infty} v^2_{[\lambda_1 T]+1} / h_v + [\lambda_1 W(1) - W(\lambda_1)]^2 / [\lambda_1 (1 - \lambda_1)]$$

where $v_t = \Delta u_t$ and $h_v = \sigma^2_e / \zeta(1)^2$. Things are fine for the second component, but the first component involves two complications. The first is that h_v is not the proper scaling, instead one needs $\sigma^2_v = var(v_t)$ to have $v^2_{[\lambda_1 T]+1} / \sigma^2_v$ be $\chi^2(1)$. The other complication is that $v^2_{[\lambda_1 T]}$ is not *i.i.d.* as in the case of the AR(1) specification. This matters because the limit distribution of the test with an unknown break date will be different from that tabulated earlier.

Now recall that $\widetilde{\Psi} = (\widetilde{\mu}_0, \widetilde{\mu}_1, \widetilde{\beta}_0)'$ for Model I and $\widetilde{\Psi} = (\widetilde{\mu}_0, \widetilde{\mu}_1, \widetilde{\beta}_0, \widetilde{\beta}_1)'$ for Model III. To achieve the desired corrections one needs to replace $\widetilde{\mu}_1$ by μ^*_1 which involves two modifications. The first is to get back $e^2_{[\lambda_1 T]+1}$ instead of $v^2_{[\lambda_1 T]+1}$ in the limit distribution. Denote the sequence of estimates $\widetilde{\mu}_1$ for different values of the break date T_1 by $\widetilde{\mu}_1(T_1)$ for $T_1 = [\epsilon T], ..., [(1 - \epsilon) T]$. In large samples and under the null hypothesis of no break, $\widetilde{\mu}_1(T_1) \approx v_{[\lambda_1 T]+1}$ and, hence,

$$\hat{\zeta}(L) \widetilde{\mu}_1(T_1) = \widetilde{\mu}_1(T_1) - \hat{\zeta}_1 \widetilde{\mu}_1(T_1 - 1) \cdots - \hat{\zeta}_k \widetilde{\mu}_1(T_1 - k) \approx e_{[\lambda_1 T]+1}$$

in large samples. The second modification is to ensure the proper scaling. To that effect we define $\mu^*_1 = \hat{h}^{1/2}_v \hat{\zeta}(L) \widetilde{\mu}_1(T_1) / \hat{\sigma}_{ek}$. Hence, the form of the Wald test used is

$$W^*_{RQF}(\lambda_1) = [R(\widetilde{\Psi}^* - \Psi)]' [\hat{h}_v R (X'X)^{-1} R']^{-1} [R(\widetilde{\Psi}^* - \Psi)]$$

where $\widetilde{\Psi}^* = (\widetilde{\mu}_0, \mu^*_1, \widetilde{\beta}_0)'$ for Model I and $\widetilde{\Psi}^* = (\widetilde{\mu}_0, \mu^*_1, \widetilde{\beta}_0, \widetilde{\beta}_1)'$ for Model III, and X is as defined in (11). Note that in the case of Model I with hypothesis testing pertaining only to the shift in intercept μ_1, the statistic reduces to, with $R = [0, 1, 0]$,

$$\begin{aligned} W^*_{RQF}(\lambda_1) &= (\mu_1^*)^2/[\hat{h}_v R(X'X)^{-1}R'] \\ &= (\hat{\zeta}(L)\widetilde{\mu}_1(T_1)/\hat{\sigma}^2_{ek})/[R(X'X)^{-1}R'] \end{aligned}$$

and the estimate $\hat{h}_v$ is not needed.

Again, the decision rule to select whether to use the form of the statistic corresponding to the $I(0)$ or $I(1)$ case depends on whether the truncated value $\widetilde{\alpha}_{MS}$ is 1 or not. Asymptotically, this results in a correct classification and a consistent testing procedure. Hence, the procedure we recommend is the following:

1. Detrend the data by OLS to obtain residuals $\hat{u}_t$;

2. Consider the autoregression (9) with k selected using an information criterion (we recommend the Bayesian Information Criterion, BIC, with k allowed to be in the range $[0, 12(T/100)^{1/4}]$, see Schwarz, 1978). The corresponding estimate is denoted $\widetilde{\alpha}$. If the order selected is $k = 0$, the procedure described in Section 2.5 applies, otherwise, the next steps are applied.

3. Construct the bias corrected version of $\widetilde{\alpha}$, $\widetilde{\alpha}_M$, as defined by (7) described in Section 2.5. Then apply the truncation

$$\widetilde{\alpha}_{MS} = \begin{cases} \widetilde{\alpha}_M & \text{if } |\widetilde{\alpha}_M - 1| > T^{-1/2} \\ 1 & \text{if } |\widetilde{\alpha}_M - 1| \leq T^{-1/2} \end{cases}$$

4. Apply the quasi GLS procedure with $\widetilde{\alpha}_{MS}$ to obtain the estimate of Ψ and construct the Wald-statistic W_{RQF} or W^*_{RQF} depending on the Model and value of $\widetilde{\alpha}_{MS}$, using one of the two versions of $\hat{h}_v$ suggested to construct the estimate of (2π times) the spectral density function at frequency zero of v_t.

5. With an unknown break, the test statistic needs to be evaluated for each break date candidate and the Exp functional defined by (6) is evaluated.

4.3 Finite Sample Simulations

We present results about the finite sample size and power of our test with an $AR(2)$ error component generated by:

$$u_t = \alpha u_{t-1} + \psi(u_{t-1} - u_{t-2}) + e_t \qquad (14)$$

where $e_t \sim i.i.d.\ N(0,1)$ and $u_0 = u_{-1} = 0$. The number of replications is 1000 and we assess the size properties of our test at the nominal 5% level for the following specifications: $\alpha = 1, 0.95, 0.90, 0.80$, $\psi = 0.0, 0.3, 0.5, 0.7$, and $T = 100$. We consider positive AR coefficients since this is the most relevant case in practice.

We consider three versions of our test that varies with the choice of $\hat{h}_v$. In the first case, the estimate (12) is used and is referred to as "NP" (for Non Parametric). The second is the autoregressive based estimates with k chosen by BIC. It is referred to as "AR" (for Autoregressive). The third is a mixture of NP and AR. We use NP for the case $|\widetilde{\alpha}_{MS}| < 1$ and AR for $\widetilde{\alpha}_{MS} = 1$. It is referred to "AN" (for Autoregressive and Non Parametric) [4].

4.3.1 Known Break Date

Table 3 presents the size of the tests. The results suggest that the "AR" version can have serious size distortions in some of the $I(0)$ cases considered and that the "NP" specifications in turn leads to size distortions in the $I(1)$ case. Overall, the mixed method "AN" is a good compromise and has acceptable size properties, though somewhat conservative when α is close to but not equal to 1.

We now consider the power of our test with a break occurring at mid-sample, i.e., $T_1 = [0.5T]$. We consider only the specification that uses the mixed method "AN" to estimate h_v and compare its properties with the $T^{-1}W_T$ and the PS_T tests of Vogelsang (2001), again using a 5% nominal size so that their properties pertain to the case when both tests are applied independently. Table 4 presents the power results for $\eta = 0, 0.1, 0.3, 0.5$. Our test again has good properties and dominates the others for all models and values of α and ψ.

4.3.2 Unknown Break Date

Table 5 shows the size properties of our test in the case of an unknown break date when the Exp functional is used. The results suggest again that the mixed method "AN" to estimate h_v is preferable. The test based on the "AR" specification is too liberal especially when $\alpha = 0.8, 0.9$, and $\psi > 0$, and the test based on the "NP" specification is too liberal when $\alpha = 1$ and $\psi > 0$.

Next, we consider the power of our test with the "AN" specification. We only consider the Exp versions of the $T^{-1}W_T$ and the PS_T tests of Vogelsang (2001) as they work best for

[4] In the case of Model I, when $\widetilde{\alpha}_{MS} = 1$, we do not have to estimate the long-run variance. Therefore, AN is the same as NP.

all models (compared to the Mean and Sup functionals). Table 6 presents the power results for $\eta = 0, 0.1, 0.3, 0.5$. Our test again dominates the others in most cases. For Models I and III, our test is considerably superior and for Model II, it is competitive to the best of the $T^{-1}W_T$ and PS_T tests.

5 The Multiple Breaks Case

Our testing procedure extends, in principle, naturally to the case of multiple breaks. This is important since, as discussed in Perron (2005), most tests may exhibit non monotonic power functions if the number of breaks present under the alternative is greater than the number of breaks explicitly accounted for in the construction of the tests. Consider the following extended versions of our three Models for m breaks denoted $(T_1, ..., T_m)$ with corresponding break fractions $\lambda_i = T_i/T$ $(i = 1, ..., m)$.

Model I (Multiple structural breaks in intercepts): $x_t = (1, DU_{1t}, ..., DU_{mt}, t)'$, $\Psi = (\mu_0, \mu_1, ..., \mu_m, \beta_0)'$ where $DU_{it} = 1(t > T_i)$. Here, the hypothesis of interest is $\mu_1 = ... = \mu_m = 0$.

Model II (Multiple structural breaks in slopes): $x_t = (1, t, DT_{1t}, ..., DT_{mt})'$, $\Psi = (\mu_0, \beta_0, \beta_1, ..., \beta_m)'$ where $DT_{it} = 1(t > T_i)(t - T_i)$. The hypothesis of interest is $\beta_1 = ... = \beta_m = 0$.

Model III (Multiple structural breaks both in intercepts and slopes): $x_t = (1, DU_{1t}, ..., DU_{mt}, t, DT_{1t}, ..., DT_{mt})'$, $\Psi = (\mu_0, \mu_1, ..., \mu_m, \beta_0, \beta_1, ..., \beta_m)'$. The hypothesis of interest is $\mu_1 = ... = \mu_m = \beta_1 = ... = \beta_m = 0$.

All theoretical results discussed for the case of a single break continue to hold with minor modifications, as stated in the following Theorem.

Theorem 3 *Consider first the case of known break dates. Let $W_{FS}(\lambda)$ denote the Wald test for testing null hypothesis relevant to Models I, II or III. Under the data generating process (1), when $|\alpha| < 1$, the results of Theorem 1 continue to hold with $F(r, \lambda_1)$ replaced by the following: for Model I, $F(r, \lambda_1, ..., \lambda_m) = [1, 1(r > \lambda_1), ..., 1(r > \lambda_m), r]'$; for Model II, $F(r, \lambda_1, ..., \lambda_m) = [1, r, 1(r > \lambda_1)(r - \lambda_1), ..., 1(r > \lambda_m)(r - \lambda_m)]'$; and for Model III, $F(r, \lambda_1, ..., \lambda_m) = [1, 1(r > \lambda_1), ..., 1(r > \lambda_m), r, 1(r > \lambda_1)(r - \lambda_1), ..., 1(r > \lambda_m)(r - \lambda_m)]'$. If*

$\alpha = 1$,

$$W_{FS}(\lambda) \Rightarrow \begin{cases} V_1(\lambda) & \textit{for Model I} \\ V_2(\lambda) & \textit{for Model II} \\ V_1(\lambda) + V_2(\lambda) & \textit{for Model III} \end{cases}$$

where

$$\begin{aligned} V_1(\lambda) &= \lim_{T\to\infty} (e^2_{[\lambda_1 T]+1} + ... + e^2_{[\lambda_m T]+1})/\sigma^2 \\ V_2(\lambda) &= [R^* H(\lambda)^{-1} Q(\lambda)]' [R^* H(\lambda)^{-1} R^{*\prime}]^{-1} [R^* H(\lambda)^{-1} Q(\lambda)] \end{aligned}$$

with $R^* = \left[\underset{\sim}{0},\ I_{(m)}\right]$ a $(m \times m+1)$ matrix where $I_{(m)}$ is the identitiy matrix of order m and $\underset{\sim}{0}$ is a $(m \times 1)$ zero vector; $Q(\lambda) = [W(1), W(1) - W(\lambda_1), ..., W(1) - W(\lambda_m)]'$ is a $(m+1 \times 1)$ vector; $H(\lambda)$ is a $(m+1 \times m+1)$ matrix defined by

$$H(\lambda) = \begin{bmatrix} 1 & 1-\lambda_1 & ... & 1-\lambda_m \\ 1-\lambda_1 & 1-\lambda_1 & ... & 1-\lambda_m \\ \vdots & \vdots & \ddots & \vdots \\ 1-\lambda_m & 1-\lambda_m & ... & 1-\lambda_m \end{bmatrix}.$$

Also, the Exp-test defined by

$$Exp\text{-}W_{FS} = \log \left[T^{-m} \sum_{\Lambda} \exp\left(\frac{1}{2} W_{FS}(\lambda')\right)\right]$$

where $\Lambda = \{(\lambda'_1, ..., \lambda'_m);\ \lambda'_1 \geq \epsilon,\ 1-\epsilon \geq \lambda'_m,\ \lambda'_i - \lambda'_{i-1} \geq \epsilon \text{ for } i = 2, ..., m\}$ has a limit distribution given by $Exp\text{-}W_{FS} \Rightarrow \log\left[\int_\Lambda \exp\left(\frac{1}{2} g(\lambda')\right) d\lambda'\right]$, where $g(\lambda')$ is the relevant limit function of the Wald test.

Remark 5 *$Q(\lambda) \sim N(0, H(\lambda))$. Therefore, $R^* H(\lambda)^{-1} Q(\lambda) \sim N(0, R^* H(\lambda)^{-1} R^{*\prime})$ and $V_2(\lambda)$ is a chi-square random variable with m degrees of freedom, which is independent of $V_1(\lambda)$. It is also easy to see that when $m = 1$, we recover the results of Theorem 1.*

While the theoretical extensions to the case with multiple breaks are straightforward, there remains an important problem for practical implementations. The reason is that the *Exp*-W_{FS} (and the Mean and Sup versions) depends on Wald tests evaluated at a number of partitions (or combinations of break dates) that is of order $O(T^m)$. This becomes prohibitive for common sample sizes once m is greater than 2. For the Sup-Wald test, an efficient solution to find the partition that corresponds to the maximal value of the Wald test has been devised based on a dynamic programming algorithm (see Bai and Perron, 2003). However, no such efficient algorithm exists to compute the Mean and Exp-Wald tests in the case of multiple breaks. So a full treatment will need to wait for advances in this respect. Nevertheless, the case with two breaks is computationally feasible and also important in practice. For example, Lumsdaine and Papell (1997) showed evidence that the Nelson and Plosser (1982) data may have two breaks in a trend function. To that effect, Table 7 presents the relevant quantiles of the limiting distributions of the $Exp\text{-}W_{FS}$ test statistic for both the $I(0)$ and $I(1)$ cases. Again, one should use the largest of the two when performing hypothesis testing. Note that, here also, the test is only slightly conservative (asymptotically) for the cases that do not correspond to that from which the critical values are selected.

6 Empirical Applications

This section considers empirical applications related to real GDP series for several countries. As discussed in the introduction, assessing the stability of the trend function of series of aggregate economic activity is an important practical question. We present evidence using different sets of series. The first relates to historical series for a variety of countries spanning the period 1870 to 1986 and considers both real GDP series and their per capita counterparts. The second set considers postwar real GDP series for the G7 countries.

6.1 Historical real GDP series

We consider a historical data set of (log) real GDP series and their per capita counterparts from 1870 to 1986 for 10 different countries: Australia, Canada, Denmark, France, Germany, Italy, Norway, Sweden, the United Kingdom and the United States [5]. The series are presented in Figures 8.a and 8.b. From these figures, it seems clear that most series are characterized by at least one (and in most cases only one) major shift in the slope and/or intercept of the

[5] This data set is the same as used by Kormendi and Meguire (1990), Perron (1992) and Perron and Zhu (2005), and was obtained through the *Journal of Money, Credit and Banking* editorial office. All series are real GDP except for the United States for which real GNP is used. For the United States, the series is real GNP from the National Income and Products Accounts for the period 1929-1986, spliced to Romer's (1989) estimates for the period 1870-1928. For the United Kingdom, the series is real GDP from Feinstein (1972) for the period 1870-1947 spliced to the International Financial Statistics (IFS) series of the IMF for the period 1948-1986. For the remaining countries, the series are indices of annual real GDP from Madison (1982) spliced to the postwar IFS data. The population series used are from the same sources. All series are analyzed with a logarithmic transformation.

trend function. The dates of the breaks are not common across countries. They occur at the time of World War II for France, Germany and Italy; at the beginning of the 30's for Australia and around the time of World War I for Sweden and the United Kingdom. While these visual inspections are revealing, our test can give a more precise statement about whether a break in the trend function exists or not, without any prior knowledge about the $I(0)$ or $I(1)$ nature of the noise component.

To carry the testing procedure we used Model III, given the nature of the series. The results are presented in Table 8. They clearly point to a strong rejection of the null hypothesis that the trend function is stable in favor of a trend function with a shift for all countries, with the exception of Canada. This conclusion remains whether we use real GDP series or their per capita counterparts. Table 8 also presents an estimate of the break date obtained by minimizing the sum of squared residuals from a regression of the relevant series on a constant, a time trend, a level shift dummy and a slope shift dummy. As shown in Perron and Zhu (2005), selecting the break date in this fashion leads to a consistent estimate whether the errors are $I(0)$ or $I(1)$. In general, the break date selected is the same for the real GDP series and its per capita counterpart (the exceptions are Norway and the U.S., and to a some extent Germany). Also, these dates correspond to plausible events, most notably WWII for the European Countries and WW I for the U.K.. Also, presented are the pre- and post-break annual rates of growth. The differences are in many cases very large and, with the exception of the U.S., the post break rates of growth are larger.

A related paper is that of Ben-David and Papell (1995). They analyzed real GDP and per capita real GDP for 16 industrial countries over the period 1860-1989 (including many of the countries considered here) and tested for the presence of a structural break using the tests of Vogelsang (1997). Comparing their results to ours, we see stronger rejections for more countries using our test.

6.2 Postwar Real GDP for the G7 Countries

We now consider quarterly (log) real GDP series for the G7 countries. For all countries except the U.S., the data are from the International Financial Statistics (IFS) database. The series start at different dates for each countries but all end in 2002:4 (see Table 9). For the U.S., the data is from the Citibase databank and the sample period is 1947:1-1998:2 (it is the same series used in Perron and Wada (2005) who consider the related issue of trend-cycle decompositions). Graphs of the series are presented in Figure 9. They show that all countries experienced a decline in the rate of growth, the change occurring mostly in the seventies. To document whether such changes are significant, we report the results of our test statistic using Model III. While some series do not show discontinuity in the form of a level shift, using Model III may be more appropriate to crudely account for the gradual nature of the change in slope. In any event, the conclusions are unchanged using Model II in most cases, the exception being the U.S. which we discuss in more detail below.

The results are presented in Table 9. For Canada, France, Germany, Italy and Japan, the evidence in favor of a structural change in the trend function is very strong with test statistics having p-values well below 1%. For the U.K., a rejection is possible only at the 10% level. Table 9 also shows the estimates of the break dates (obtained as discussed above) and the pre- and post-break annual rates of growth. For these six series, the break dates are close to the oil price shock of 1973 and vary between 1972:1 (Japan) to 1978:1 (Italy). The changes in the rates of growth are also very large. The smallest change is for the U.K, from 3.2% to 2.4%, which may explain the marginal rejection as being due to a potentially low power of the test in this case. The largest change is for Japan which went from a 9.8% to a 2.6% growth rate.

We consider now the results for the U.S.. Here things are quite different when the statistics are based on Model II or III. With Model III, a rejection at the 5% level is possible but the break date is estimated to be at 1965:4 with the growth rate going from 3.5% to 2.8%. With Model II, the break date occurs at 1973:2 and the growth rate goes from 3.8% to 2.7%, a result in line with those of Perron and Wada (2005), but the statistic does not allow a rejection at any significance level. Unreported estimations point to the fact that the results are also quite sensitive to the exact sample period used. Perron and Wada's (2005) results are more in line with those of Model II. They document the fact that the change occurs over several quarters around 1973:1, though the bulk of it is within a one year period. Also, their results imply that the noise component is stationary, the sum of the autoregressive coefficients being close to 0.9, a region of the parameter space where our test, or any other, has low power.

7 Conclusion

We proposed new test statistics for structural breaks in a trend function when it is a priori unknown whether the series is trend-stationary or contains an autoregressive unit root. The test statistics are based on a Feasible Quasi Generalized Least Squares procedure with a superefficient estimate of the sum of the autoregressive parameters α when $\alpha = 1$. With known break dates, the Wald test is asymptotically distributed as a chi-square random variable for any values of α. On the other hand, with unknown break dates, the limiting distributions of test statistics still depend on the I(0) or I(1) dichotomy. Nevertheless, for the Exp version the asymptotic critical values are very close for all significance levels, thereby allowing a procedure with nearly the same asymptotic size in both the $I(0)$ and $I(1)$ cases. Simulations have shown its usefulness and that it provides substantial improvements over existing tests. Our empirical applications reveal that structural changes in the trend function of real GDP series (or their per capita counterparts) are widespread across historical episodes and countries.

Appendix: Technical Derivations

Proof of equation (3): The t-statistic for β_1 is given by:

$$
\begin{aligned}
t_F &= \frac{T^{-1/2}\sum_{t=1}^{T}(DT_t-\hat{\alpha}DT_{t-1})(u_t-\hat{\alpha}u_{t-1})}{(s^2T^{-1}\sum_{t=1}^{T}(DT_t-\hat{\alpha}DT_{t-1})^2)^{1/2}} \\
&= \Bigg\{T^{-1/2}\sum_{t=T_1+1}^{T} e_t - T(\hat{\alpha}-1)T^{-3/2}\sum_{t=T_1+1}^{T} u_{t-1} \\
&\quad -T(\hat{\alpha}-1)[T^{-3/2}\sum_{t=T_1+1}^{T}(t-T_1)e_t - T(\hat{\alpha}-1)T^{-5/2}\sum_{t=T_1+1}^{T}(t-T_1)u_{t-1}]\Bigg\} \\
&\quad / \left\{s^2[(1-\lambda_1)-T(\hat{\alpha}-1)(1-\lambda_1)^2+T^2(\hat{\alpha}-1)^2(1-\lambda_1)^3/3]\right\}^{1/2} + o_p(1).
\end{aligned}
$$

The result follows using the facts that $T^{-1/2}\sum_{t=1}^{[rT]} e_t \Rightarrow \sigma W(r)$, $T^{-3/2}\sum_{t=T_1+1}^{T}(t-T_1)e_t \Rightarrow \sigma\int_{\lambda_1}^{1}(r-\lambda_1)dW(r)$, $T^{-3/2}\sum_{t=T_1+1}^{T}u_t \Rightarrow \sigma\int_{\lambda_1}^{1}W(r)dr$, $T^{-5/2}\sum_{t=T_1+1}^{T}(t-T_1)u_t \Rightarrow \sigma\int_{\lambda_1}^{1}(r-\lambda_1)W(r)dr$, and $s^2=\sigma^2+o_p(1)$.

Proof of Theorem 1: We give the proof only for Model II, the derivations being similar for all Models. We have

$$
\hat{\Psi}-\Psi = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{12} & q_{22} & q_{23} \\ q_{13} & q_{23} & q_{33} \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix}
$$

where

$$
q_{11} = 1+(T-1)(\hat{\alpha}_S-1)^2
$$

$$
q_{12} = 1+(1-\hat{\alpha}_S)\sum_{t=2}^{T}(t-\hat{\alpha}_S(t-1)) = 1-(T-1)\hat{\alpha}_S(\hat{\alpha}_S-1)+(\hat{\alpha}_S-1)^2\sum_{t=2}^{T}t,
$$

$$
\begin{aligned}
q_{13} &= (1-\hat{\alpha}_S)\sum_{t=1}^{T}(DT_t-\hat{\alpha}_S DT_{t-1}) = -(\hat{\alpha}_S-1)\sum_{t=1}^{T-T_1}(\hat{\alpha}_S-(\hat{\alpha}_S-1)t) \\
&= -(T-T_1)\hat{\alpha}_S(\hat{\alpha}_S-1)+(\hat{\alpha}_S-1)^2\sum_{t=1}^{T-T_1}t,
\end{aligned}
$$

$$
q_{22} = \sum_{t=1}^{T}(\hat{\alpha}_S-(\hat{\alpha}_S-1)t)^2 = T\hat{\alpha}_S^2-2\hat{\alpha}_S(\hat{\alpha}_S-1)\sum_{t=1}^{T}t+(\hat{\alpha}_S-1)^2\sum_{t=1}^{T}t^2,
$$

$$q_{33} = \sum_{t=1}^{T_1}(\hat{\alpha}_S - (\hat{\alpha}_S - 1)t)^2 = (T - T_1)\hat{\alpha}_S^2 - 2\hat{\alpha}_S(\hat{\alpha}_S - 1)\sum_{t=1}^{T-T_1} t + (\hat{\alpha}_S - 1)^2 \sum_{t=1}^{T-T_1} t^2,$$

$$\begin{aligned} q_{23} &= \sum_{t=2}^{T}(\hat{\alpha}_S - (\hat{\alpha}_S - 1)t)(DT_t - \hat{\alpha}_S DT_{t-1}) \\ &= \sum_{t=1}^{T-T_1}(\hat{\alpha}_S - (\hat{\alpha}_S - 1)T_1 - (\hat{\alpha}_S - 1)t)(\hat{\alpha}_S - (\hat{\alpha}_S - 1)t) = T_1 q_{13} + q_{33} \end{aligned}$$

$$r_1 = u_1 - (\hat{\alpha}_S - 1)\sum_{t=2}^{T} u_t^*,$$

$$r_2 = u_1 + \sum_{t=2}^{T} u_t^*(\hat{\alpha}_S - (\hat{\alpha}_S - 1)t) = u_1 + \hat{\alpha}_S \sum_{t=2}^{T} u_t^* - (\hat{\alpha}_S - 1)\sum_{t=2}^{T} t u_t^*,$$

$$r_3 = \sum_{t=1}^{T-T_1} u_{T_1+t}^*(\hat{\alpha}_S - (\hat{\alpha}_S - 1)t) = \hat{\alpha}_S \sum_{t=1}^{T-T_1} u_{T_1+t}^* - (\hat{\alpha}_S - 1)\sum_{t=1}^{T-T_1} t u_{T_1+t}^*$$

where $u_t^* = u_t - \hat{\alpha}_S u_{t-1}$.

Stationary Case ($|\alpha| < 1$). The convergence results for each components are as follows:

$$\begin{aligned} T^{-1/2}\sum_{t=1}^{[Tr]} u_t^* &= T^{-1/2}\sum_{1}^{[Tr]}(e_t - (\hat{\alpha}_S - \alpha)u_{t-1}) \\ &= T^{-1/2}\sum_{t=1}^{[Tr]} e_t - T^{-1/2}(T^{1/2}(\hat{\alpha}_S - \alpha))(T^{-1/2}\sum_{t=1}^{[Tr]} u_{t-1}) \\ &= T^{-1/2}\sum_{t=1}^{[Tr]} e_t + o_p(1) \Rightarrow \sigma W(r), \end{aligned}$$

$T^{-1}q_{11} \Rightarrow (1-\alpha)^2$, $T^{-2}q_{12} \Rightarrow (1-\alpha)^2 \int_0^1 r dr$, $T^{-2}q_{13} \Rightarrow (1-\alpha)^2 \int_0^1 (r-\lambda_1)1(r>\lambda_1)dr$, $T^{-3}q_{22} \Rightarrow (1-\alpha)^2 \int_0^1 r^2 dr$, $T^{-3}q_{33} \Rightarrow (1-\alpha)^2 \int_0^1 (r-\lambda_1)^2 1(r>\lambda_1)dr$, $T^{-3}q_{23} \Rightarrow (1-\alpha)^2 \int_0^1 r(r-\lambda_1)1(r>\lambda_1)dr$, $T^{-1/2}r_1 \Rightarrow \sigma(1-\alpha)W(1)$, $T^{-3/2}r_2 \Rightarrow \sigma(1-\alpha)\int_0^1 r dW(r)$, $T^{-3/2}r_3 \Rightarrow \sigma(1-\alpha)\int_0^1 (r-\lambda_1)1(r>\lambda_1)dW(r)$. Let $\Upsilon_T = diag(T^{1/2}, T^{3/2}, T^{3/2})$. Then, we have

$$\Upsilon_T^{-1}(X'X)\Upsilon_T^{-1} = \begin{bmatrix} T^{-1}q_{11} & T^{-2}q_{12} & T^{-2}q_{13} \\ & T^{-3}q_{22} & T^{-3}q_{23} \\ & & T^{-3}q_{33} \end{bmatrix}$$

$$\Rightarrow (1-\alpha)^2 \begin{bmatrix} 1 & \int_0^1 rdr & \int_0^1 (r-\lambda_1)1(r>\lambda_1)dr \\ & \int_0^1 r^2 dr & \int_0^1 r(r-\lambda_1)1(r>\lambda_1)dr \\ & & \int_0^1 (r-\lambda_1)^2 1(r>\lambda_1)dr \end{bmatrix}$$

$$= (1-\alpha)^2 \int_0^1 F(r)F(r)'dr$$

$$\Upsilon_T^{-1}X'U = \begin{bmatrix} T^{-1/2}r_1 \\ T^{-3/2}r_2 \\ T^{-3/2}r_3 \end{bmatrix} \Rightarrow \sigma(1-\alpha)\begin{bmatrix} \int_0^1 dW(r) \\ \int_0^1 rdW(r) \\ \int_0^1 (r-\lambda_1)1(r>\lambda_1)dW(r) \end{bmatrix}$$

$$= \sigma(1-\alpha)\int_0^1 F(r)dW(r)$$

$$\Upsilon_T(\hat{\Psi}-\Psi) = (\Upsilon_T^{-1}X'X\Upsilon_T^{-1})^{-1}(\Upsilon_T^{-1}X'U)$$

$$\Rightarrow ((1-\alpha)^2\int_0^1 F(r)F(r)'dr)^{-1}(\sigma(1-\alpha)\int_0^1 F(r)dW(r))$$

$$= \frac{\sigma}{1-\alpha}(\int_0^1 F(r)F(r)'dr)^{-1}\int_0^1 F(r)dW(r)$$

The result stated in Theorem 1 follows using the convergence results stated above noting that we can express the Wald tests as:

$$W_{FS}(\lambda_1) = [R\Upsilon_T(\hat{\Psi}-\Psi)]'[s^2 R\Upsilon_T(X'X)^{-1}\Upsilon_T R']^{-1}[R\Upsilon_T(\hat{\Psi}-\Psi)] \quad \text{(A.1)}$$

Unit Root Case ($\alpha = 1$). Using the fact that $T(\hat{\alpha}_S - 1) \to^p 0$, the convergence results for each elements are:

$$T^{-1/2}\sum_{t=1}^{[Tr]} u_t^* = T^{-1/2}\sum_{t=1}^{[Tr]}(e_t - (\hat{\alpha}_S - 1)u_{t-1})$$

$$= T^{-1/2}\sum_{t=1}^{[Tr]} e_t - T(\hat{\alpha}_S - 1)(T^{-1}\sum_{t=1}^{[Tr]} T^{-1/2}u_{t-1}) \Rightarrow \sigma W(r),$$

$q_{11} \Rightarrow 1$, $q_{12} \Rightarrow 1$, $q_{13} \Rightarrow 0$, $T^{-1}q_{22} \Rightarrow 1$, $T^{-1}q_{33} \Rightarrow (1-\lambda_1)$, $T^{-1}q_{23} \Rightarrow (1-\lambda_1)$, $r_1 = u_1 + o_p(1)$,$T^{-1/2}r_2 \Rightarrow \sigma W(1)$,$T^{-1/2}r_3 \Rightarrow \sigma \int_{\lambda_1}^1 dW(r)$. Let $\Upsilon_T = diag(1, T^{1/2}, T^{1/2})$. Then we have

$$\begin{aligned}
\Upsilon_T(\hat{\Psi}(\lambda) - \Psi) &= (\Upsilon_T^{-1}X'X\Upsilon_T^{-1})^{-1}(\Upsilon_T^{-1}X'U) \\
&= \begin{bmatrix} q_{11} & T^{-1/2}q_{12} & T^{-1/2}q_{13} \\ & T^{-1}q_{22} & T^{-1}q_{23} \\ & & T^{-1}q_{33} \end{bmatrix}^{-1} \begin{bmatrix} r_1 \\ T^{-1/2}r_2 \\ T^{-1/2}r_3 \end{bmatrix} \\
&\Rightarrow \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 1-\lambda_1 \\ & & 1-\lambda_1 \end{bmatrix}^{-1} \begin{bmatrix} e_1 \\ \sigma W(1) \\ \sigma(W(1) - W(\lambda_1)) \end{bmatrix} \\
&= \begin{bmatrix} e_1 \\ \sigma \frac{W(\lambda_1)}{\lambda_1} \\ \sigma \frac{\lambda_1 W(1) - W(\lambda_1)}{\lambda_1(1-\lambda_1)} \end{bmatrix}
\end{aligned}$$

The result stated in Theorem 1 follows using the convergence results stated above and the representation of the Wald test stated in (A.1).

Near Unit Root Case ($\alpha_T = 1 + c/T$, Proof of Theorem 2). As shown in Perron and Yabu (2005), $T(\hat{\alpha}_S - 1) \to_p 0$. Now, the true value of α is in a T^{-1} neighborhood of 1 so that in large sample $\hat{\alpha}$ is always truncated to take value one. Then, we have the following limit results:

$$\begin{aligned}
T^{-1/2}\sum_{t=1}^{[Tr]} u_t^* &= T^{-1/2}\sum_{t=1}^{[Tr]}(e_t + \frac{c}{T}u_{t-1} - (\hat{\alpha}_S - 1)u_{t-1}) \\
&= T^{-1/2}\sum_{t=1}^{[Tr]} e_t + cT^{-3/2}\sum_{t=1}^{[Tr]} u_{t-1} - T(\hat{\alpha}_S - 1)(T^{-1}\sum_{t=1}^{[Tr]} T^{-1/2}u_{t-1}) \\
&\Rightarrow \sigma[W(r) + c\int_0^r J_c(s)ds] = \sigma J_c(r),
\end{aligned}$$

$q_{11} \Rightarrow 1$, $q_{12} \Rightarrow 1$, $q_{13} \Rightarrow 0$, $T^{-1}q_{22} \Rightarrow 1$, $T^{-1}q_{33} \Rightarrow (1-\lambda_1)$, $T^{-1}q_{23} \Rightarrow (1-\lambda_1)$, $r_1 = u_1 + o_p(1)$,$T^{-1/2}r_2 \Rightarrow \sigma J_c(1)$,$T^{-1/2}r_3 \Rightarrow \sigma \int_{\lambda_1}^1 dJ_c(r)$. Let $\Upsilon_T = diag(1, T^{1/2}, T^{1/2})$. The result stated in Theorem 2 follows using the convergence results stated above and the representation of the Wald test stated in (A.1).

References

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

[2] Andrews, D.W.K., and W. Ploberger (1994): "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica* 62, 1383-1414.

[3] Bai, J., and P. Perron (2003): "Computation and Analysis of Multiple Structural Change Models," *Journal of Applied Econometrics* 18, 1-22.

[4] Ben-David, D., and D.H. Papell (1995): "The Great Wars, the Great Crash, and Steady State Growth: Some New Evidence about an Old Stylized Fact," *Journal of Monetary Economics* 36, 453-475.

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

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

[7] Feinstein, C.H. (1972): "National Income Expenditure and Output of the United Kingdom," vol. 6 in *Studies in the National Income and Expenditure of the United Kingdom*, Richard Stone (series editor), London: Cambridge University Press.

[8] Gardner, L.A. (1969): "On Detecting Changes in the Mean of Normal Variates," *The Annals of Mathematical Statistics* 40, 116-126.

[9] Kim, D., and P. Perron (2005): "Unit Root Tests with a Consistent Break Fraction Estimator," Unpublished Manuscript, Department of Economics, Boston University.

[10] Kormendi, R.C., and P. Meguire (1990): "A Multicountry Characterization of the Nonstationarity of Aggregate Output," *Journal of Money, Credit and Banking* 22, 77-93.

[11] Lumsdaine, R.L, and D.H. Papell (1997): "Multiple Trend Breaks and the Unit-Root Hypothesis," *Review of Economics and Statistics* 79, 212-218.

[12] Maddison, A. (1982): *Phases of Capitalist Development*, London: Oxford University Press.

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

[14] Perron, P. (1991): "A Test for Changes in a Polynominal Trend Function for a Dynamic Time Series," Research Memorandum No. 363, Econometric Research Program, Princeton University.

[15] Perron, P. (1992): "Trend, Unit Root and Structural Change: A Multi-Country Study with Historical Data," *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 144-149.

[16] Perron, P. (2005): "Dealing with Structural Breaks," forthcoming in the *Palgrave Handbook of Econometrics*, Volume 1: Econometric Theory.

[17] Perron, P., and T. Wada (2005): "Trends and Cycles: A New Approach and Explanations of some Old Puzzles," Unpublished Manuscript, Department of Economics, Boston University.

[18] Perron, P., and T. Yabu (2005): "Estimating Deterministic Trends with an Integrated or Stationary Noise Component," Unpublished Manuscript, Department of Economics, Boston University.

[19] Perron, P., and X. Zhu (2005): "Structural Breaks with Deterministic and Stochastic Trends," forthcoming in the *Journal of Econometrics*.

[20] Romer, C. (1989): "The Prewar Business Cycle Reconsidered: New Estimates of Gross National Product, 1969-1928," *Journal of Political Economy* 97, 1-37.

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

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

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

[24] Vogelsang, T.J. (1997): "Wald-Type Tests for Detecting Breaks in the Trend Function of a Dynamic Time Series," *Econometric Theory* 13, 818-849.

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

[26] Vogelsang, T.J. (2001): "Testing for a Shift in Ttrend When Serial Correlation is of Unknown Form," Unpublished Manuscript, Department of Economics, Cornell University.

[27] Vogelsang, T.J., and P. Perron (1998): "Additional Tests for a Unit Root Allowing the Possibility of Breaks in the Trend Function," *International Economic Review* 39, 1073-1100.

[28] Yabu, T. (2005): "Essays on Theoretical and Empirical Aspects of Structural Change Models," Unpublished PhD Dissertation, Department of Economics, Boston University.

[29] Zivot, E., and D.W.K. Andrews (1992): "Further Evidence on the Great Crash, the Oil Price Shock and the Unit Root Hypothesis," *Journal of Business and Economic Statistics* 10, 251-270.

Table 1: Asymptotic Distributions for One Break Occurring at an Unknown Date

	$\epsilon = 0.01$						$\epsilon = 0.15$					
	sup-W_{FS}		Mean-W_{FS}		Exp-W_{FS}		sup-W_{FS}		Mean-W_{FS}		Exp-W_{FS}	
%	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
Model I												
0.900	9.99	∞	1.72	0.98	1.59	1.60	8.96	∞	1.31	0.70	1.22	1.26
0.950	11.64	∞	2.11	0.98	2.07	1.92	10.60	∞	1.64	0.70	1.74	1.58
0.975	13.08	∞	2.50	0.98	2.60	2.32	12.17	∞	1.98	0.70	2.30	1.99
0.990	15.03	∞	3.04	0.98	3.33	2.99	14.30	∞	2.47	0.70	3.12	2.64
Model II												
0.900	6.38	8.86	2.12	1.93	1.45	1.52	4.91	7.14	1.66	1.50	1.07	1.13
0.950	7.78	10.42	2.84	2.49	1.97	2.02	6.27	8.68	2.28	2.01	1.61	1.67
0.975	9.19	11.94	3.59	3.06	2.55	2.57	7.63	10.24	2.92	2.54	2.18	2.26
0.990	11.00	13.97	4.54	3.90	3.30	3.37	9.33	12.18	3.82	3.19	2.97	3.06
Model III												
0.900	12.79	∞	3.36	2.91	2.68	2.96	11.11	∞	2.58	2.20	2.25	2.48
0.950	14.52	∞	4.16	3.47	3.34	3.55	12.86	∞	3.17	2.71	2.84	3.12
0.975	16.13	∞	4.96	4.04	3.94	4.15	14.38	∞	3.83	3.24	3.52	3.75
0.990	18.13	∞	6.11	4.88	4.67	5.02	16.66	∞	4.70	3.89	4.35	4.47

Table 2.a: Asymptotic Distribution of the Exp Test: Model I

	$\epsilon = 0.01$		$\epsilon = 0.05$		$\epsilon = 0.10$		$\epsilon = 0.15$		$\epsilon = 0.25$	
%	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0.900	1.59	1.60	1.47	1.52	1.33	1.41	1.22	1.26	0.83	0.91
0.905	1.60	1.61	1.51	1.55	1.36	1.43	1.26	1.28	0.86	0.93
0.910	1.62	1.62	1.56	1.57	1.40	1.45	1.30	1.30	0.90	0.96
0.915	1.66	1.65	1.59	1.59	1.44	1.46	1.35	1.34	0.93	0.98
0.920	1.71	1.67	1.63	1.62	1.50	1.49	1.40	1.36	0.98	1.01
0.925	1.76	1.70	1.67	1.65	1.55	1.52	1.46	1.40	1.03	1.04
0.930	1.81	1.74	1.71	1.69	1.59	1.55	1.50	1.43	1.09	1.08
0.935	1.85	1.78	1.75	1.72	1.66	1.58	1.56	1.47	1.15	1.11
0.940	1.92	1.81	1.82	1.76	1.72	1.62	1.61	1.51	1.20	1.16
0.945	1.98	1.86	1.89	1.82	1.80	1.66	1.69	1.56	1.28	1.21
0.950	2.07	1.92	1.97	1.86	1.88	1.70	1.74	1.58	1.33	1.26
0.955	2.14	1.97	2.05	1.92	1.96	1.76	1.85	1.67	1.41	1.33
0.960	2.22	2.04	2.16	1.99	2.05	1.82	1.93	1.73	1.51	1.40
0.965	2.33	2.11	2.25	2.07	2.18	1.88	2.02	1.81	1.61	1.49
0.970	2.44	2.23	2.40	2.16	2.30	1.98	2.14	1.88	1.73	1.58
0.975	2.60	2.32	2.53	2.29	2.45	2.07	2.30	1.99	1.92	1.68
0.980	2.74	2.50	2.69	2.41	2.60	2.19	2.50	2.14	2.11	1.81
0.985	3.02	2.81	2.90	2.60	2.79	2.39	2.74	2.36	2.35	2.01
0.990	3.33	2.99	3.24	2.81	3.05	2.67	3.12	2.64	2.83	2.32
0.995	3.91	3.62	3.84	3.26	3.60	3.29	3.55	3.28	3.33	2.91

Table 2.b: Asymptotic Distribution of the Exp Test: Model II

	$\epsilon = 0.01$		$\epsilon = 0.05$		$\epsilon = 0.10$		$\epsilon = 0.15$		$\epsilon = 0.25$	
%	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0.900	1.45	1.52	1.34	1.40	1.20	1.28	1.07	1.13	0.71	0.74
0.905	1.47	1.58	1.38	1.45	1.24	1.33	1.11	1.19	0.74	0.79
0.910	1.51	1.62	1.42	1.49	1.29	1.36	1.15	1.23	0.79	0.84
0.915	1.55	1.65	1.47	1.53	1.32	1.41	1.19	1.27	0.83	0.89
0.920	1.60	1.70	1.53	1.57	1.37	1.46	1.24	1.32	0.87	0.93
0.925	1.64	1.74	1.58	1.62	1.43	1.51	1.30	1.37	0.92	0.98
0.930	1.70	1.80	1.64	1.68	1.47	1.57	1.34	1.44	0.97	1.02
0.935	1.76	1.86	1.70	1.73	1.54	1.62	1.39	1.49	1.03	1.08
0.940	1.84	1.92	1.76	1.78	1.61	1.69	1.47	1.53	1.11	1.14
0.945	1.89	1.98	1.84	1.86	1.68	1.78	1.54	1.59	1.19	1.22
0.950	1.97	2.02	1.90	1.93	1.75	1.86	1.61	1.67	1.25	1.28
0.955	2.06	2.10	1.98	2.01	1.85	1.93	1.71	1.73	1.32	1.37
0.960	2.19	2.19	2.07	2.10	1.94	2.05	1.80	1.80	1.40	1.46
0.965	2.30	2.29	2.19	2.23	2.06	2.14	1.91	1.91	1.52	1.59
0.970	2.41	2.46	2.29	2.36	2.19	2.24	2.06	2.10	1.65	1.73
0.975	2.55	2.57	2.42	2.51	2.34	2.39	2.18	2.26	1.81	1.84
0.980	2.74	2.82	2.61	2.71	2.56	2.55	2.42	2.39	1.97	2.02
0.985	3.00	3.09	2.83	2.98	2.79	2.81	2.64	2.61	2.20	2.24
0.990	3.30	3.37	3.07	3.27	3.20	3.18	2.97	3.06	2.60	2.61
0.995	3.88	4.01	3.61	3.95	3.80	3.68	3.62	3.46	3.28	3.27

Table 2.c: Asymptotic Distribution of the Exp Test: Model III

	$\epsilon = 0.01$		$\epsilon = 0.05$		$\epsilon = 0.10$		$\epsilon = 0.15$		$\epsilon = 0.25$	
%	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
0.900	2.68	2.96	2.51	2.82	2.35	2.65	2.25	2.48	1.86	2.15
0.905	2.70	2.98	2.56	2.86	2.39	2.70	2.27	2.53	1.90	2.20
0.910	2.74	3.00	2.60	2.90	2.45	2.74	2.31	2.57	1.96	2.25
0.915	2.81	3.02	2.65	2.94	2.50	2.79	2.36	2.63	2.01	2.29
0.920	2.85	3.07	2.71	3.00	2.56	2.83	2.40	2.67	2.06	2.35
0.925	2.92	3.12	2.75	3.05	2.62	2.87	2.46	2.76	2.11	2.41
0.930	2.99	3.17	2.82	3.11	2.68	2.93	2.53	2.83	2.18	2.47
0.935	3.05	3.23	2.89	3.17	2.74	2.98	2.59	2.90	2.26	2.54
0.940	3.12	3.32	2.97	3.24	2.79	3.03	2.66	2.97	2.35	2.62
0.945	3.22	3.40	3.04	3.30	2.88	3.10	2.74	3.05	2.43	2.69
0.950	3.34	3.55	3.12	3.36	2.98	3.16	2.84	3.12	2.50	2.79
0.955	3.37	3.60	3.23	3.46	3.10	3.25	2.92	3.20	2.62	2.88
0.960	3.50	3.66	3.36	3.57	3.22	3.34	3.03	3.32	2.76	3.01
0.965	3.63	3.77	3.46	3.68	3.33	3.44	3.14	3.47	2.89	3.15
0.970	3.76	3.93	3.63	3.81	3.48	3.59	3.30	3.60	3.06	3.31
0.975	3.94	4.15	3.83	3.99	3.67	3.77	3.52	3.75	3.24	3.50
0.980	4.13	4.28	4.06	4.15	3.86	4.00	3.62	3.96	3.46	3.69
0.985	4.44	4.54	4.39	4.42	4.11	4.29	3.92	4.23	3.73	4.06
0.990	4.67	5.02	4.78	4.76	4.57	4.59	4.35	4.47	4.04	4.57
0.995	5.55	5.72	5.42	5.42	5.15	5.16	5.02	5.25	4.78	5.58

Table 3: Finite Sample Null Rejection Probability of W_{RQF} with 5% Nominal Size, T = 100, Known Break Date

AR(2) case: $u_t = \alpha u_{t-1} + \psi(u_{t-1} - u_{t-2}) + e_t$

		Model I		Model II			Model III		
α	ψ	AR	AN	AR	NP	AN	AR	NP	AN
1.00	0.00	0.06	0.06	0.09	0.08	0.10	0.08	0.09	0.09
	0.30	0.07	0.06	0.12	0.12	0.10	0.08	0.11	0.09
	0.50	0.05	0.06	0.10	0.13	0.09	0.10	0.12	0.09
	0.70	0.04	0.07	0.10	0.17	0.12	0.10	0.14	0.11
0.95	0.00	0.07	0.08	0.04	0.04	0.05	0.06	0.05	0.05
	0.30	0.07	0.05	0.04	0.03	0.03	0.05	0.04	0.04
	0.50	0.07	0.05	0.03	0.01	0.02	0.04	0.03	0.03
	0.70	0.06	0.07	0.02	0.01	0.01	0.03	0.02	0.02
0.90	0.00	0.06	0.08	0.04	0.04	0.03	0.05	0.05	0.05
	0.30	0.07	0.05	0.04	0.03	0.03	0.05	0.04	0.04
	0.50	0.08	0.05	0.05	0.01	0.02	0.05	0.03	0.03
	0.70	0.10	0.04	0.07	0.01	0.01	0.07	0.02	0.02
0.80	0.00	0.07	0.07	0.05	0.05	0.07	0.07	0.06	0.06
	0.30	0.08	0.05	0.06	0.03	0.03	0.07	0.04	0.04
	0.50	0.10	0.03	0.07	0.02	0.02	0.10	0.03	0.03
	0.70	0.18	0.06	0.10	0.03	0.04	0.18	0.04	0.04

Table 4: Finite Sample Power, Known Break Date, T=100

AR(2) case: $u_t = \alpha u_{t-1} + \psi(u_{t-1} - u_{t-2}) + e_t$

			Model I			Model II			Model III		
α	ψ	η	PS	$T^{-1}W$	W_{RQF}	PS	$T^{-1}W$	W_{RQF}	PS	$T^{-1}W$	W_{RQF}
1.00	0.00	0.10	0.04	0.05	0.18	0.05	0.07	0.11	0.06	0.08	0.20
		0.30	0.08	0.14	0.85	0.11	0.24	0.35	0.12	0.27	0.90
		0.50	0.12	0.28	1.00	0.17	0.54	0.71	0.20	0.60	1.00
	0.30	0.10	0.03	0.06	0.18	0.04	0.09	0.12	0.04	0.08	0.19
		0.30	0.05	0.10	0.84	0.06	0.17	0.23	0.06	0.17	0.85
		0.50	0.07	0.19	1.00	0.09	0.34	0.45	0.11	0.40	1.00
	0.50	0.10	0.03	0.06	0.19	0.03	0.08	0.10	0.02	0.09	0.17
		0.30	0.03	0.09	0.84	0.03	0.13	0.17	0.03	0.15	0.83
		0.50	0.04	0.15	1.00	0.05	0.23	0.31	0.05	0.29	1.00
	0.70	0.10	0.01	0.07	0.19	0.01	0.11	0.11	0.01	0.13	0.18
		0.30	0.01	0.09	0.86	0.02	0.13	0.14	0.02	0.16	0.81
		0.50	0.02	0.12	1.00	0.02	0.16	0.21	0.02	0.20	1.00
0.95	0.00	0.10	0.07	0.05	0.20	0.05	0.02	0.07	0.05	0.02	0.18
		0.30	0.10	0.15	0.86	0.14	0.21	0.34	0.16	0.24	0.87
		0.50	0.18	0.33	1.00	0.28	0.65	0.77	0.27	0.70	1.00
	0.30	0.10	0.05	0.05	0.17	0.03	0.01	0.05	0.03	0.01	0.13
		0.30	0.08	0.11	0.82	0.07	0.09	0.17	0.08	0.09	0.84
		0.50	0.12	0.21	0.99	0.16	0.35	0.50	0.16	0.41	1.00
	0.50	0.10	0.05	0.04	0.18	0.01	0.00	0.02	0.02	0.01	0.10
		0.30	0.06	0.08	0.82	0.04	0.04	0.10	0.05	0.03	0.78
		050	0.09	0.14	099	0.10	0.17	0.23	0.10	0.18	100
	0.70	0.10	0.04	0.04	0.18	0.00	0.00	0.02	0.00	0.00	0.08
		0.30	0.05	0.05	0.81	0.02	0.01	0.04	0.02	0.01	0.72
		0.50	0.05	0.08	0.99	0.04	0.04	0.09	0.03	0.04	0.99
0.90	0.00	0.10	0.08	0.04	0.19	0.06	0.00	0.09	0.08	0.00	0.15
		0.30	0.15	0.16	0.83	0.27	0.23	0.33	0.26	0.21	0.89
		0.50	0.29	0.41	1.00	0.43	0.79	0.85	0.43	0.82	1.00
	0.30	0.10	0.07	0.03	0.16	0.04	0.00	0.06	0.04	0.00	0.12
		0.30	0.13	0.09	0.81	0.17	0.07	0.24	0.17	0.07	0.83
		0.50	0.21	0.23	0.99	0.33	0.48	0.55	0.34	0.50	1.00
	0.50	0.10	0.06	0.02	0.16	0.02	0.00	0.04	0.03	0.00	0.10
		0.30	0.09	0.05	0.76	0.13	0.02	0.21	0.13	0.02	0.76
		0.50	0.16	0.14	0.99	0.29	0.22	0.37	0.26	0.22	1.00
	0.70	0.10	0.04	0.00	0.14	0.02	0.00	0.04	0.01	0.00	0.08
		0.30	0.07	0.02	0.66	0.08	0.00	0.20	0.07	0.00	0.72
		0.50	0.13	0.06	0.95	0.23	0.04	0.35	0.18	0.03	0.99
0.80	0.00	0.10	0.09	0.01	0.18	0.12	0.00	0.22	0.15	0.00	0.24
		0.30	0.28	0.15	0.85	0.60	0.30	0.53	0.58	0.28	0.90
		0.50	0.54	054	1.00	0.81	0.97	0.92	0.81	0.98	1.00
	0.30	0.10	0.07	0.00	0.15	0.11	0.00	0.17	0.11	0.00	0.21
		0.30	0.24	0.07	0.77	0.56	0.09	0.59	0.54	0.07	0.89
		0.50	0.49	0.32	1.00	0.83	0.83	0.81	0.76	0.79	1.00
	0.50	0.10	0.06	0.00	0.10	0.09	0.00	0.16	0.08	0.00	0.17
		0.30	0.20	0.03	0.64	0.58	0.02	0.72	0.55	0.01	0.91
		0.50	0.46	0.18	0.97	0.85	0.52	0.84	0.78	0.43	1.00
	0.70	0.10	0.04	0.00	0.11	0.08	0.00	0.21	0.04	0.00	0.19
		0.30	0.15	0.00	0.55	0.61	0.00	0.86	0.52	0.00	0.94
		0.50	0.35	0.05	0.94	0.89	0.09	0.94	0.81	0.08	1.00

Table 5: Finite Sample Null Rejection Probability of Exp-W_{RQF} with 5% Nominal Size, T = 100, Unknown Break Date

AR(2) Case: $u_t = \alpha u_{t-1} + \psi(u_{t-1} - u_{t-2}) + e_t$

		Model I		Model II			Model III		
α	ψ	AR	AN	AR	NP	AN	AR	NP	AN
1.00	0.00	0.09	0.09	0.09	0.09	0.11	0.10	0.10	0.10
	0.30	0.10	0.07	0.14	0.18	0.13	0.13	0.14	0.11
	0.50	0.08	0.05	0.13	0.16	0.13	0.12	0.18	0.11
	0.70	0.03	0.05	0.14	0.26	0.13	0.18	0.27	0.15
0.95	0.00	0.10	0.09	0.04	0.03	0.03	0.05	0.05	0.07
	0.30	0.08	0.08	0.05	0.04	0.03	0.07	0.04	0.04
	0.50	0.09	0.06	0.03	0.04	0.03	0.06	0.03	0.03
	0.70	0.08	0.04	0.04	0.02	0.01	0.04	0.04	0.01
0.90	0.00	0.10	0.10	0.02	0.03	0.02	0.05	0.05	0.04
	0.30	0.12	0.09	0.04	0.03	0.02	0.08	0.03	0.03
	0.50	0.11	0.06	0.05	0.02	0.01	0.08	0.03	0.02
	0.70	0.19	0.04	0.13	0.05	0.02	0.15	0.03	0.02
0.80	0.00	0.14	0.11	0.04	0.03	0.04	0.07	0.06	0.05
	0.30	0.17	0.08	0.06	0.02	0.03	0.10	0.03	0.04
	0.50	0.26	0.05	0.12	0.02	0.03	0.19	0.02	0.04
	0.70	0.71	0.09	0.31	0.08	0.08	0.57	0.08	0.09

Table 6: Finite Sample Power, Unknown Break Date, T=100

AR(2) Case: $u_t = \alpha u_{t-1} + \psi(u_{t-1} - u_{t-2}) + e_t$

			Model I			Model II			Model III		
α	ψ	η	PS	$T^{-1}W$	W_{RQF}	PS	$T^{-1}W$	W_{RQF}	PS	$T^{-1}W$	W_{RQF}
1.00	0.00	0.10	0.05	0.06	0.08	0.06	0.06	0.12	0.06	0.08	0.15
		0.30	0.04	0.04	0.36	0.10	0.24	0.29	0.11	0.24	0.52
		0.50	0.06	0.04	0.92	0.16	0.53	0.59	0.17	0.49	0.98
	0.30	0.10	0.04	0.09	0.06	0.05	0.08	0.14	0.03	0.09	0.15
		0.30	0.04	0.07	0.37	0.05	0.15	0.25	0.05	0.16	0.43
		0.50	0.03	0.08	0.92	0.07	0.36	0.41	0.10	0.32	0.94
	0.50	0.10	0.01	0.11	0.07	0.02	0.10	0.12	0.02	0.10	0.10
		0.30	0.03	0.11	0.37	0.03	0.14	0.16	0.02	0.14	0.38
		0.50	0.02	0.10	0.91	0.04	0.27	0.28	0.04	0.29	0.95
	0.70	0.10	0.01	0.18	0.05	0.02	0.14	0.14	0.01	0.14	0.15
		0.30	0.01	0.17	0.36	0.02	0.16	0.18	0.01	0.17	0.41
		0.50	0.01	0.17	0.96	0.02	0.25	0.21	0.02	0.24	0.93
0.95	0.00	0.10	0.08	0.02	0.10	0.04	0.01	0.05	0.06	0.01	0.08
		0.30	0.07	0.01	0.38	0.11	0.15	0.17	0.11	0.14	0.42
		0.50	0.07	0.01	0.91	0.21	0.58	0.56	0.23	0.53	0.98
	0.30	0.10	0.06	002	0.10	0.02	0.01	0.06	0.02	0.01	0.06
		0.30	0.05	0.01	0.35	0.06	0.06	0.12	0.07	0.06	0.32
		0.50	0.04	0.01	0.90	0.13	0.32	0.29	0.15	0.30	0.95
	0.50	0.10	0.03	0.01	0.07	0.01	0.01	0.02	0.01	0.01	0.03
		0.30	0.03	0.01	0.33	0.03	0.02	0.05	0.05	0.03	0.24
		050	0.03	0.01	0.90	0.07	0.14	0.10	0.09	0.12	0.88
	0.70	0.10	0.02	0.00	0.04	0.01	0.00	0.02	0.01	0.00	0.02
		0.30	0.02	0.00	0.30	0.01	0.00	0.03	0.01	0.00	0.17
		0.50	0.03	0.00	0.91	0.04	0.03	0.05	0.03	0.02	0.82
0.90	0.00	0.10	0.09	0.00	0.09	0.05	0.00	0.05	0.07	0.00	0.07
		0.30	0.07	0.00	0.35	0.19	0.14	0.13	0.22	0.13	0.44
		0.50	0.08	0.00	0.92	0.30	0.70	0.51	0.36	0.59	0.98
	0.30	0.10	0.04	0.00	0.08	0.02	0.00	0.04	0.04	0.00	0.05
		0.30	0.05	0.00	0.33	0.12	0.03	0.10	0.13	0.04	0.30
		0.50	0.05	0.00	0.88	0.25	0.34	0.28	0.25	0.29	0.95
	0.50	0.10	0.03	0.00	0.06	0.02	0.00	0.02	0.02	0.00	0.02
		0.30	0.04	0.00	0.30	0.11	0.01	0.12	0.13	0.01	0.26
		0.50	0.03	0.00	0.86	0.21	0.11	0.16	0.21	0.08	0.88
	0.70	0.10	0.02	0.00	0.04	0.01	0.00	0.04	0.01	0.00	0.02
		0.30	0.02	0.00	0.24	0.06	0.00	0.15	0.06	0.00	0.20
		0.50	0.02	0.00	0.77	0.18	0.01	0.25	0.15	0.01	0.81
0.80	0.00	0.10	0.08	0.00	0.11	0.09	0.00	0.12	0.13	0.00	0.13
		0.30	0.07	0.00	0.36	0.19	0.14	0.26	0.48	0.09	0.54
		0.50	0.16	0.00	0.89	0.65	0.91	0.50	0.65	0.80	0.98
	0.30	0.10	0.05	0.00	0.09	0.06	0.00	0.09	0.09	0.00	0.10
		0.30	0.06	0.00	0.29	0.12	0.02	0.37	0.44	0.02	0.51
		0.50	0.08	0.00	0.86	0.61	0.58	0.50	0.63	0.46	0.96
	0.50	0.10	0.03	0.00	0.05	0.06	0.00	0.08	0.07	0.00	0.08
		0.30	0.04	0.00	0.16	0.11	0.00	0.51	0.45	0.00	0.61
		0.50	0.06	0.00	0.70	0.69	0.22	0.67	0.63	0.14	0.94
	0.70	0.10	0.01	0.00	0.08	0.04	0.00	0.17	0.04	0.00	0.18
		0.30	0.02	0.00	0.16	0.06	0.00	0.76	0.41	0.00	0.65
		0.50	0.03	0.00	0.54	0.74	0.02	0.91	0.70	0.00	0.96

.**Table 7: Asymptotic Distribution of the Exp-W_{FS} Test for Two Breaks**

	$\epsilon = 0.01$		$\epsilon = 0.05$		$\epsilon = 0.10$		$\epsilon = 0.15$		$\epsilon = 0.25$	
%	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)	I(0)	I(1)
Model I										
0.900	2.10	2.39	1.68	2.06	1.39	1.70	0.83	1.18	-0.81	-0.45
0.950	2.62	2.81	2.24	2.43	2.18	2.18	1.47	1.69	-0.15	0.02
0.975	3.14	3.22	2.87	3.09	2.67	2.67	2.21	2.27	0.51	0.54
0.990	4.03	3.99	3.53	4.25	3.47	3.43	3.27	2.96	1.49	1.35
Model II										
0.900	1.75	1.97	1.44	1.58	1.10	1.20	0.50	0.73	-1.10	-0.98
0.950	2.51	2.56	2.06	2.06	1.77	1.90	1.36	1.23	-0.35	-0.35
0.975	3.10	3.16	2.62	2.65	2.44	2.60	1.93	1.80	0.30	0.38
0.990	3.96	3.96	3.40	3.54	3.31	3.52	2.77	2.69	1.22	1.22
Model III										
0.900	3.81	4.68	3.43	4.24	3.16	3.81	2.57	3.39	0.81	1.49
0.950	4.75	5.33	4.23	4.96	3.94	4.55	3.38	4.14	1.57	2.23
0.975	5.62	6.11	5.01	5.75	4.67	5.23	4.19	4.87	2.23	2.96
0.990	6.91	6.85	5.87	6.82	5.60	6.04	5.39	5.88	3.23	4.07

Table 8: Historical Annual GDP Series: 1870-1986

		GDP				Per Capita GDP			
				Growth Rates				Growth Rates	
Country	Model	W_{RQF}	Break	Pre-	Post-	W_{RQF}	Break	Pre-	Post-
Australia	III	3.9^b	1929	2.4%	4.0%	3.2^b	1929	0.2%	2.2%
Canada	III	1.6	1930	3.3%	4.8%	1.5	1930	1.6%	2.9%
Denmark	III	8.9^a	1939	2.7%	3.7%	9.4^a	1939	1.6%	2.9%
France	III	27.2^a	1943	1.1%	5.0%	25.3^a	1943	1.0%	4.1%
Germany	III	32.1^a	1954	2.2%	3.6%	55.7^a	1945	1.4%	4.3%
Italy	III	15.9^a	1943	1.6%	4.9%	17.1^a	1943	0.9%	4.3%
Norway	III	7.3^a	1948	2.2%	4.2%	7.7^a	1925	1.0%	2.9%
Sweden	III	12.9^a	1916	2.5%	3.3%	13.6^a	1916	1.9%	2.7%
U.K.	III	11.8^a	1919	1.8%	2.3%	11.0^a	1919	1.0%	1.8%
U.S.	III	6.1^a	1929	3.7%	3.4%	5.6^a	1940	1.6%	1.5%

Note: a, b and c denote a statistic significant at the the 1%, 5% and 10% level, respectively.

Table 9: Post War Real GDP Series for the G7 Countries

						Annual Growth Rate	
Country	Period	T	Model	W_{RQF}	Break Date	Pre-	Post-
Canada	1957:1-2002:4	184	III	5.2^a	1973:4	5.5%	2.2%
France	1965:1-2002:4	152	III	37.1^a	1974:2	5.4%	2.1%
Germany	1960:1-2002:4	172	III	34.9^a	1974:4	5.1%	2.8%
Italy	1960:1-2002:4	172	III	97.7^a	1978:1	6.1%	2.4%
Japan	1957:1-2002:4	184	III	40.4^a	1972:1	9.8%	2.6%
U.K.	1957:1-2002:4	184	III	3.2^c	1974:1	3.2%	2.4%
U.S.	1947:1-1998:2	206	II	0.5	1973:2	3.8%	2.7%
			III	3.8^b	1965:4	3.5%	2.8%

Note: a, b and c denote a statistic significant at the the 1%, 5% and 10% level, respectively.

Figure 1: Finite Sample Size of W_{FMS} Test

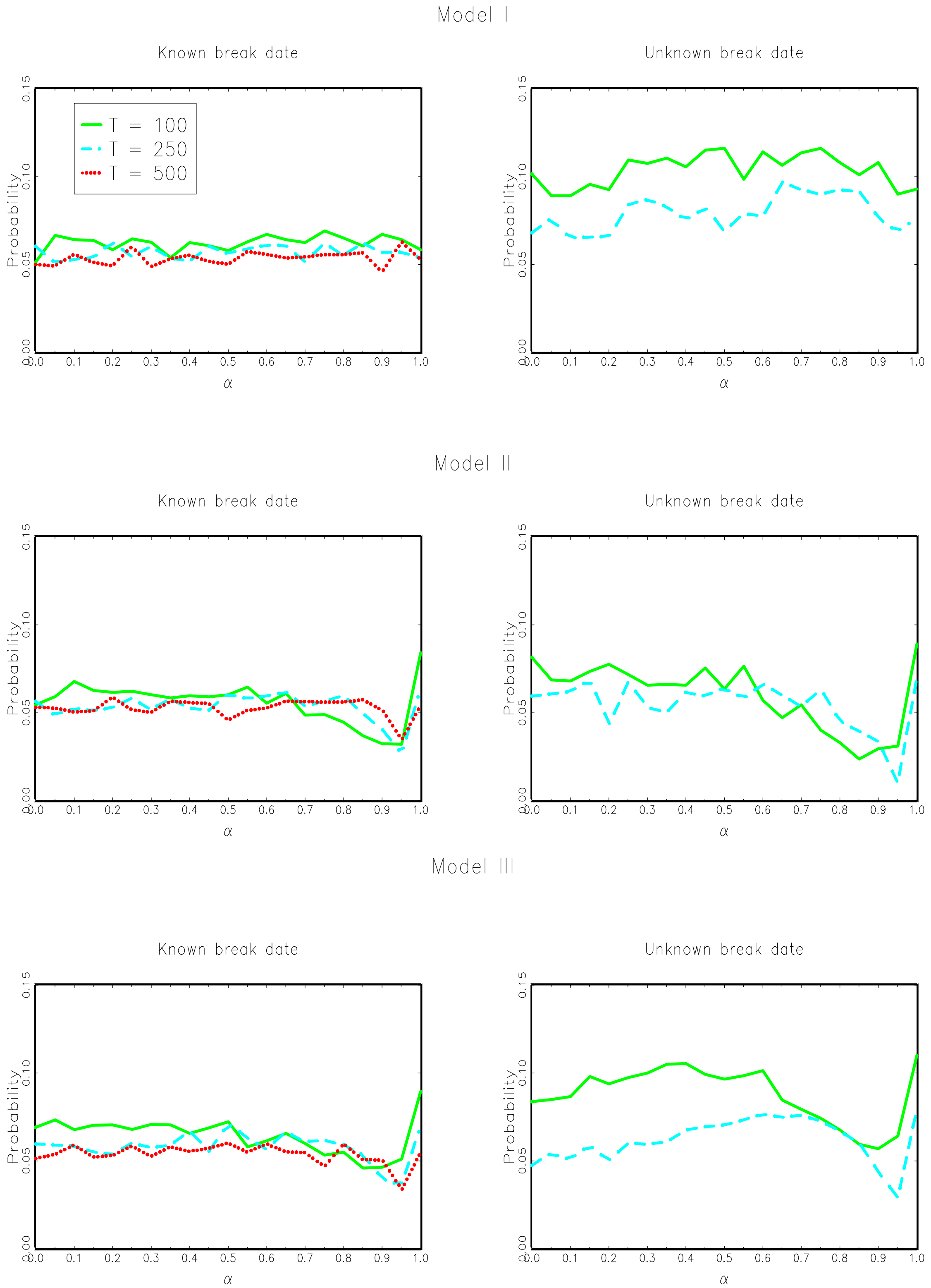

Figure 2: Finite Sample Power

Model I, Known Break Date, T=100

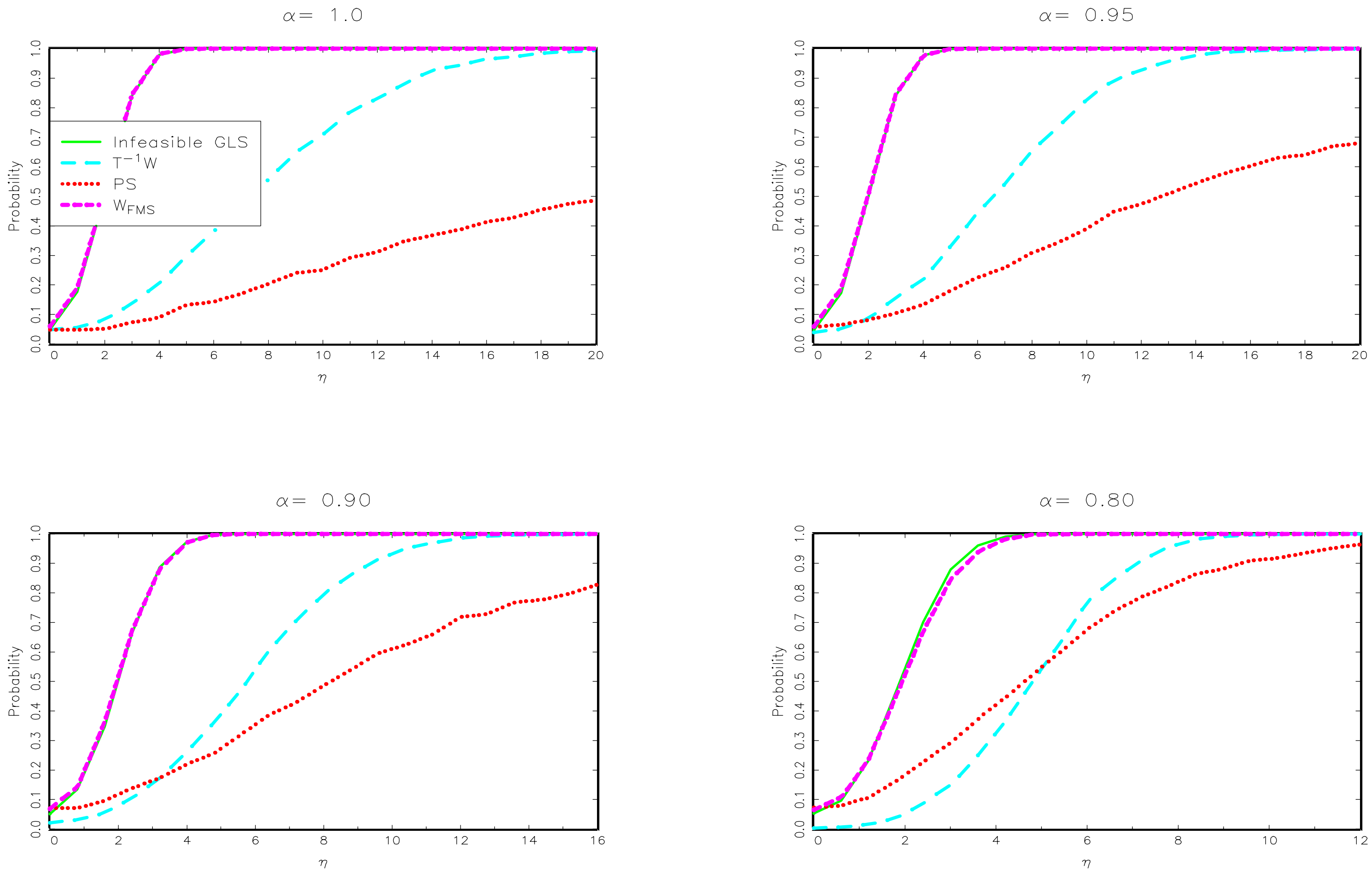

Figure 3: Finite Sample Power of the Exp Tests

Model I, Unknown Break Date, T=100

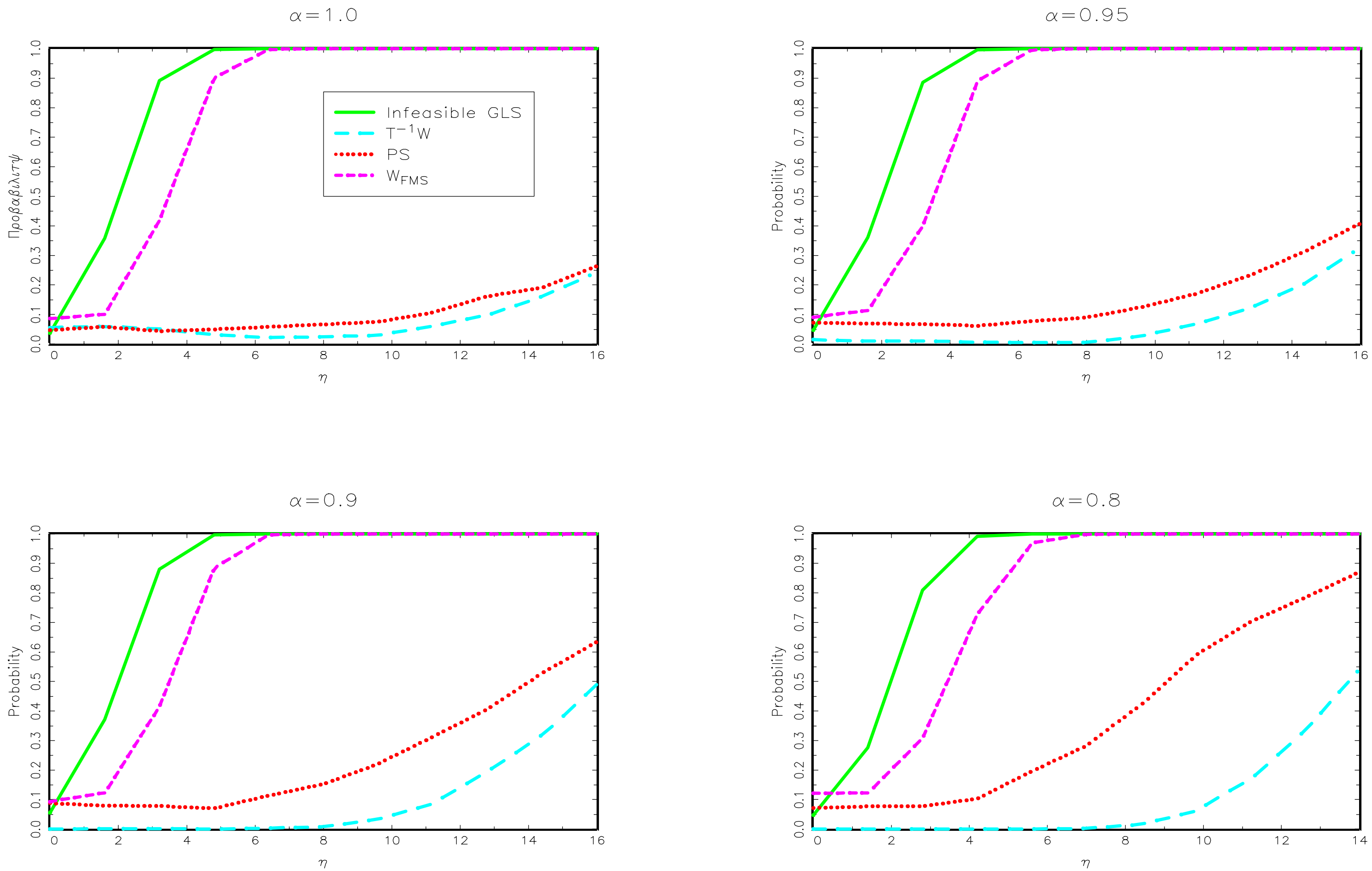

Figure 4.a: Finite Sample Power

Model II, Known Break Date, T=100

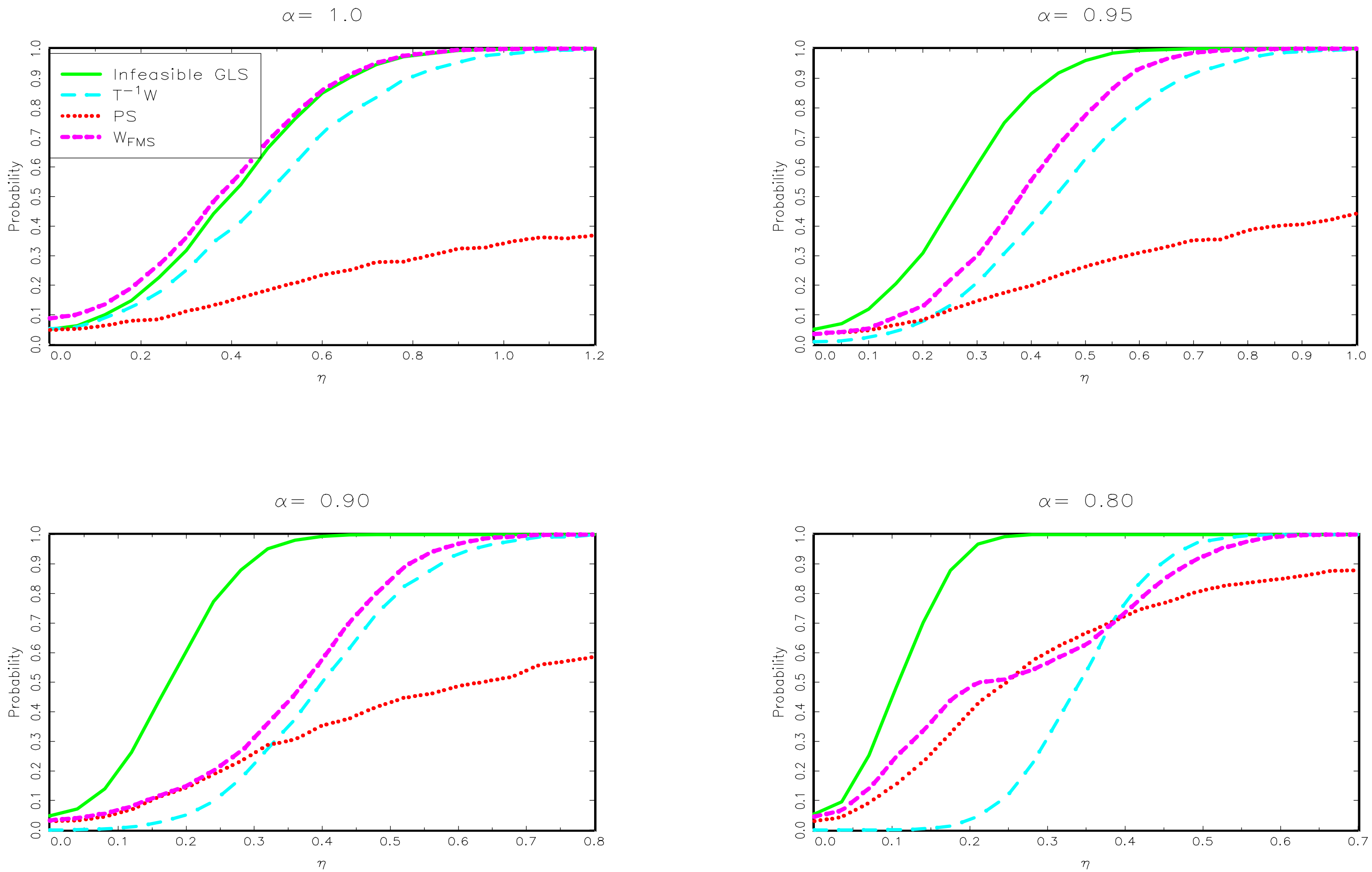

Figure 4.b: Finite Sample Power

Model II, Known Break Date, T=250

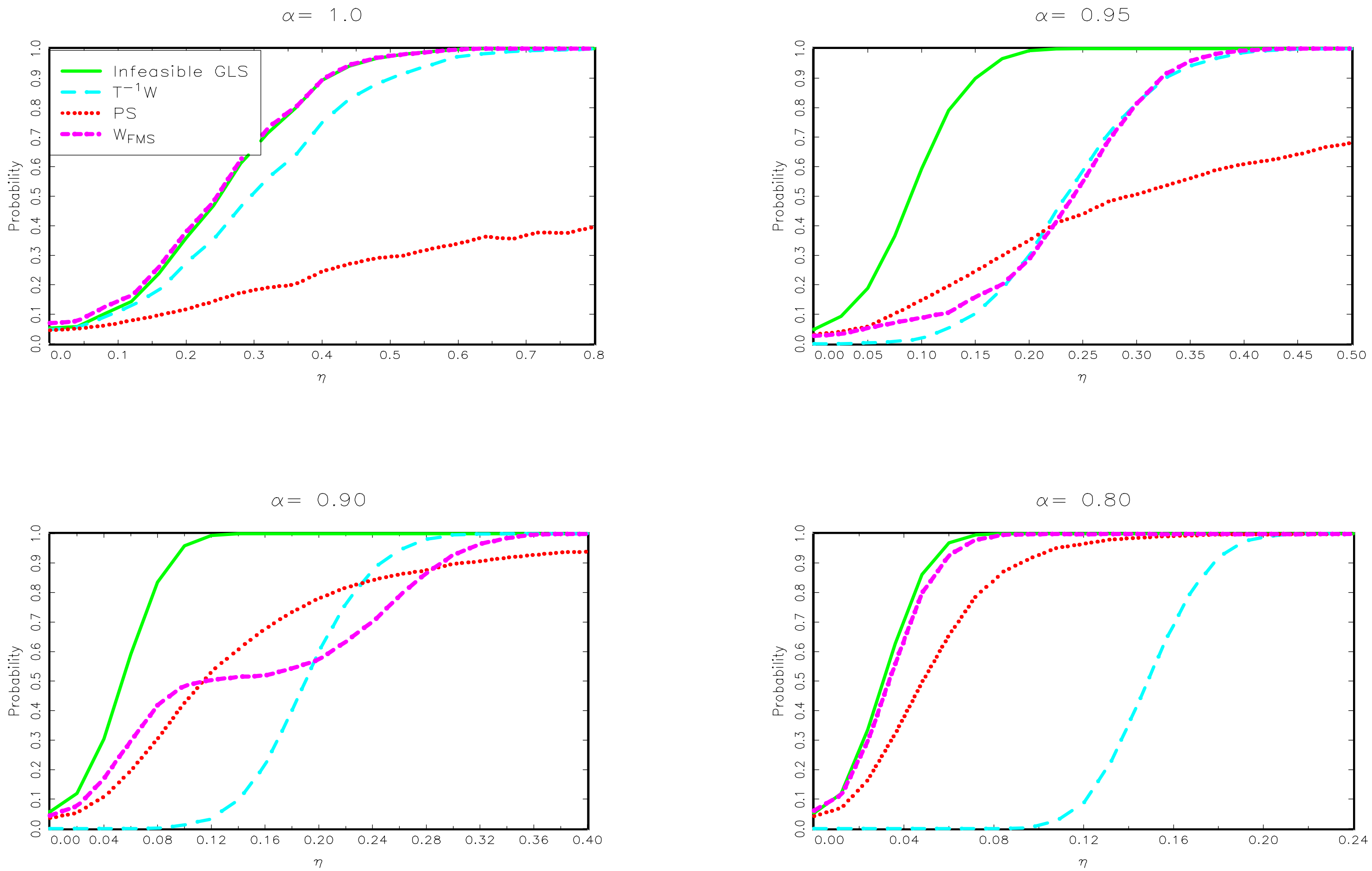

Figure 4.c: Finite Sample Power

Model II, Known Break Date, T=500

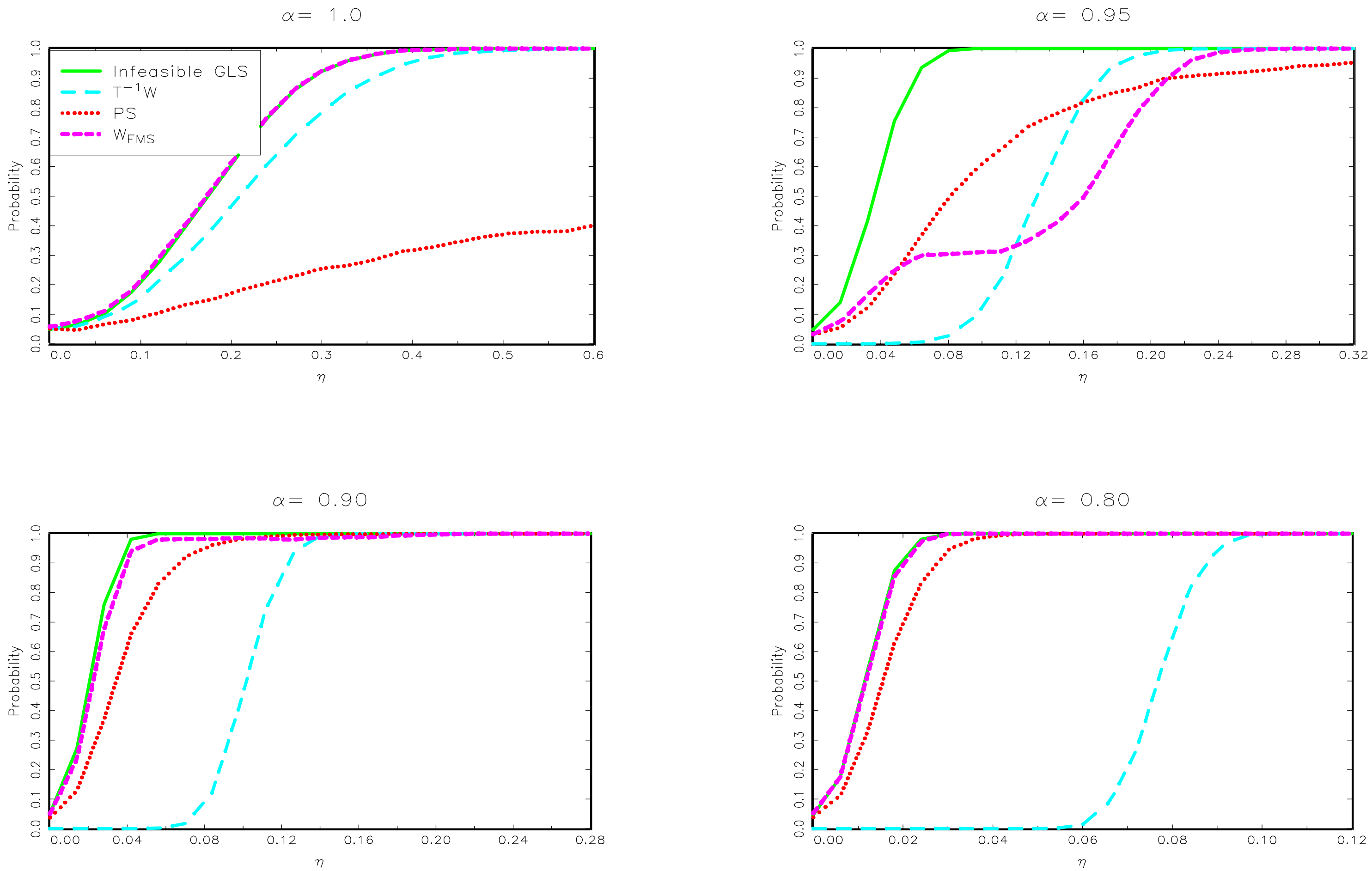

Figure 5.a: Finite Sample Power of the Exp Tests
Model II, Unknown Break Date, T=100

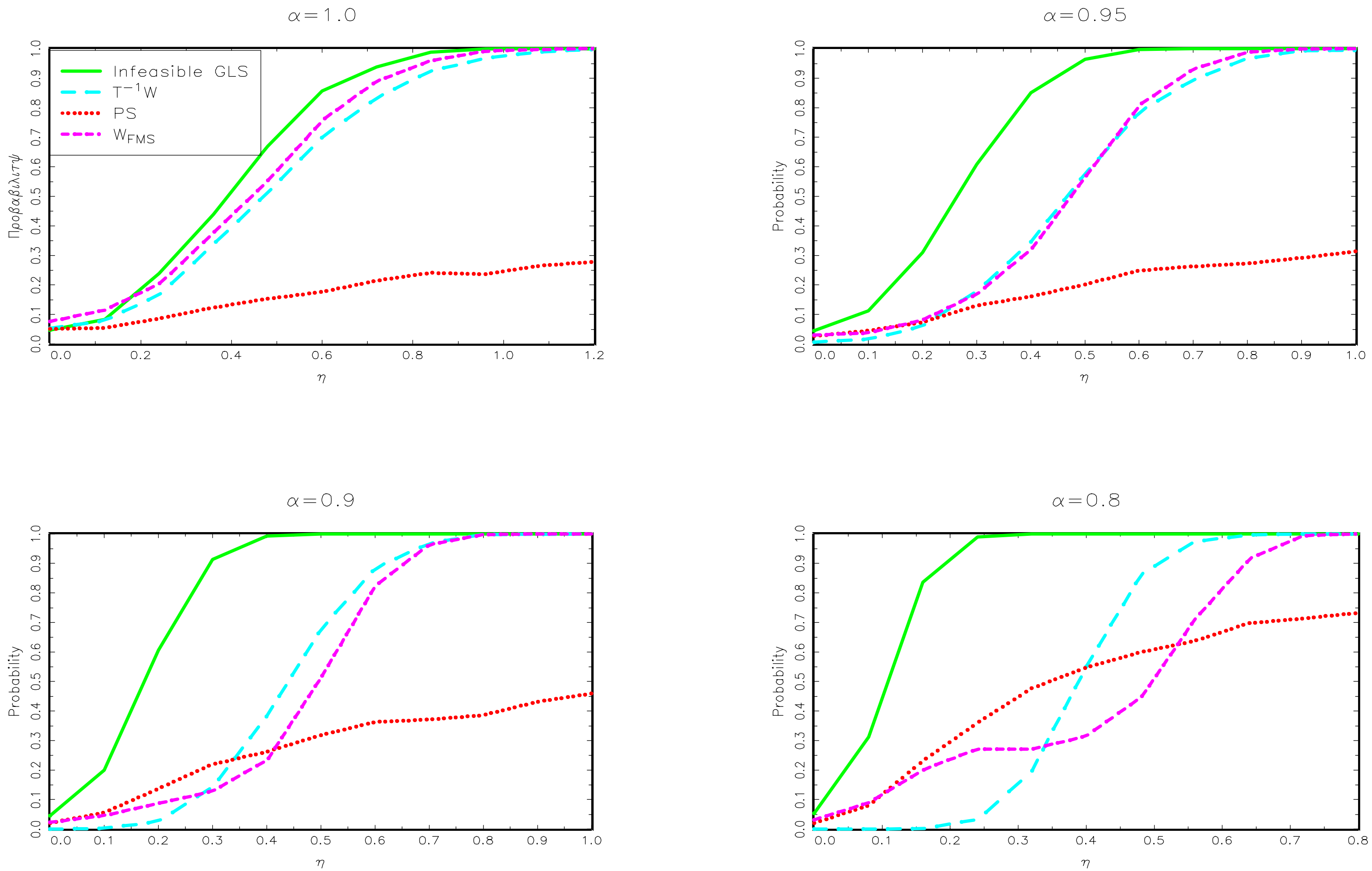

Figure 5.b: Finite Sample Power of the Exp Tests
Model II, Unknown Break Date, T=250

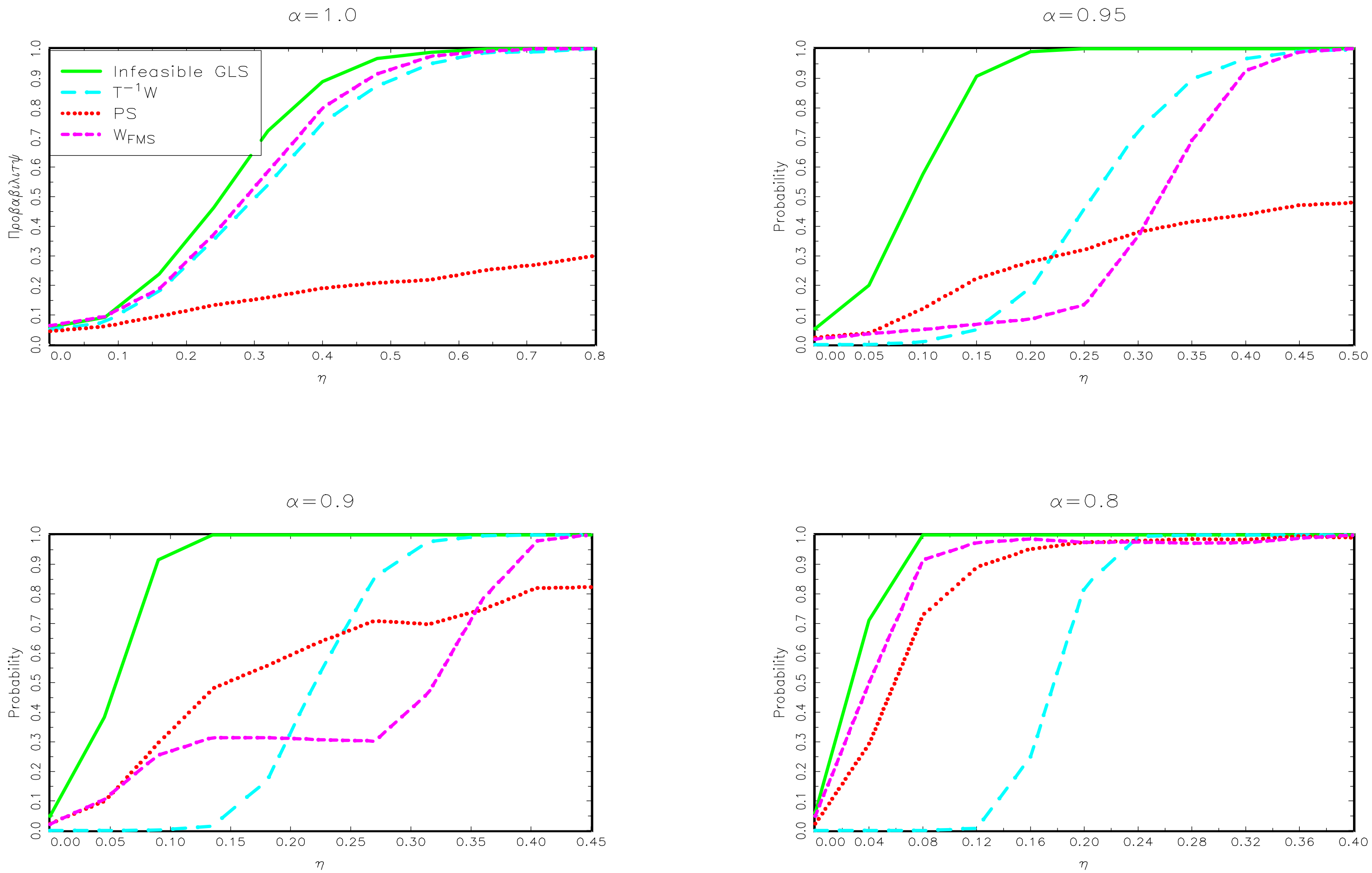

Figure 6.a: Finite Sample Power

Model III, Known Break Date, T=100

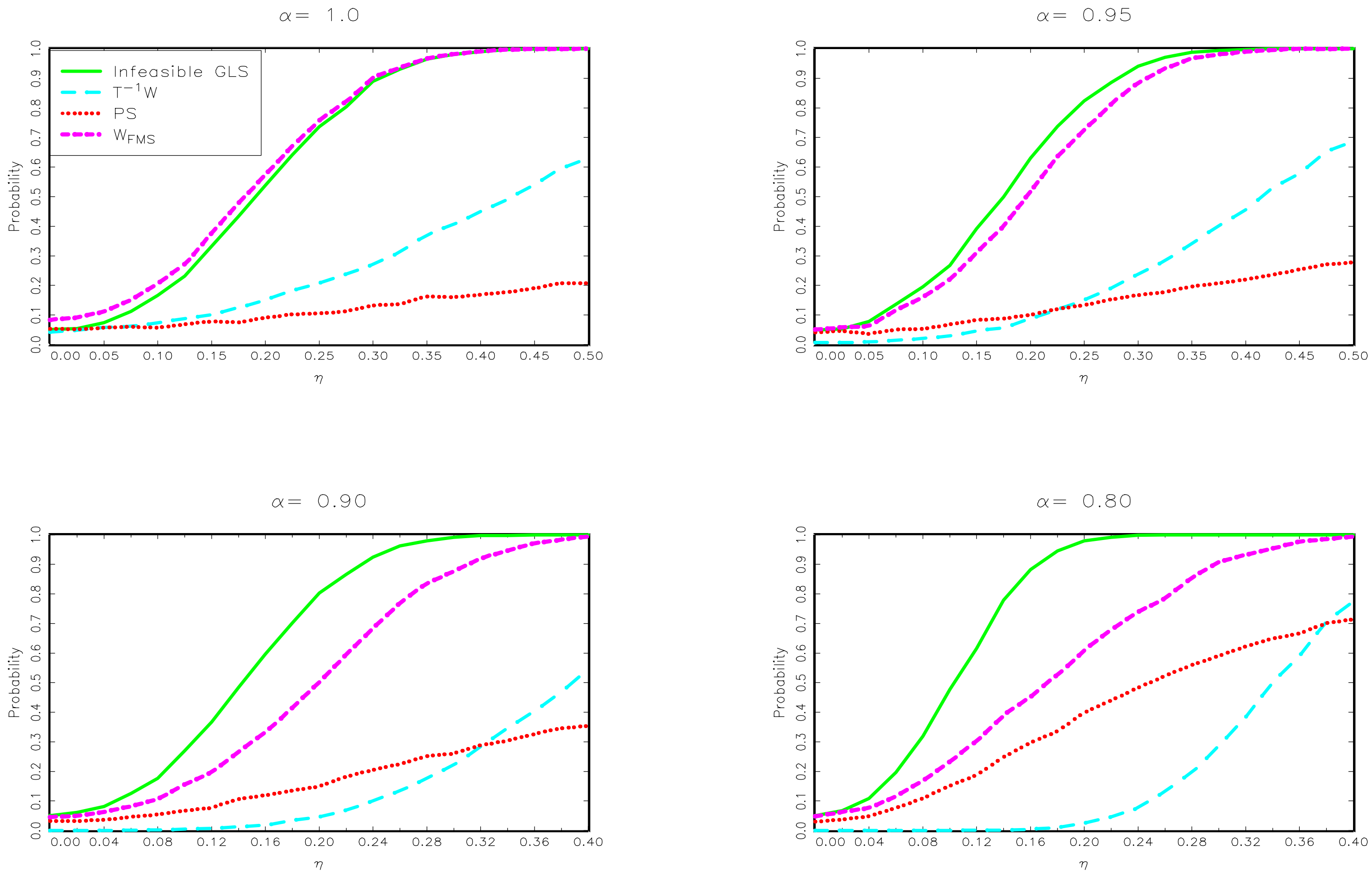

Figure 6.b: Finite Sample Power

Model III, Known Break Date, T=250

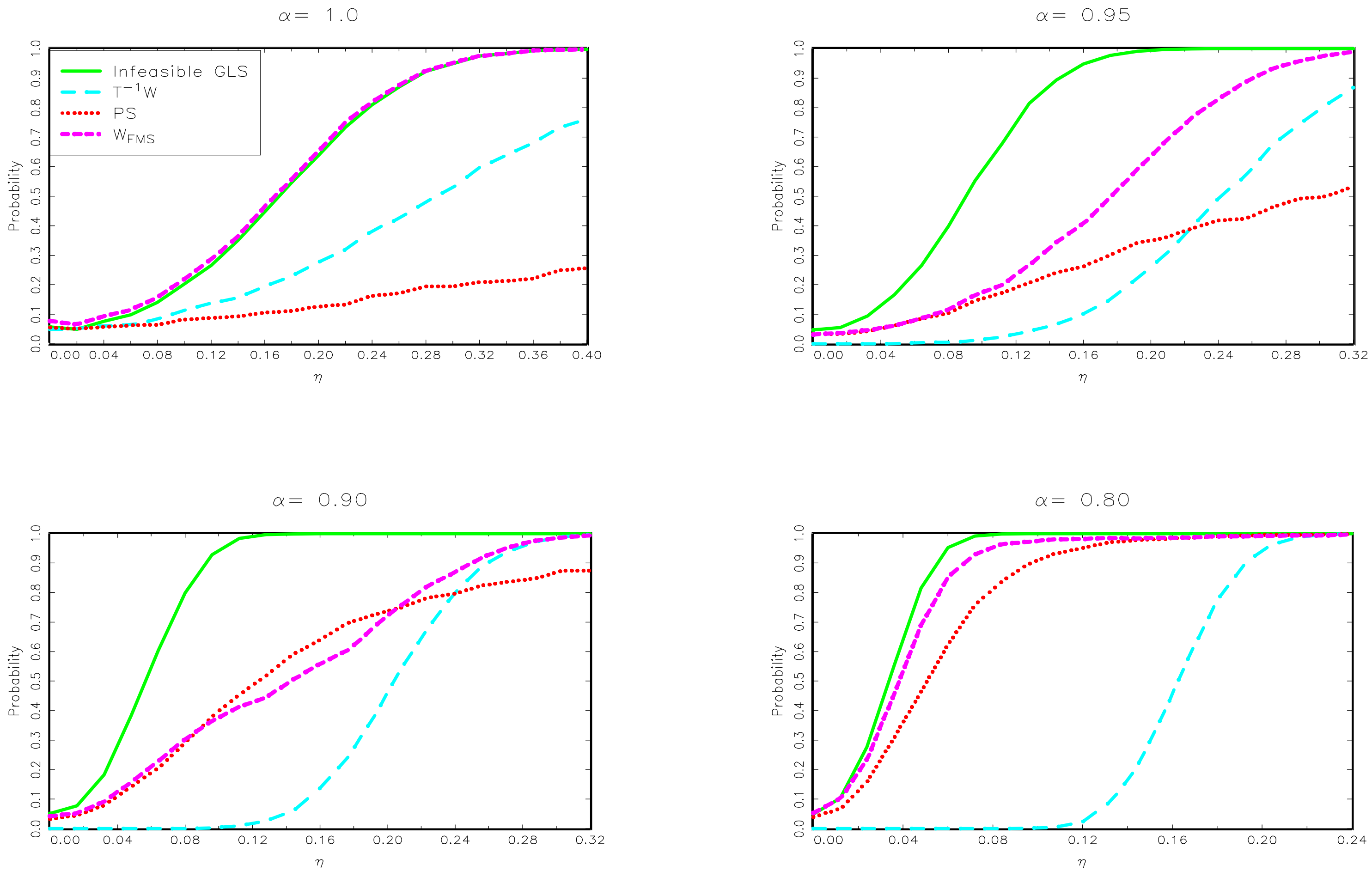

Figure 6.c: Finite Sample Power

Model III, Known Break Date, T=500

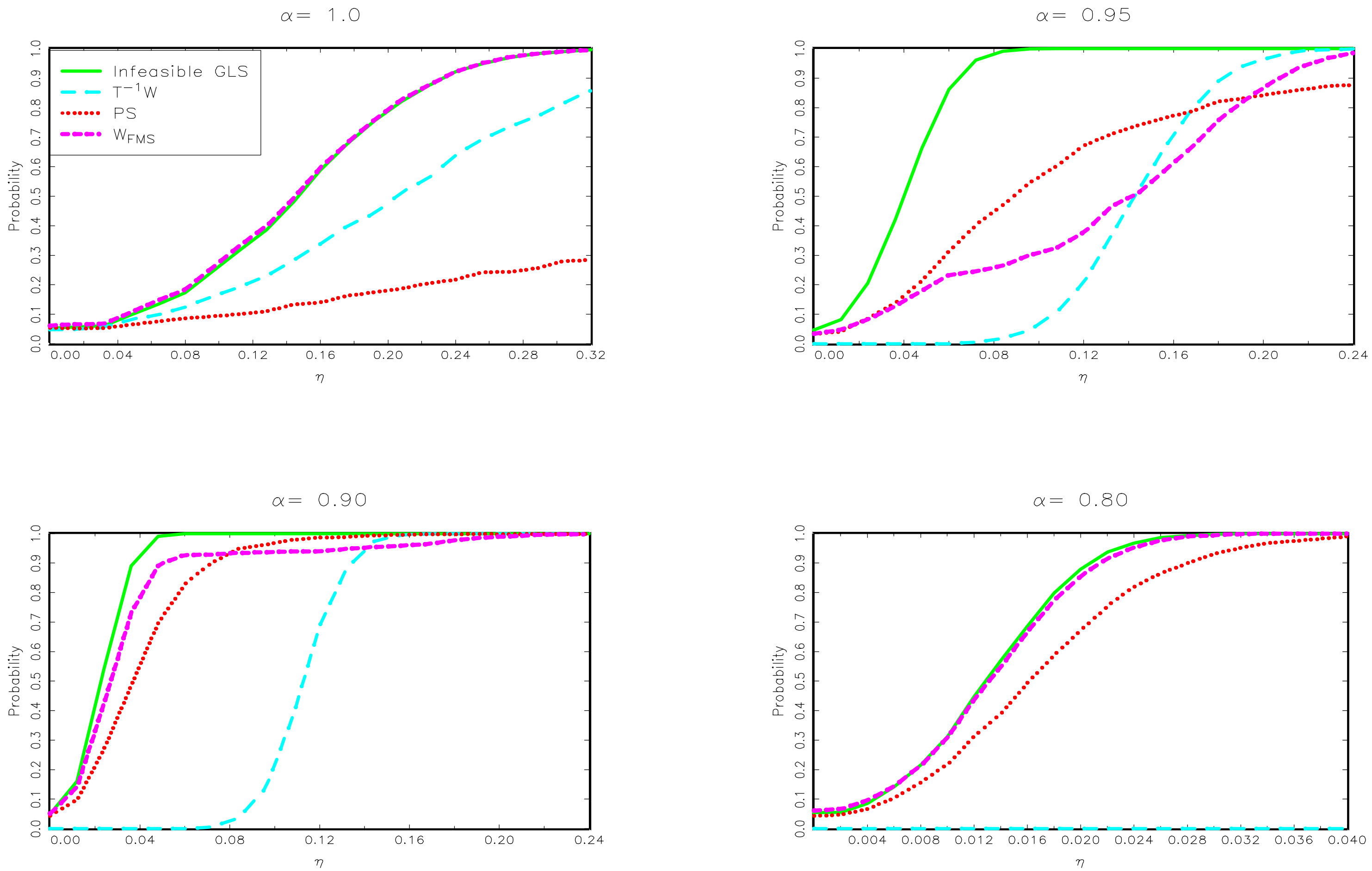

Figure 7.a: Finite Sample Power of the Exp Tests

Model III, Unknown Break Date, T=100

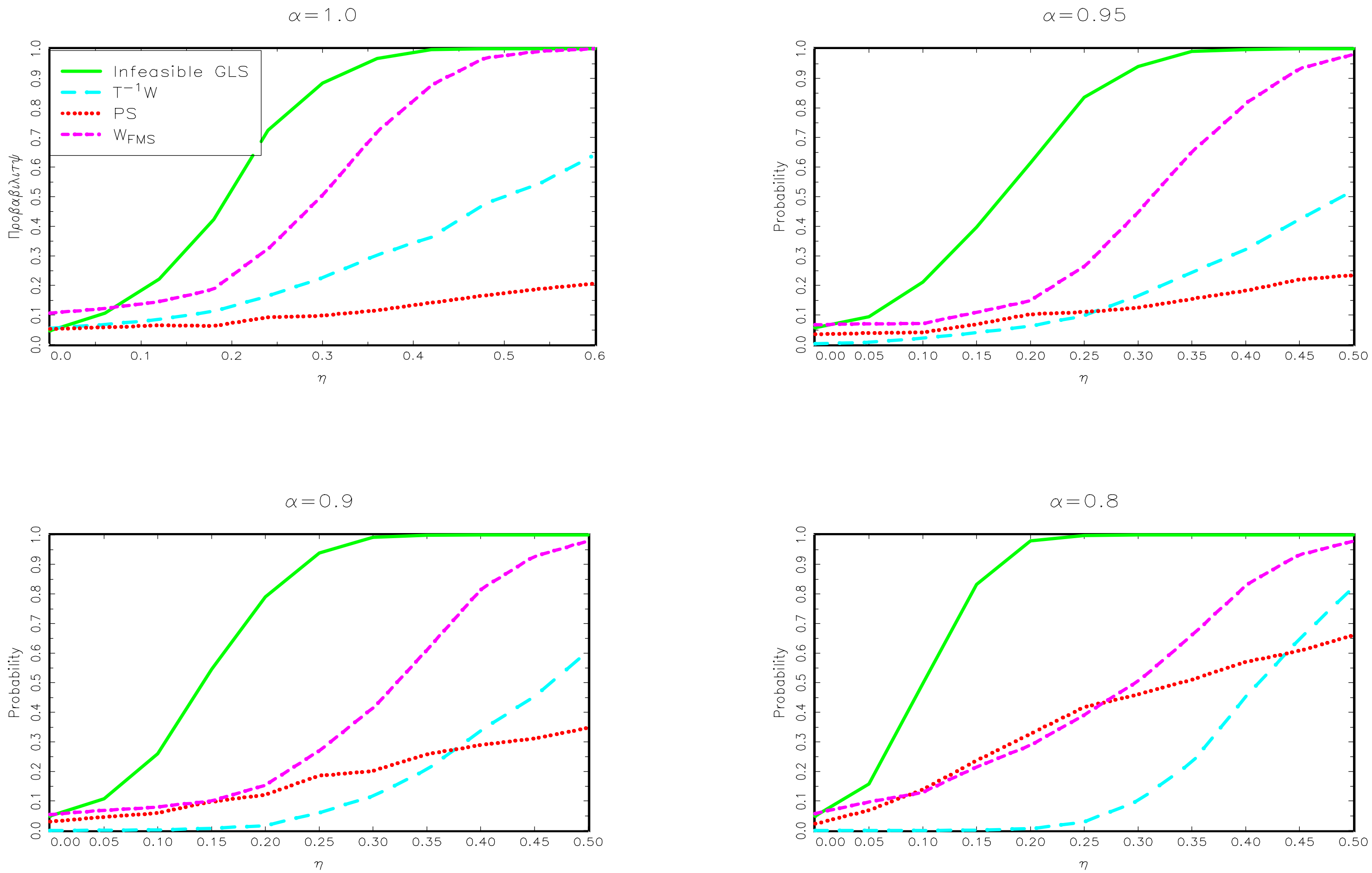

Figure 7.b: Finite Sample Power of the Exp Tests

Model III, Unknown Break Date, T=250

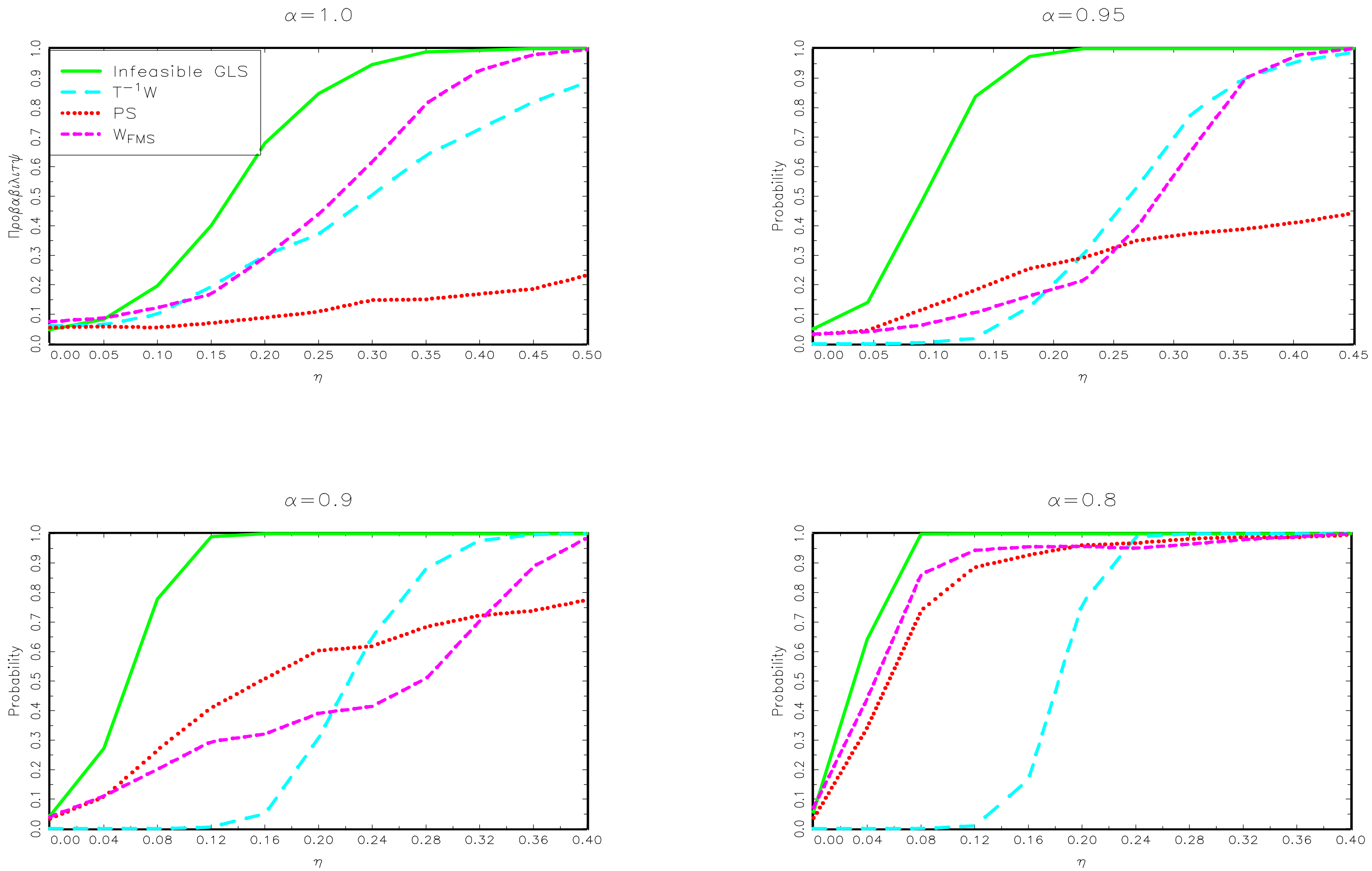

Figure 8.a: Historical Real GDP Series, 1870-1986

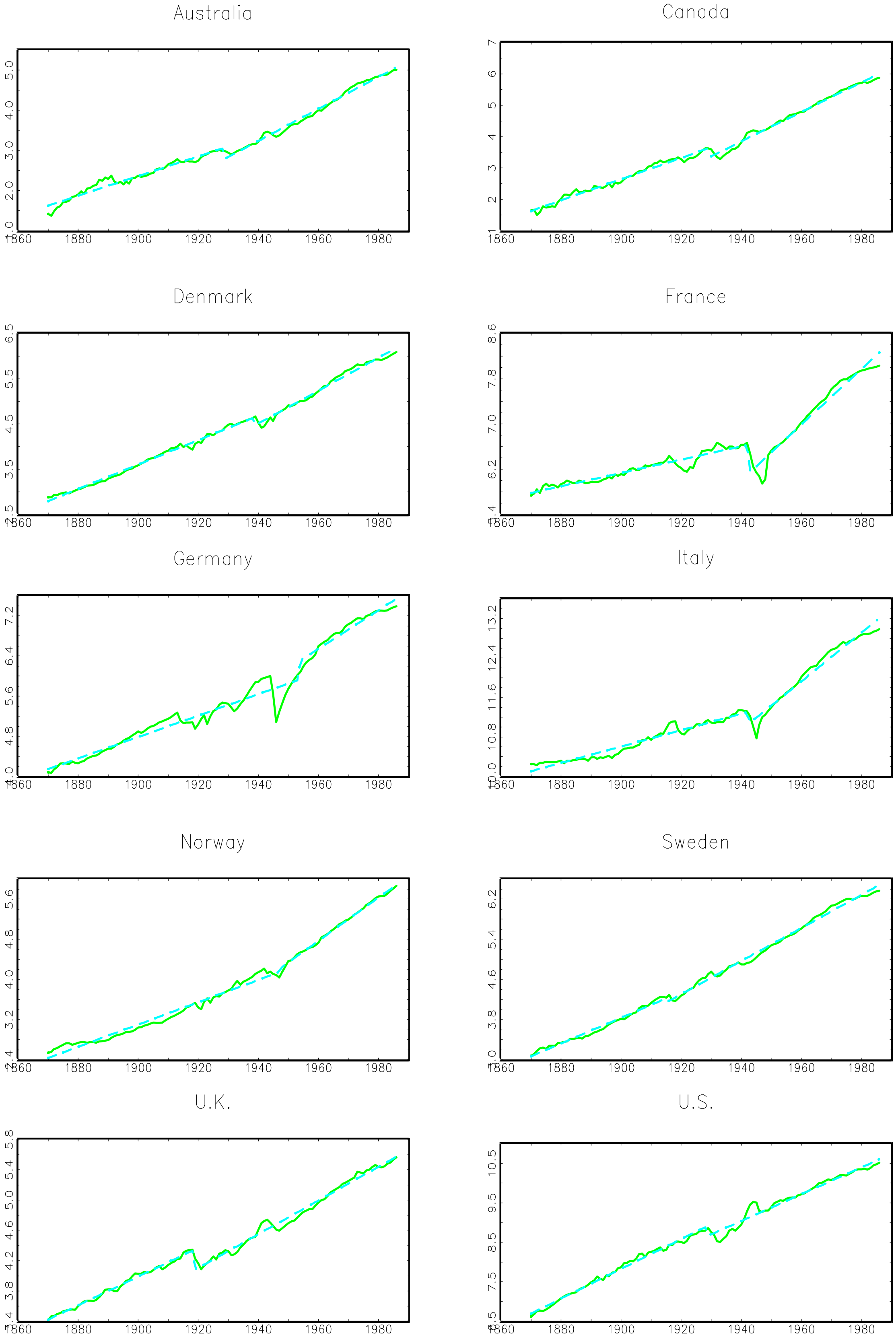

Figure 8.b: Historical Per Capita Real GDP Series, 1870-1986

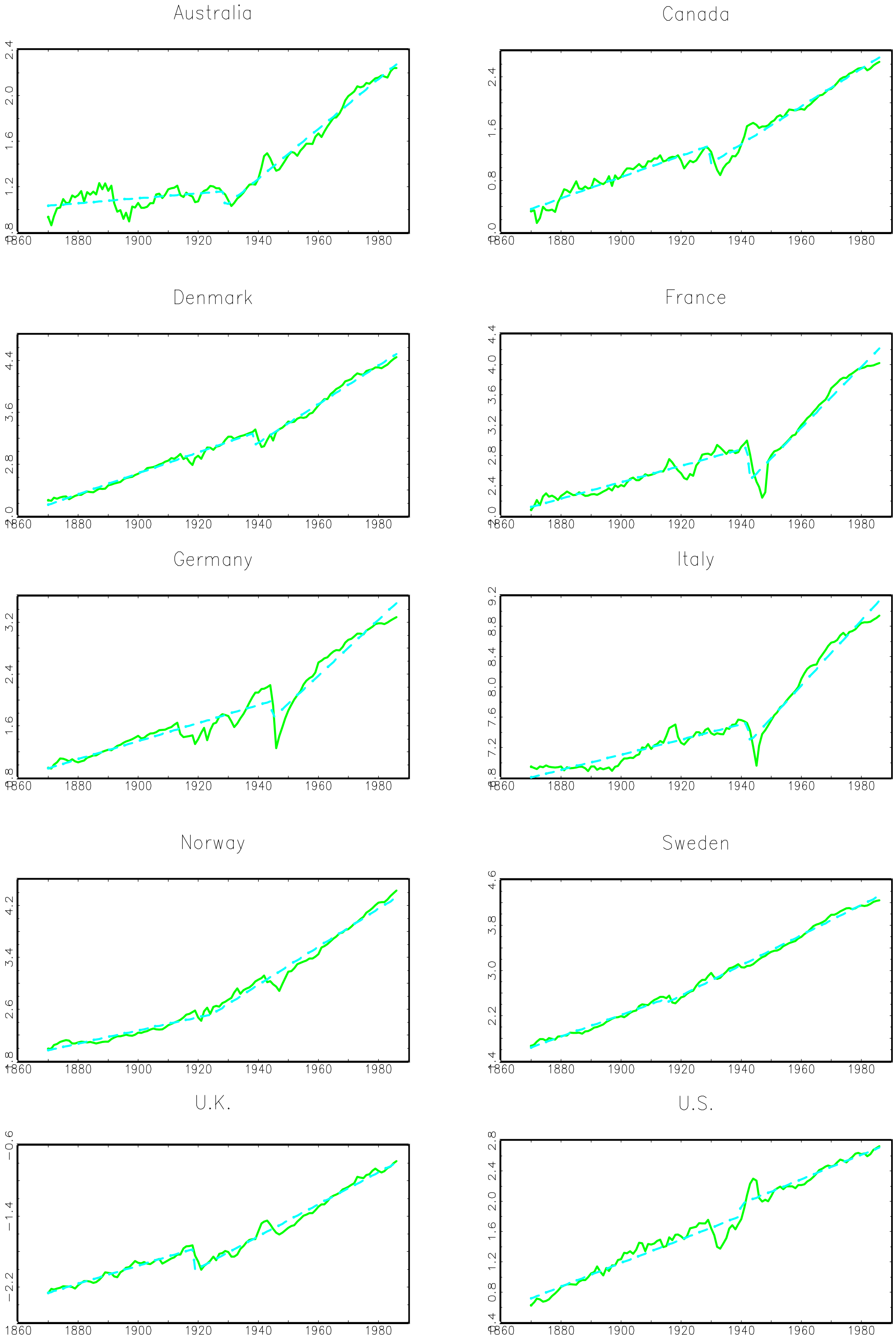

Figure 9: Postwar Quarterly Real GDP for the G7 Countries