Granger Centre
for time series econometrics

Granger Centre Discussion Paper Series

On the behaviour of fixed-*b* trend break tests under fractional integration

by

Fabrizio Iacone, Stephen J. Leybourne and A. M. Robert Taylor

---

Granger Centre Discussion Paper No. 11/03

The University of Nottingham

On the behaviour of fixed-b trend break tests under fractional integration*

Fabrizio Iacone[a], Stephen J. Leybourne[b] and A.M. Robert Taylor[b]

[a] Department of Economics and Related Studies, University of York

[b] School of Economics and Granger Centre for Time Series Econometrics, University of Nottingham

November 2011

Abstract

Testing for the presence of a broken linear trend when the nature of the persistence in the data is unknown is not a trivial problem, since the test needs to be both asymptotically correctly sized and consistent, regardless of the order of integration of the data. In a recent paper, Sayginsoy and Vogelsang (2011) [SV] show that tests based on fixed-b asymptotics provide a useful solution to this problem in the case where the shocks may be either weakly dependent or display strong dependence within the near-unit root class. In this paper we analyse the performance of these tests when the shocks may be fractionally integrated, an alternative model paradigm which allows for either weak or strong dependence in the shocks. We demonstrate that the fixed-b trend break statistics converge to well-defined limit distributions under both the null and local alternatives in this case (and retain consistency against fixed alternatives), but that these distributions depend on the fractional integration parameter δ. As a result, it is only when δ is either zero or one that the SV critical values yield correctly sized tests. Consequently, we propose a procedure which employs δ-adaptive critical values to remove the size distortions in the SV test. In addition, use of δ-adaptive critical values also allows us to consider a simplification of the SV test which is (asymptotically) correctly sized across δ but can also provide a significant increase in power over the standard SV test when $\delta = 1$.

Keywords: Trend break; fractional integration; fixed-b asymptotics.

JEL classification: C22.

*We thank two anonymous referees for their helpful and constructive comments on an earlier version of this paper. We also than participants at the conference in honour of Sir David F. Hendry held at St Andrews University on July 22-23 2010 for their comments. Correspondence to: Robert Taylor, School of Economics, The Sir Clive Granger Building, University of Nottingham, Nottingham NG7 2RD, U.K. E-mail: `Robert.Taylor@nottingham.ac.uk`

1 Introduction

Testing for the presence of a broken linear trend at an unknown break date when the nature of the persistence of the data is not known is not a trivial problem because the test needs to be asymptotically correctly sized and consistent regardless of the order of integration of the data. If the order of integration were known one could simply difference the data to the appropriate order and test the trend break hypothesis within the resulting model for the differenced data.

Such an approach, however, requires that the order of integration of the data is known. In practice this will not be the case and we would need to estimate the order of integration from the data. Even where we constrain attention to the case where the data are either weakly dependent, $I(0)$, or admit a unit root, $I(1)$, a circular testing problem arises. This is because, as demonstrated in Perron (1989), unit root tests are highly sensitive to trend breaks, and failure to account for them results in spurious acceptance of the unit root hypothesis even when the data are $I(0)$. On the other hand, unnecessarily allowing for a trend break in unit root tests seriously reduces their power, other things being equal, especially when the location of the potential break is treated as unknown; see, *inter alia*, Perron (2005), Zivot and Andrews (1992) and Harris, Harvey, Leybourne and Taylor (2009). At the same time, as Harvey, Leybourne and Taylor (2009) [HLT] show, conventional regression-based tests for a trend break which assume the data are $I(0)$ spuriously reject when the data are $I(1)$.

Two recent papers have appeared which, under the assumption that the data are either $I(0)$ or $I(1)$, attempt to circumvent this circular inference problem. First, HLT propose a trend break statistic which is a weighted average of the largest (in absolute value) of the sequences over all possible break dates of conventional regression-based t-ratios for a trend break in the levels and first-difference regressions. The weight function employed is such that the statistic reduces to the largest t-ratio from the sequence of levels regressions when the data are $I(0)$ and to the largest t-ratio from the sequence of first difference regressions when the data are $I(1)$. Sayginsoy and Vogelsang (2011) [SV] take a different approach and base their trend break test on a function of the Wald-type statistic from the sequence of levels regressions alone, again taken over all possible break dates. HLT use separate conventional long run variance estimators for the $I(0)$ and $I(1)$ t-ratios, while the statistics proposed in SV embody the use of long run variance estimators employing the fixed-b bandwidth approach of Kiefer and Vogelsang (2002) to yield a statistic which has a well-defined limiting null distribution in both $I(0)$ and $I(1)$ environments. In both HLT and SV, a significance level-specific multiplicative correction factor is also required due to the fact that their statistics possess different limiting null distributions under $I(0)$ and $I(1)$ shocks.

The tests proposed in HLT and SV are therefore designed to be asymptotically correctly sized irrespective of whether the data are $I(0)$ or $I(1)$ (they are also consistent against fixed trend break alternatives). Under near-$I(1)$ shocks both tests display rejection frequencies which fall below the nominal asymptotic level (but remain consistent against fixed alternatives). Consequently, under either weak dependence or strong dependence of the (near-) $I(1)$ type, these tests are asymptotically well behaved and thereby avoid the circular testing problem noted above. These tests are therefore referred to as being 'robust' to strong correlation in the data. While this is true for strong correlation of the (near-) $I(1)$ type, an important extension to this paradigm is to consider the more general case of fractionally integrated, $I(\delta)$, shocks which also display strong dependence whenever $\delta > 0$. This is also relevant because of the ongoing debate as to whether it is actually unmodelled structural breaks that give rise to strong dependencies often found in observed data; see Sibbertsen (2004) and the references therein. En route to settling the debate, an obvious requirement it to be able to test whether structural change has actually occurred, where inference is not made contingent on any underlying assumption of weak or strong dependence in the data generation process.

Iacone *et al.* (2011) [ILT] show that when δ is fractional, the statistic proposed in HLT either collapses to zero or diverges to positive infinity in the limit. As a consequence the HLT test is rendered unworkable as a trend break test once we allow for the possibility of fractional integration. In contrast we show in this paper that under fractional integration the SV statistic converges to a well-defined limit distribution under both the null of no trend break and local trend break alternatives, but is consistent against fixed alternatives, regardless of δ. This limiting null distribution depends on the long memory parameter δ, however, so that the asymptotic size of the test when run using the asymptotic critical values given in SV is no longer the anticipated level, unless δ is zero or one. To counter this, we suggest a procedure which employs δ-adaptive critical values to remove these size distortions in the SV test. Two possible estimators of δ are explored, one which imposes the trend break null hypothesis and one which does not. Both are shown to yield asymptotically correctly sized and consistent tests.

The use of δ-adaptive critical values also allows us to consider a logical simplification of the SV test which is correctly sized in the limit across δ. The simplification arises from the fact that once δ-adaptive critical values are employed, the multiplicative correction factor embedded in the SV test (to line up the limiting null distribution of the test statistic under $I(0)$ and $I(1)$ errors at a given significance level) is made redundant. We show that the removal of this redundant term also opens the potential for a significant increase in power when $\delta > 0$.

After outlining our fractionally integrated trend break data generating process [DGP] in section 2 and detailing the fixed-b trend break tests of SV in section 3, we present the asymptotic properties of SV in section 4. These results cover the null, local trend break and fixed trend break alternatives. Section 5 introduces the δ-adaptive critical values schemes. The simplified SV tests are considered in section 6. Simulations are conducted to allow comparison of the relative sizes of the various procedures. In section 7 we examine the local power properties of each procedure. Our results on size and power allow us to make a concrete recommendation on the best procedure for use in practice. Concluding comments are given in section 8. The proof of our main theorem is provided in a mathematical appendix.

In what follows we use the notation: $x := y$ ($x =: y$) to indicate that x is defined by y (y is defined by x); $\lfloor \cdot \rfloor$ to denote the integer part of the argument; $\mathbb{I}(\cdot)$ to denote the indicator function which returns the value one when its argument is true and zero otherwise; $\xrightarrow{p}$ to denote convergence in probability and $\xrightarrow{d}$ to denote convergence in the Skorohod J_1 topology of $D[0,1]$, the space of real-valued functions on $[0,1]$ which are continuous on the right and with finite left limit, respectively; $\sim$ to indicate that the ratio between left- and right-sides tends to a finite positive constant. Finally, reference to a variable being $O_e(T^k)$, $k > 0$, denotes exact order in probability, in that the variable is of $O_p(T^k)$ but is not of $o_p(T^k)$.

2 The Fractionally Integrated Trend Break Model

Following Perron (1989), HLT and SV, *inter alia*, we consider the following trend break DGP

$$y_t = \alpha + \beta t + \beta^* DT_t(\tau_0) + u_t, \quad t = 1, ..., T, \tag{1}$$

where

$$DT_t(\tau_0) := \begin{cases} 0, & t \leq \lfloor \tau_0 T \rfloor \\ t - \lfloor \tau_0 T \rfloor, & t > \lfloor \tau_0 T \rfloor \end{cases}$$

for $\tau_0 \in [\tau_L, \tau_U] =: \Lambda \subseteq (0,1)$. The quantities τ_L and τ_U are called *trimming* parameters. Here u_t is taken to be a zero mean $I(\delta)$ process. In this model there is a break in trend at time $\lfloor \tau_0 T \rfloor$ when $\beta^* \neq 0$.

Our interest in this paper centres on testing the null hypothesis $H_0 : \beta^* = 0$ against the two-sided alternative hypothesis $H_1 : \beta^* \neq 0$, without assuming knowledge of the fractional integration parameter, δ, or the putative break fraction, τ_0. This generalizes the hypothesis structure considered in HLT and SV for $\delta \in \{0, 1\}$ to the case of fractional δ.

The trend break model is completed by formalizing the $I(\delta)$ properties of u_t.

Assumption 1 *For some process η_t which satisfies the conditions of Assumption 2 below, the process u_t is such that, for $\delta \in [0, 3/2)$, $u_t = \Delta^{-\delta} \eta_t \mathbb{I}(t > 0)$ where $\Delta^\delta = (1 - L)^\delta$, with L being the usual lag operator, and $\Delta^{-\delta} = \sum_{j=0}^{\infty} \psi_j L^j$ with $\psi_j := \Gamma(j + \delta) / (\Gamma(\delta) \Gamma(j + 1))$, $\Gamma(\cdot)$ denoting the Gamma function.*

Assumption 2 *η_t is a linear process satisfying $\eta_t = A(L) \varepsilon_t := \sum_{j=0}^{\infty} A_j \varepsilon_{t-j}$. The weights $\{A_j\}$ are such that $A(1)^2 > 0$ and $\sum_{l=0}^{\infty} l |A_l| < \infty$ and ε_t is an independent, identically distributed (i.i.d.) sequence with $E(\varepsilon_t) = 0$, $E(\varepsilon_t^2) = 1$ and $E(|\varepsilon_t|^q) < \infty$ for some $q > 2$.*

Remark 1. Assumptions 1 excludes the case $\delta < 0$, which we consider to be reasonable given that the trend break model in (1) is intended for the levels (rather than the differences) of the data. Note also that Assumption 1 is based on the definition of a Type II fractionally integrated process; see, for example, Marinucci and Robinson

(1999). Assumption 2 places conditions on the weights A_l which coincide with those made in SV and HLT, while the innovations, ε_t, are assumed to be i.i.d. The latter is commonly assumed in the fractional integration literature as it allows us to appeal to the functional central limit theorem of Marinucci and Robinson (2000) for fractionally integrated processes. A consequence of Assumption 2 is that the long run variance of η_t is given by $\sigma^2 := A(1)^2$. Moreover, we also define $\sigma := \sqrt{\sigma^2}$.

Remark 2. The model in (1) embodies a break in trend. SV additionally develop tests for a break in trend with a simultaneous break in level, and for a break in level (about a constant, possibly zero, trend). In the interests of brevity we focus attention in this paper only on (1), the leading case of interest. It would be entirely straightforward to apply a similar analysis to these other models and the resulting tests considered in SV.

3 Fixed-b Trend Break Tests

For a given possible break date $T_a \in \Lambda^* := \{\lfloor \tau_L T \rfloor, ..., \lfloor \tau_U T \rfloor\}$, SV recommend a trend break test based on the average, taken across Λ^*, of the Wald-type statistics of the null hypothesis $H_0 : \beta^* = 0$ formed from the levels data and using a long run variance estimator constructed using a Daniell kernel with bandwidth $M = \lfloor bT \rfloor$, $b \in (0, 1]$. To achieve robustness to the possibility of $I(1)$ shocks, this average statistic is then multiplied by a correction factor based on a unit root statistic, following the approach of Vogelsang (1998).

To be more precise, first let $\tau \in \Lambda$ be a generic break fraction such that $T_a = \lfloor \tau T \rfloor$ and let $f_1(t) := (1, t)'$ and $f_2(t, \tau) := DT_t(\tau)$. SV then consider the levels regression

$$y_t = \beta_1' f_1(t) + \beta_\tau^* f_2(t, \tau) + u_t, \tag{2}$$

for a 2×1 vector β_1 and for a scalar β_τ^*. The regression in (2) does of course coincide with (1) in the case where $\tau = \tau_0$, in which case $\beta_1 = (\alpha, \beta)'$ and $\beta_\tau^* = \beta^*$. Now let $\widetilde{f}_2(t, \tau)$ and $\widetilde{y}_t$ denote the residuals from the regression of $f_2(t, \tau)$ and y_t, respectively, on $f_1(t)$, and, hence, by the Frisch-Waugh theorem the OLS estimator of β_τ^* from (2) is given by

$$\widehat{\beta}_\tau^* := \left(\sum_{t=1}^T \widetilde{f}_2(t, \tau) \widetilde{y}_t \right) / \left(\sum_{t=1}^T \widetilde{f}_2(t, \tau)^2 \right).$$

Next let $\widehat{u}_{t,\tau}$ be the OLS residuals from estimating the regression (2), and let the corresponding sample autocovariances be denoted by $\widehat{\gamma}_{j,\tau} := T^{-1} \sum_{t=j+1}^T \widehat{u}_{t,\tau} \widehat{u}_{t-j,\tau}$. Using the estimated autocovariances, form the long-run variance estimator

$$\widehat{\sigma}_\tau^2 := \widehat{\gamma}_{0,\tau} + 2 \sum_{j=1}^{T-1} k(j/M) \widehat{\gamma}_{j,\tau}$$

where $k(j/M)$ is a kernel function and M is a bandwidth parameter such that $M = \lfloor bT \rfloor$, $b \in (0, 1]$ is assumed. SV provide a generic characterisation of $k(j/M)$, but recommend the use of the Daniell kernel which we follow here.

For a given break location, T_a, the resulting Wald-type statistic is then given by

$$W(T_a) := \left(\sum_{t=1}^{T} \widetilde{f}_2(t,\tau)^2 \right) \left(\widehat{\beta}_\tau^* \right)^2 / \widehat{\sigma}_\tau^2. \tag{3}$$

Since the putative break fraction, τ_0, is assumed unknown SV follow Andrews and Ploberger (1994) and take the average of the statistic in (3) across Λ^*; viz,

$$\mathcal{MW} := T^{-1} \sum_{T_a \in \Lambda^*} W(T_a). \tag{4}$$

The limiting null distribution of $\mathcal{MW}$ of (4) depends on whether u_t is $I(0)$ or $I(1)$. Consequently, SV suggest modifying $\mathcal{MW}$ in such a way that for a given nominal asymptotic size the asymptotic critical value for the test coincides in the $I(0)$ and $I(1)$ cases. To that end, for each possible break date $T_a \in \Lambda^*$ SV recommend running the regression

$$y_t = \beta_1' f_1(t) + \beta_\tau^* f_2(t,\tau) + \sum_{i=2}^{9} \pi_i t^i + u_t, \tag{5}$$

to obtain the unrestricted residual sum of squares, RSS_U, and second to estimate (5) under the set of restrictions $\pi_2 = \cdots = \pi_9 = 0$ imposed to yield the restricted residual sum of squares, RSS_R. Then construct the statistic

$$J(T_a) := \frac{RSS_R - RSS_U}{RSS_U}.$$

Now take the minimum of these statistics across Λ^*, viz., $J^* := \inf_{T_a \in \Lambda^*} J(T_a)$. The correction factor recommended by SV for the model in (1), is then given by $\exp(-cJ^*)$, where c is a scaling constant.

Finally then, the SV trend break test rejects for large values of the statistic

$$\mathcal{SV} := \mathcal{MW} \times \exp(-cJ^*). \tag{6}$$

The correction factor is such that $\exp(-cJ^*) \xrightarrow{p} 1$ when u_t is $I(0)$, while $\exp(-cJ^*)$ has a non-degenerate limit distribution (which is free of nuisance parameters) when u_t is $I(1)$. The constant c in the correction factor $\exp(-cJ^*)$ is chosen so that, for a given significance level, the asymptotic null critical value of $\mathcal{SV}$ is the same in the $I(0)$ and $I(1)$ cases for u_t. Practical recommendations regarding the choice of the tuning parameter b, the resulting scaling constant c, and asymptotic critical value needed to implement the test are given in SV.

In the next section we analyse the large sample behaviour of $\mathcal{SV}$ when u_t in (1) is fractionally integrated.

4 Large Sample Behaviour of $\mathcal{SV}$

In this section we derive the large sample properties of $\mathcal{SV}$ of (6) under both a local alternative and under the fixed alternative, $H_1 : \beta^* \neq 0$. In the case of the former, the relevant local alternative is of the form $H_\kappa : \beta^* = \kappa T^{-3/2+\delta}$, where κ is a finite constant (the Pitman drift). Notice that on setting $\kappa = 0$, H_κ reduces to $H_0 : \beta^* = 0$, our null case. Where $\kappa \neq 0$, a local trend break occurs at time $\lfloor \tau_0 T \rfloor$.

In order to characterise the limit distribution of the $\mathcal{SV}$ we need to introduce some notation. First, let $W_{\delta+1}(r)$ denote a Type II fractional Brownian motion, as in Marinucci and Robinson (1999). Then for the fractionally integrated process u_t satisfying Assumptions 1 and 2, we have the following invariance principle from Marinucci and Robinson (1999,2000)

$$\frac{1}{T^{1/2+\delta}} \sum_{t=1}^{\lfloor rT \rfloor} u_t \xrightarrow{d} \sigma W_{\delta+1}(r), \ r \in [0,1]. \tag{7}$$

Now let $F_1(r) := (1, r)'$ and $F_2(r, \tau) := (r - \tau)\mathbb{I}(r > \tau)$, where τ is such that $T_a/T \to \tau$ as $T \to \infty$. Using these we define,

$$\begin{aligned}
\widetilde{F}_2(r,\tau) :=& \quad F_2(r,\tau) - \left(\int_0^1 F_2(r,\tau) F_1(r)' dr\right)\left(\int_0^1 F_1(r) F_1(r)' dr\right)^{-1} F_1(r), \\
P_\delta(\tau) :=& \quad \left(\int_0^1 \widetilde{F}_2(r,\tau)^2 dr\right)^{-1}\left(\int_0^1 \widetilde{F}_2(r,\tau) dW_{\delta+1}(r)\right) \\
Q_\delta(r,\tau) :=& \quad W_{\delta+1}(r) - \int_0^r \widetilde{F}_2(s,\tau) ds P_\delta(\tau) \\
G(r) :=& \quad \left(1, r, ..., r^9\right)', \\
\breve{F}_2(r,\tau) :=& \quad F_2(r,\tau) - \left(\int_0^1 F_2(r,\tau) G(r)' dr\right)\left(\int_0^1 G(r) G(r)' dr\right)^{-1} G(r) \\
\Psi(\tau,\tau_0) :=& \quad \left(\int_0^1 \widetilde{F}_2(r,\tau)^2 dr\right)^{-1} \int_0^1 \widetilde{F}_2(r,\tau)\widetilde{F}_2(r,\tau_0) dr \\
\breve{\Psi}(\tau,\tau_0) :=& \quad \left(\int_0^1 \breve{F}_2(r,\tau)^2 dr\right)^{-1} \int_0^1 \breve{F}_2(r,\tau)\breve{F}_2(r,\tau_0) dr \\
\Xi(r,\tau,\tau_0,\kappa/\sigma) :=& \quad \left\{\left(\int_0^r \widetilde{F}_2(s,\tau) ds\right)\Psi(\tau,\tau_0) - \left(\int_0^r \widetilde{F}_2(s,\tau_0) ds\right)\right\}\frac{\kappa}{\sigma} \\
Q_{\frac{\kappa}{\sigma},\delta}(r,\tau,\tau_0,\kappa/\sigma) :=& \quad \left(Q_\delta(r,\tau) - \Xi(r,\tau,\tau_0,\kappa/\sigma)\right).
\end{aligned}$$

Moreover, let $k(\cdot)$ denote the Daniell Kernel; that is, $k(x) := \frac{\sin(\pi x)}{\pi x}$, and let $k^*(x) := k(x/b)$ and let $k^{*\prime\prime}(\cdot)$ denote the second derivative of this function. Using the latter, we define the function

$$\Phi_{\frac{\kappa}{\sigma},\delta}(b,k,\tau,\tau_0,\kappa/\sigma) := \int_0^1 \int_0^1 -k^{*\prime\prime}(r-s) Q_{\frac{\kappa}{\sigma},\delta}(r,\tau,\tau_0,\kappa/\sigma) Q_{\frac{\kappa}{\sigma},\delta}(s,\tau,\tau_0,\kappa/\sigma) drds.$$

We are now in a position to state the limiting distribution of the $\mathcal{SV}$ statistic of (6) under the local alternative, H_κ.

Theorem 1 *Let Assumptions 1 and 2 hold. Then under the local alternative,* $H_\kappa : \beta^* = \kappa T^{-3/2+\delta}$,

(i) if $\delta \in [0, 1/2)$, *and provided that the spectral density of* η_t *is bounded and bounded away from zero at all frequencies,*

$$\mathcal{SV} \xrightarrow{d} \int_{\tau_L}^{\tau_U} Wald_{\kappa/\sigma,\delta}^{\infty}(\tau, \tau_0, b, k, \kappa/\sigma)\, d\tau,$$

where

$$Wald_{\kappa/\sigma,\delta}^{\infty}(\tau, \tau_0, b, k, \kappa/\sigma) := \frac{\left(P_\delta(\tau) + \Psi(\tau, \tau_0)\frac{\kappa}{\sigma}\right)^2 \left(\int_0^1 \widetilde{F}_2(r,\tau)^2\, dr\right)}{\Phi_{\frac{\kappa}{\sigma},\delta}(b, k, \tau, \tau_0, \kappa/\sigma)}.$$

(ii) if $\delta \in (1/2, 3/2)$ *and provided* $E(|\varepsilon_t|^q) < \infty$, *for some* $q > \max\left(2, \frac{2}{2\delta - 1}\right)$,

$$\mathcal{SV} \xrightarrow{d} \int_{\tau_L}^{\tau_U} Wald_{\kappa/\sigma,\delta}^{\infty}(\tau, \tau_0, b, k, \kappa/\sigma)\, d\tau \times \exp\left(-c \inf_{\tau \in [\tau_L, \tau_U]} J_{\kappa/\sigma,\delta}^{\infty}(\tau, \tau_0)\right),$$

where $Wald_{\kappa/\sigma,\delta}^{\infty}(\tau, \tau_0, b, k, \kappa/\sigma)$ *is as defined in part (i) above, and where*

$$J_{\kappa/\sigma,\delta}^{\infty} := \frac{\int_0^1 \left(\hat{W}_\delta(r,\tau) + \widetilde{F}_2(r,\tau)(1 - \Psi(\tau,\tau_0))\frac{\kappa}{\sigma}\right)^2 dr}{\int_0^1 \left(\breve{W}_\delta(r,\tau) + \breve{F}_2(r,\tau)\left(1 - \breve{\Psi}(\tau,\tau_0)\right)\frac{\kappa}{\sigma}\right)^2 dr} - 1$$

where $\hat{W}_\delta(r,\tau)$ *denotes the residuals from the projection of* $W_\delta(r)$ *onto the space spanned by* $\left(F_1(r)', F_2(r,\tau)\right)'$ *on* $[0,1]$, *and* $\breve{W}_\delta(r,\tau)$ *denotes the residuals from the projection of* $W_\delta(r)$ *onto the space spanned by* $\left(F_1(r)', F_2(r,\tau), r^2, ..., r^9\right)'$ *on* $[0,1]$.

Remark 3. Setting $\kappa = 0$ throughout in the representations given in Theorem 1 yields the limiting distribution of $\mathcal{SV}$ under the null hypothesis $H_0 : \beta^* = 0$. Specifically, we obtain that in part (i) of Theorem 1, where $\delta \in [0, 1/2)$, that under H_0,

$$\mathcal{SV} \xrightarrow{d} \int_{\tau_L}^{\tau_U} Wald_\delta^{\infty}(\tau, b, k)\, d\tau, \tag{8}$$

where

$$Wald_\delta^{\infty}(\tau, b, k) := \frac{P_\delta(\tau)^2 \left(\int_0^1 \widetilde{F}_2(r,\tau)^2\, dr\right)}{\Phi_\delta(b, k, \tau)}$$

$$\Phi_\delta(b, k, \tau) := \int_0^1 \int_0^1 -k^{*\prime\prime}(r - s)\, Q_\delta(r,\tau)\, Q_\delta(s,\tau)\, dr ds.$$

Similarly, in part (ii) of Theorem 1, where $\delta \in (1/2, 3/2)$, we obtain that under H_0,

$$\mathcal{SV} \xrightarrow{d} \int_{\tau_L}^{\tau_U} Wald_{\delta}^{\infty}(\tau, b, k)\, d\tau \times \exp\left(-c \inf_{\tau \in [\tau_L, \tau_U]} J_{\delta}^{\infty}\right), \tag{9}$$

where

$$J_{\delta}^{\infty} := \frac{\int_0^1 \hat{W}_{\delta}(r,\tau)^2\, dr}{\int_0^1 \breve{W}_{\delta}(r,\tau)^2\, dr} - 1.$$

Remark 4. Notice from the representations given in Theorem 1 and Remark 3 that the correction factor $\exp(-cJ^*)$ used by SV plays no role when $\delta \in [0, 1/2)$ since it converges in probability to one. However, when $\delta \in [0, 1/2)$ the limit of the $\mathcal{MW}$ component of $\mathcal{SV}$ still depends on δ. When $\delta \in (1/2, 3/2)$ the limits of both $\exp(-cJ^*)$ and $\mathcal{MW}$ depend on δ. Critical values and choices of c (for a given significance level) made on the assumption that only $\delta = 0$ or $\delta = 1$ is permissible will not therefore deliver the expected asymptotic rejection frequency when δ is fractional. The extent to which the size of the SV test based on $\mathcal{SV}$ varies across fractional values of δ for SV's recommended choices of the the scaling constant, c, and bandwidth tuning parameter, b, is explored below.

Remark 5. It can, however, be shown that $\mathcal{SV}$ is consistent against fixed alternatives, regardless of the value of δ. To do so, we need to establish the behaviour of $\mathcal{SV}$ under $H_1 : \beta^* \neq 0$. For $T_a = \lfloor \tau_0 T \rfloor$ it is easily shown that, under H_1, $\widehat{\beta}^*_{\tau_0} \xrightarrow{p} \beta^*$, and that $\sum_{t=1}^{T} \widetilde{f}_2(t,\tau)^2 = O_e(T^3)$. Consequently, and using the result established in (A.5) in the appendix, coupled with the invariance principle in (7), we then have that $\widehat{\sigma}^2_{\tau_0} = O_e(T^{2\delta})$, and, hence, that $W(T_a) = O_e(T^{3-2\delta})$. Finally, proceeding as in the proof of Theorem 1 in the appendix, we can show that $J(T_a) = O_e(1)$. It therefore follows straightforwardly that $\mathcal{SV} = O_e(T^{3-2\delta})$ under the fixed alternative, H_1, for any $\delta \in [0, 1/2) \cup (1/2, 3/2)$. Consequently, $\mathcal{SV}$ is consistent; that is, under H_1, for any $0 < K < \infty$, $P(\mathcal{SV} > K) \to 1$, as $T \to \infty$. It should also be clear that a test based on the $\mathcal{MW}$ statistic will be consistent against the fixed alternative H_1, regardless of the value of δ.

Remark 6. Parts (i) and (ii) of Theorem 1 both require some additional assumptions to hold above those stated in Assumptions 1 and 2. These assumptions are needed to establish the large sample behaviour of the correction factor, $\exp(-cJ^*)$. The additional moment condition in part (ii) of Theorem 1, which will clearly be a particularly strong assumption when δ is in the proximity of $1/2$, is required to establish the limits $T^{1/2-\delta} u_t \xrightarrow{d} \sigma W_{\delta}$ and $T^{-2\delta} \sum_{t=1}^{T} u_t^2 \xrightarrow{d} \sigma^2 \int_0^1 W_{\delta}^2 dr$ (see Johansen and Nielsen, 2011), as are required for determining the limit of the J^* statistic. However, to anticipate subsequent analysis in section 6 below, we note that, under Assumptions 1 and 2, $\mathcal{MW} \xrightarrow{d} \int_{\tau_L}^{\tau_U} Wald_{\kappa/\sigma,\delta}^{\infty}(\tau, \tau_0, b, k, \kappa/\sigma)\, d\tau$ for $\delta \in [0, 3/2)$. Similarly to Remark 3, under H_0 this simplifies to $\mathcal{MW} \xrightarrow{d} \int_{\tau_L}^{\tau_U} Wald_{\delta}^{\infty}(\tau, b, k)\, d\tau$. It is important to notice, here, that the additional conditions of Theorem 1 are not required to derive the limit distribution of the $\mathcal{MW}$ statistic, and this alone may be a reason to prefer a test based solely on the $\mathcal{MW}$ statistic (i.e. without scaling factor $\exp(-cJ^*)$), especially when δ is allowed to approach 1/2 from above.

4.1 Size of $\mathcal{SV}$

In order to run the trend break test based on the $\mathcal{SV}$ statistic of (6) the bandwidth tuning parameter, b, must be chosen which in turns determines the scaling constant, c and the asymptotic critical value for a given nominal level. The choice of b is discussed in the working paper which accompanies SV, Sayginsoy and Vogelsang (2008). They suggest a data-based procedure where we first take the value of $T_a \in \Lambda^*$ which minimises the residual sum of squares (RSS) obtained from estimating (2) by OLS. Using the corresponding residuals, say $\widehat{u}_t$, they then compute the first-order $AR(1)$ estimator, $\widehat{\alpha} := (\sum_{t=2}^{T} \widehat{u}_{t-1}^2)^{-1} \sum_{t=2}^{T} \widehat{u}_{t-1}\widehat{u}_t$. Then, defining $\widehat{\alpha}^* := T(1-\widehat{\alpha})$, Sayginsoy and Vogelsang (2008) recommend setting the bandwidth as $M = \max(2, b(\widehat{\alpha}^*)T)$, where the value of $b = b(\widehat{\alpha}^*)$ is determined by $\widehat{\alpha}^*$ according to the table below (abstracted from Table 3, page 42 of Sayginsoy and Vogelsang (2008), which is appropriate for $\Lambda = [0.1, 0.9]$:

$\widehat{\alpha}^*$	$[0,16]$	$(16,42]$	$(42,56]$	$(56,66]$	$(0.66,\infty)$
$b(\widehat{\alpha}^*)$	0.30	0.28	0.08	0.06	0.02

Corresponding approximate formulae for the associated scaling constant, c, and asymptotic critical value as twelfth-order polynomials in b are provided in Tables 1.3.1 and 1.3.2 of Sayginsoy and Vogelsang (2008,p.36).

To evaluate the size properties of $\mathcal{SV}$ under fractional integration, we now present the results from a Monte Carlo experiment. We generated data according to the DGP in (1), setting $\alpha = \beta = 0$ with no loss of generality. The fractionally integrated disturbance process in (1) was constructed as $y_t = u_t$ where

$$u_t := \Delta^{-\delta} \{\eta_t \mathbb{I}(t>0)\}, \qquad \eta_t = \rho\eta_{t-1} + \varepsilon_t, \qquad t = 1, ..., T \tag{10}$$

with ε_t distributed as $IIDN(0,1)$, setting $\rho = 0$. Results are reported for $\delta \in \{0.00, 0.33, 0.66, 1.00, 1.33\}$ with sample sizes $T \in \{128, 256, 512, 1024\}$; the largest of which we take as an approximation to the asymptotic case. All experiments are based on 10000 Monte Carlo replications. Here and throughout we set $[\tau_L, \tau_U] = [0.1, 0.9]$. The data-based tuning parameter $b = b(\widehat{\alpha}^*)$, and the associated scaling parameter, c, and asymptotic critical value were selected in the manner recommended by SV, as outlined above, using the nominal asymptotic 0.05-level. Table 1(a) reports the empirical sizes of $\mathcal{SV}$. Consistent with the Monte Carlo results reported in SV, we find that the test has approximately correct size when either $\delta = 0$ or $\delta = 1$, as would be expected. Also, when $\delta > 1$ there is little in the way of size distortion; if anything $\mathcal{SV}$ is modestly conservative here. However, when $\delta \in (0,1)$ the size of $\mathcal{SV}$ appears much larger than the theoretical 0.05 value, indicating that spurious rejections of the no trend break null will arise.

5 δ-Adaptive Critical Values

As demonstrated in the previous section, the asymptotic null distribution of the statistic $\mathcal{SV}$ of (6) depends on the long memory parameter, δ, when the shocks are fractionally integrated, while the asymptotic critical values from SV are only valid where $\delta = 0$ or $\delta = 1$. However, a key point shown by the results in Theorem 1 is that the size of $\mathcal{SV}$ test will not degenerate to zero or unity as the sample size diverges, unlike the HLT statistic, due to the fact that the limiting null distribution of $\mathcal{SV}$ is a well-defined random variable, regardless of the fractional value of δ.

If one knew the value of δ one could therefore substitute the null critical values given in SV with the null critical values appropriate for δ. These are given in Table 2(a) and were calculated using the simulation model $y_t = u_t$ for u_t generated as in (10) with $\rho = 0$, setting $T = 2000$. We provide critical values for $\delta \in \{0, 0.1, ..., 0.4, 0.5, 0.6, ..., 1.4\}$.[1] This, of course, does not yield a feasible procedure because δ is unknown. However, δ can be consistently estimated even without knowledge of the location of the putative trend break; see ILT who propose a consistent estimate of δ based on an extension of the Fully Extended Local Whittle [FELW] approach of Abadir, Di Staso and Giraitis (2007). For completeness, we now briefly outline this approach.[2]

In order to estimate δ, ILT first obtain an estimate of τ_0.They consider the minimum RSS estimator of Perron and Zhu (2005). Specifically, using $\widehat{u}_{t,\tau}$, the OLS residuals from estimating the regression (2), Perron and Zhu (2005) suggest the estimate $\widehat{\tau} := \arg\min_{T_a \in \Lambda^*} \sum_{t=1}^{T} \widehat{u}_{t,\tau}^2$. Under Assumptions 1 and 2, ILT show that $\widehat{\tau} - \tau_0 = O_p\left(T^{-3/2+\delta}\right)$ under the fixed alternative, H_1, and that $\widehat{\tau}$ is of $O_p(1)$ under H_0. The latter result can trivially be shown to also hold under the local alternative H_κ. The corresponding OLS residuals, $\widehat{u}_{t,\widehat{\tau}}$ are then used in the subsequent estimation of δ.

Following Abadir *et al.* (2007), the FELW of δ is then given by

$$\widehat{\delta} := \arg \min_{d \in [0,3/2-\iota]} U\left(I_{\widehat{u}}; d\right)$$

for ι an arbitrarily small positive constant[3], and where

$$U\left(I_{\widehat{u}}; d\right) := \ln\left(\frac{1}{m}\sum_{j=1}^{m} j^{2d} I_{\widehat{u}}\left(\lambda_j, d\right)\right) - \frac{2d}{m}\sum_{j=1}^{m} \ln j$$

[1]If required, critical values for intermediate values of δ can be obtained via linear interpolation. Notice also that our reported critical values for $\delta = 0$ and $\delta = 1$ are relatively close, although not identical, to one another and to the critical values reported in SV; these differences are attributable to simulation error.

[2]Once one has a suitable estimator of δ, an alternative to the approach taken in this section is to use a bootstrap implementation of the $\mathcal{SV}$ test. We also considered such an approach but found it to deliver slightly inferior finite sample performance to the numerically simpler δ-adaptive method discussed here. Further details are available from the authors on request.

[3]In practice the minimization is carried out over $d \in [0, 3/2]$. The numerical results reported in this paper used the Maxlik optimisation routine in GAUSS 6.0.

for $\lambda_j = \frac{2\pi j}{T-1}$, and with $I_{\widehat{u}}(\lambda, d) := |w_{\widehat{u}}(\lambda, d)|^2$ denoting the extended periodogram of $\widehat{u}_{t,\widehat{\tau}}$, where

$$w_{\widehat{u}}(\lambda, d) := \frac{1}{\sqrt{2\pi(T-1)}} \sum_{t=2}^{T} \widehat{u}_{t,\widehat{\tau}}\, e^{i\lambda(t-1)} + \left(\frac{1}{\sqrt{2\pi(T-1)}} e^{i\lambda} \left(\widehat{u}_{T,\widehat{\tau}} - \widehat{u}_{1,\widehat{\tau}}\right) \times \mathbb{I}\left(d \in (1/2, 3/2)\right) \right)$$

with complex conjugate $w^*_{\widehat{u}}(\lambda, d)$, is the extended discrete Fourier transform of $\widehat{u}_{t,\widehat{\tau}}$, $t = 1, ..., T$, at a generic frequency λ.

In order to discuss the properties of $\widehat{\delta}$, let $f_\eta(\lambda)$ be the spectral density of η_t, so that $E(\eta_t \eta_{t+h}) = \int_{-\pi}^{\pi} \cos(h\lambda) f_\eta(\lambda)\, d\lambda$. ILT demonstrate that under Assumptions 1 and 2, strengthened to also require that $f_\eta(\lambda)$ is such that $f_\eta(\lambda) \sim G \in (0, \infty)$ as $\lambda \to 0^+$ and is differentiable in the neighbourhood $(0, \omega)$ of the origin, with $\frac{df_\eta(\lambda)}{d\lambda} = O(\lambda^{-1})$ as $\lambda \to 0^+$, then provided that the bandwidth m is such that $m \to \infty$ but $m/T \to 0$ as $T \to \infty$, $\widehat{\delta}$ is a consistent estimator of δ, regardless of the value of the trend break magnitude β^*. When constructing $\widehat{\delta}$ we adopt a bandwidth of $m = \lfloor T^{0.65} \rfloor$ in what follows.

The estimator $\widehat{\delta}$ obtained in this manner is consistent for δ under H_0, H_κ and H_1. We can also consider the FELW estimator based on the OLS residuals, $\widehat{u}_{t,0}$ say, obtained from estimating the regression (2) under H_0; that is, *without* the trend break regressor $f_2(t, \tau)$. This estimator, $\widehat{\delta}_0$ say, is obviously rather simpler to compute than $\widehat{\delta}$ as no trend break point needs to be searched for. It should be clear that $\widehat{\delta}_0$ and $\widehat{\delta}$ share the same consistency properties under H_0, and it can be shown that this is also the case under H_κ. Given the consistency properties of $\widehat{\delta}$ and $\widehat{\delta}_0$ under H_0, we can use them to obtain asymptotically correctly sized inference in an obvious way; we compare $\mathcal{SV}$ not with the critical values given in SV, but rather with critical values from Table 2(a) , using $\widehat{\delta}$ or $\widehat{\delta}_0$ to proxy the unknown δ. It should also be clear from Remark 5 that this test procedure will be consistent against the fixed alternative H_1, regardless of the value of δ or which of $\widehat{\delta}$ or $\widehat{\delta}_0$ is used.

5.1 Size of $\mathcal{SV}$ with δ-Adaptive Critical Values

Table 1(a) gives the size of $\mathcal{SV}$ with δ-adaptive critical values, with $\widehat{\delta}$ and $\widehat{\delta}_0$ used in place of the unknown δ, respectively, using the simulation DGP of section 4.1. Results are shown for nominal 0.05-level tests using the critical values from Table 2(a). With an obvious notation, we denote these test procedures as $\mathcal{SV}^A$ and $\mathcal{SV}_0^A$. It is clear that both schemes lead to a significant improvement in size over the results seen in Table 1(a) for the non-adaptive SV test, particularly for the larger sample sizes and in those cases where $\delta \in (0, 1)$. Between $\mathcal{SV}^A$ and $\mathcal{SV}_0^A$, there seems little to choose as to which gives most accurate sizes, although we observe that the test based on $\widehat{\delta}_0$, $\mathcal{SV}_0^A$, appears less prone to over-sizing.

6 A Simplified Variant of $\mathcal{SV}$ with δ-Adaptive Critical Values

Once we employ δ-adaptive critical values for $\mathcal{SV}$ we can consider a simplification of the SV test which involves dropping the correction factor $\exp(-cJ^*)$. As noted in Remark 4, this term plays no asymptotic role when $\delta \in [0, 1/2)$ since it converges in probability to one. It is active for $\delta \in (1/2, 3/2)$ and used to line up the asymptotic critical values (only when $\delta = 1$) but clearly becomes redundant once we use δ-adaptive critical values. We can therefore simply use δ-adaptive critical values applied to the $\mathcal{MW}$ component of $\mathcal{SV}$ alone. Moreover, rather than just being redundant when $\delta \in (1/2, 3/2)$, the presence of $\exp(-cJ^*)$ may potentially lead to a sacrifice in (otherwise) realisable power and, additionally, requires us to make stronger assumptions about the errors; cf. Remark 6.

When we consider fractional values of δ there is little reason to choose b according to $b(\widehat{\alpha}^*)$ since $\widehat{\alpha}^*$ is only an appropriate estimator in the case of (near) $I(1)$ errors. In view of this, Table 2(b) gives the δ-adaptive critical values for $\mathcal{MW}$, computed from the same DGP as in section 5, for a range of values of fixed b. For $\delta \in [0, 1/2)$ since $\exp(-cJ^*) \xrightarrow{p} 1$, such that $\mathcal{MW} - \mathcal{SV} \xrightarrow{p} 0$ we would expect the critical values in Table 2(a) and 2(b) to coincide for the common value of $b = 0.30$. They largely do so unless δ is close to $1/2$, in which case our asymptotic approximation using $T = 2000$ is obviously not based on a large enough sample for these near-boundary cases. What is very noticeable, however, is that for $\delta \in (1/2, 3/2)$ the critical values of $\mathcal{MW}$ and $\mathcal{SV}$ with $b = 0.30$ differ substantially, due to the asymptotic contribution of $\exp(-cJ^*)$ in the latter.

As in section 5 we can make testing with δ-adaptive critical values using $\mathcal{MW}$ a feasible procedure by estimating δ using either $\widehat{\delta}$ or $\widehat{\delta}_0$. We will denote these two procedures as $\mathcal{MW}^A$ and $\mathcal{MW}_0^A$, respectively.[4]

6.1 Size of $\mathcal{MW}$ with δ-Adaptive Critical Values

Table 1(a) shows the size of $\mathcal{MW}^A$ and $\mathcal{MW}_0^A$ for the simulation DGP of section 4.1. Results are shown for nominal 0.05-level tests using critical values taken from Table 2(b). Clearly $\mathcal{MW}^A$ can be quite badly over-sized when $\delta > 0$ for small T unless $b = 0.20$ or $b = 0.30$ (although its size does improve fairly rapidly as T increases). On the other hand, the size of $\mathcal{MW}_0^A$ is actually rather good for all δ, T and b.

To check all the tests' size behaviour outside of white noise η_t, Table 1 (b) repeats this simulation exercise when $\rho = 0.5$ in (10). We see here that pretty much all the tests except $\mathcal{SV}$ display decent size control, in the context of not being significantly over-sized. In Table 1(c) we set $\rho = -0.5$. Of the SV tests, only $\mathcal{SV}_0^A$ now adequately

[4]Our analysis focuses on the 'Mean Wald' statistic of (4). We also experimented with the 'Sup Wald' and 'Mean Exp Wald' variants, the former also considered in SV, but found no substantive differences between the three.

controls size, although $\mathcal{SV}^A$ is not too inferior in this respect. Size control is also decent for $\mathcal{MW}^A$ with $b = 0.20$ and $b = 0.30$, and all of the $\mathcal{MW}_0^A$ tests, particularly for $b = 0.05$ and above.

As a final check on size properties outside of the fractional case, we consider local-to-unit root errors. Specifically, the simulation DGP is now (10) with $\delta = 0$ and $\rho = 1 - c^*/T$ where $c^* > 0$. The results are shown in Table 1(d). Here we see that over-sizing does not arise as an issue for any of the tests; rather it is under-sizing that appears rather more prevalent, particularly for the $\mathcal{MW}^A$ and $\mathcal{MW}_0^A$ tests.

7 Local Power

The results in the previous section would suggest that both the $\mathcal{SV}^A$ and $\mathcal{SV}_0^A$ tests, the $\mathcal{MW}^A$ with $b = 0.20$ and $b = 0.30$, and all of the $\mathcal{MW}_0^A$ tests appear to display decent finite sample size control. Choosing between these tests therefore, in terms of recommending a test to use in practice, effectively boils down to a comparison of their finite sample power performances. Because of their dependence on b, we begin by comparing the seven $\mathcal{MW}$-based tests alone. To that end, we simulate data from the DGP

$$y_t = \kappa T^{-3/2+\delta} DT_t(\tau_0) + u_t \tag{11}$$

for $t = 1, ..., T$, using (10) with $\rho = 0$ to generate u_t, for a sample size of $T = 256$, and consider the case where a break occurs at the sample midpoint by setting $\tau_0 = 0.5$.[5] For the local break magnitude κ we use $\kappa \in \{0, 1, 2, ...100\}$, noting that $\kappa = 0$ represents H_0 rather than H_κ. We plot the local powers as functions of κ and these are given in Figure 1(a)-(e), for each of the five values of δ considered (the plotted values for $\kappa = 0$ coincide with the sizes given in Table 1). It is fairly unequivocal from Figure 1 that the two best performing tests are $\mathcal{MW}_0^A$ with $b = 0.025$ and $b = 0.05$. There is very little difference in the powers of these two tests in any of the plots, and certainly no systematic dominance by either one. The next best performing test after these two is $\mathcal{MW}_0^A$ with $b = 0.10$, though the power gap is easily discernable. Any of the tests using $b = 0.20$ or $b = 0.30$ fall well behind these three. Taking into account the size properties of the $\mathcal{MW}$-based tests elicited in the previous section, our preference among these tests would therefore be for $\mathcal{MW}_0^A$ with $b = 0.05$.

We next compare the performance of our preferred $\mathcal{MW}$-based test ($\mathcal{MW}_0^A$ with $b = 0.05$, simply referred to as $\mathcal{MW}_0^A$ in what follows) with $\mathcal{SV}^A$ and $\mathcal{SV}_0^A$, using the same set of simulation DGPs as outlined in the last paragraph. These results are shown in Figure 2(a)-(e). We also include $\mathcal{SV}$ as a point of comparison (notwithstanding the fact that it is over-sized when $\delta = 0.33, 0.66$). We observe that there is not a great deal of difference in power between $\mathcal{MW}_0^A$, $\mathcal{SV}^A$ and $\mathcal{SV}_0^A$ when $\delta = 0$ or $\delta = 0.33$. Partly, this might be expected by the fact that the multiplicative correction factor, $\exp(-cJ^*)$, is not (asymptotically) operative for the latter two tests under these values of δ. The

[5] Our findings are qualitatively unchanged for other choices of the sample size and break location; details are available upon request.

relatively high power seen for $\mathcal{SV}$ when $\delta = 0.33$ is simply an artefact of its over-size here. For $\delta = 0.66, 1.0$ and 1.33, we find that $\mathcal{MW}_0^A$ is (outside of some small values of κ) by far the most powerful test of the four. This provides very clear empirical evidence supporting our conjecture in section 6 that, once δ-adaptive critical values are being used, and when $\delta \in (1/2, 3/2)$, employing the multiplicative correction factor $\exp(-cJ^*)$ is not a benign measure, but actually has a highly inimical effect on power.

Finally, the powers of these same four tests are considered when the errors are local to a unit root; i.e., when u_t is generated as in (10) with $\delta = 0$ and $\rho = 1 - c^*/T$. Accordingly, we now set $\delta = 1$ in (11). Figure 3(a)-(c) shows these results. Here $\mathcal{MW}_0^A$ emerges as clearly the most powerful test. Again, the reason for this is that the correction factor is compromising the powers of the three $\mathcal{SV}$-based tests.

Taking both the size and power performance of all our candidate test procedures across the range of values of δ into consideration, it is $\mathcal{MW}_0^A$ with $b = 0.05$ that emerges as the best performing procedure. Its size is always more than acceptable and, among the $\mathcal{MW}$-based tests we have considered, ranks as arguably the most powerful. Compared to the $\mathcal{SV}$-based tests, when $\delta \in (1/2, 3/2)$, power considerations alone are sufficient to justify its recommendation.

8 Conclusions

Sayginsoy and Vogelsang (2011) [SV] proposed trend break tests which are (asymptotically) robust as to whether the shocks are $I(0)$ or $I(1)$. In this paper we have demonstrated that the SV tests are, however, not robust to the order of integration δ of the shocks when this is fractional: their limiting null distributions depending on the long memory parameter. Crucially, however, and in contrast to other available $I(0)$-$I(1)$ robust trend break tests, we have shown that these limiting null distributions are well-defined, regardless of the value of the long memory parameter. As a consequence we have shown that implementing the SV tests with δ-adaptive critical values (these being based on estimating the long memory parameter) yields tests which are (asymptotically) size-robust to the long memory parameter without requiring knowledge of that parameter. The use of δ-adaptive critical values additionally allowed us to consider simplified variants of the SV test which omit the multiplicative correction factor present in the original specification. In the form of $\mathcal{MW}_0^A$, the simplified statistic with δ estimated under the no trend break null hypothesis, this approach yielded a statistic with very good finite sample size control and power levels across δ when implemented using a bandwidth of $b = 0.05$.

References

Abadir, K.M., W. Di Staso and L. Giraitis, 2007. Nonstationarity-extended local Whittle estimation, *Journal of Econometrics* 141, 1353–1384.

Bunzel, H. and T.J. Vogelsang, 2005. Powerful trend function tests that are robust to strong serial correlation with an application to the Prebisch-Singer hypothesis. *Journal of Business and Economic Statistics* 23, 381-394.

Hannan, E.J., 1970. *Multiple Time Series*, Wiley, New York.

Harris, D., D. Harvey, S.J. Leybourne and A.M.R. Taylor, 2009. Testing for a unit root in the presence of a possible break in trend, *Econometric Theory* 25, 1545-1588.

Harvey, D.I., S.J. Leybourne and A.M.R. Taylor, 2009. Simple, robust and powerful tests of the breaking trend hypothesis, *Econometric Theory* 25, 995–1029.

Iacone, F., S.J. Leybourne and A.M.R. Taylor, 2011. Testing for a break in the trend when the order of integration is unknown, *Mimeo, University of Nottingham.*

Johansen, S., and M. Nielsen, 2011. A necessary moment condition for the fractional functional central limit theorem, forthcoming in *Econometric Theory.*

Kiefer, N.M., and T.J. Vogelsang, 2002. Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation, *Econometrica* 70, 2093-2095.

Marinucci, D., and P.M. Robinson, 1999. Alternative forms of fractional Brownian motion, *Journal of Statistical Planning and Inference* 80, 111-122.

Marinucci, D., and P.M. Robinson, 2000. Weak convergence of multivariate fractional processes, *Stochastic Processes and their Applications* 86, 103-120.

Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis, *Econometrica* 57, 1361-1401.

Perron, P., 2005. Dealing with structural breaks, *Palgrave Handbook of Econometrics, Vol. 1: Econometric Theory.*

Perron, P., and X. Zhu, 2005. Structural breaks with deterministic and stochastic trends, *Journal of Econometrics* 129, 65–119.

Robinson, P.M., and F. Iacone, 2005. Cointegration in fractional systems with deterministic trends, *Journal of Econometrics* 129, 263–298.

Sayginsoy, Ö., and T. Vogelsang, 2008. Testing for a shift in trend at an unknown date: a fixed-b analysis of heteroskedasticity autocorrelation robust OLS based tests, *Working paper*.

Sayginsoy, Ö., and T. Vogelsang, 2011. Testing for a shift in trend at an unknown date: a fixed-b analysis of heteroskedasticity autocorrelation robust OLS based tests, *Econometric Theory*, forthcoming.

Sibbertsen, P., 2004. Long memory versus structural breaks: and overview, *Statistical Papers* 45, 465-515.

Vogelsang, T.J., 1998. Testing for a shift in mean without having to estimate serial correlation parameters, *Journal of Business and Economic Statistics* 16, 73-80.

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.

Appendix

Proof of Theorem 1

We initially discuss the case where $\kappa = 0$. For any T_a, recall that $\widehat{\alpha}$, $\widehat{\beta}$ and $\widehat{\beta}_\tau^*$ are used to denote the OLS estimates of α, β and β^*, respectively, from (2). On scaling we therefore have (where we have suppressed the dependence of $DT_t(\tau)$ on τ purely for notational convenience) that,

$$T^{-\delta}\left[\begin{array}{ccc} T^{-1/2} & 0 & 0 \\ 0 & T^{-3/2} & 0 \\ 0 & 0 & T^{-3/2} \end{array}\right]\left(\begin{array}{c} \widehat{\alpha}-\alpha \\ \widehat{\beta}-\beta \\ \widehat{\beta}_\tau^*-\beta^* \end{array}\right)$$

$$= \left(\left[\begin{array}{ccc} T^{-1/2} & 0 & 0 \\ 0 & T^{-3/2} & 0 \\ 0 & 0 & T^{-3/2} \end{array}\right]\left(\sum_{t=1}^{T}\left[\begin{array}{ccc} 1 & t & DT_t \\ t & t^2 & t\times DT_t \\ DT_t & t\times DT_t & DT_t^2 \end{array}\right]\right)\left[\begin{array}{ccc} T^{-1/2} & 0 & 0 \\ 0 & T^{-3/2} & 0 \\ 0 & 0 & T^{-3/2} \end{array}\right]\right)^{-1}$$

$$\times T^{-\delta}\left[\begin{array}{ccc} T^{-1/2} & 0 & 0 \\ 0 & T^{-3/2} & 0 \\ 0 & 0 & T^{-3/2} \end{array}\right]\left(\begin{array}{c} \sum_{t=1}^{T} u_t \\ \sum_{t=1}^{T} tu_t \\ \sum_{t=1}^{T} DT_t u_t \end{array}\right) =: A\times B$$

where A and B are implicitly defined. In the case of A, entirely standard results establish that

$$A\rightarrow\left(\begin{array}{ccc} \int_0^1 1dr & \int_0^1 rdr & \int_\tau^1 (1-\tau)\, rdr \\ \int_0^1 rdr & \int_0^1 r^2dr & \int_\tau^1 r^2(1-\tau)\, dr \\ \int_\tau^1 (1-\tau)\, rdr & \int_\tau^1 r^2(1-\tau)\, dr & \int_\tau^1 r^2(1-\tau)^2\, dr \end{array}\right)^{-1}.$$

Turning to B, we may characterise the limits of each summation using the invariance principle in (7). First,

$$T^{-1/2-\delta}\sum_{t=1}^{T} u_t \xrightarrow{d} \sigma W_{\delta+1}(1)$$

and, by (A.1) of Robinson and Iacone (2005),

$$T^{-3/2-\delta}\sum_{t=1}^{T} tu_t \xrightarrow{d} \sigma\int_0^1 sdW_{\delta+1}(s).$$

By the same reference, it also follows that, pointwise in τ

$$T^{-3/2-\delta}\sum_{t=\lfloor\tau T\rfloor+1}^{T}\left(t-\lfloor\tau T\rfloor\right)u_t \xrightarrow{d} \sigma\int_\tau^1 (s-\tau)\, dW_{\delta+1}(s). \qquad \text{(A.1)}$$

To verify that this convergence also holds in the Skorohod topology for the functional of τ, we rewrite

$$T^{-3/2-\delta}\sum_{t=\lfloor\tau T\rfloor+1}^{T}\left(t-\lfloor\tau T\rfloor\right)u_t = T^{-3/2-\delta}\sum_{t=1}^{T} tu_t + T^{-3/2-\delta}\sum_{t=1}^{\lfloor\tau T\rfloor-1}\sum_{s=1}^{t} u_s - T^{-1}\lfloor\tau T\rfloor T^{-1/2-\delta}\sum_{t=1}^{T} u_t. \qquad \text{(A.2)}$$

The first and third summation on the right hand side of (A.2) do not depend on τ, and of course $T^{-1}\lfloor \tau T \rfloor \to \tau$ uniformly, so we are only concerned with the convergence in the Skorohod topology of the second term. Letting $U_t := \sum_{s=1}^{t} u_s$, then U_t is an $I(\delta+1)$ process, and, by (7), we have the functional convergence

$$T^{-3/2-\delta} \sum_{t=1}^{\lfloor \tau T \rfloor -1} \sum_{s=1}^{t} u_s \xrightarrow{d} \sigma W_{\delta+2}(\tau).$$

The convergence of the functional of τ as in (A.1) then follows by the Cramer Wold device, and proceeding as in Marinucci and Robinson (2000).

Collecting these results together, it is then straightforward to establish that

$$T^{3/2-\delta}\left(\widehat{\beta}_{\tau}^{*} - \beta^{*}\right) \xrightarrow{d} \sigma P_{\delta}(\tau)$$

Now consider $\widehat{u}_{t,\tau}$, the OLS residual from estimating (2), and define

$$\widehat{S}_j := \sum_{t=1}^{j} \widehat{u}_{t,\tau}.$$

with $P_{\delta}(\tau)$ defined as in the main text. Following the approach taken in SV, using the familiar Frisch-Waugh Theorem, we may use the foregoing results and the continuous mapping theorem [CMT] to establish that

$$T^{-1/2-\delta}\widehat{S}_{\lfloor Tr \rfloor} \xrightarrow{d} \sigma Q_{\delta}(r,\tau) \tag{A.3}$$

where $Q_{\delta}(r,\tau)$ is as defined in the main text. Next, and following Bunzel and Vogelsang (2005,p.392), let

$$\Delta^2 \kappa_{ij} := \left\{ k\left(\frac{i-j}{\lfloor bT \rfloor}\right) - k\left(\frac{i-j-1}{\lfloor bT \rfloor}\right)\right\} - \left\{ k\left(\frac{i-j+1}{\lfloor bT \rfloor}\right) - k\left(\frac{i-j}{\lfloor bT \rfloor}\right)\right\}. \tag{A.4}$$

Using (A.4), we then re-write the scaled long run variance estimator as

$$T^{-2\delta}\widehat{\sigma}^2 = -T^{-1}\sum_{i=1}^{T-1} T^{-1}\sum_{j=1}^{T-1} T^2 \Delta^2 \kappa_{ij}\left(T^{-1/2-\delta}\widehat{S}_i\right)\left(T^{-1/2-\delta}\widehat{S}_j\right) \tag{A.5}$$

from which it follows straightforwardly using (A.3) and the CMT that

$$T^{-2\delta}\widehat{\sigma}^2 \xrightarrow{d} \sigma^2 \Phi_{\delta}(b,k,\tau) \tag{A.6}$$

where $\Phi_{\delta}(b,k,\tau)$ is as defined in the main text. Finally, we therefore have from (A.3) and (A.6) and the CMT that

$$W(T_a) \xrightarrow{d} Wald_{\delta}^{\infty}(\tau,b,k)$$

and, consequently, that

$$\mathcal{MW} \xrightarrow{d} \int_{\tau_L}^{\tau_U} Wald_{\delta}^{\infty}(\tau, b, k)\, d\tau$$

with this result holding for either $\delta \in [0, 1/2)$ or $\delta \in (1/2, 3/2)$.

In order to complete the proof we need to establish the behaviour of the scaling term, $\exp(-cJ^*)$. To that end, consider first part (i) where $\delta \in [0, 1/2)$, and let

$$\upsilon_t := \Delta^{-\delta}\eta_t,$$

so that υ_t is the Type I fractionally integrated version of the process u_t, and notice that, by a law of large numbers for stationary ergodic processes (see Hannan, 1970, pp. 203-204 for a discussion of sufficient conditions for ergodicity, but note that these are met under the conditions of part (i) of Theorem 1), then $T^{-1}\sum_{t=1}^{T}\upsilon_t^2 \xrightarrow{p} E(\upsilon_t^2)$. Rewriting

$$T^{-1}\sum_{t=1}^{T} u_t^2 = T^{-1}\sum_{t=1}^{T}(u_t - \upsilon_t)^2 + 2T^{-1}\sum_{t=1}^{T}(u_t - \upsilon_t)\upsilon_t + T^{-1}\sum_{t=1}^{T}\upsilon_t^2,$$

since $E(\upsilon_t - u_t)^2 = O(t^{2\delta-1})$, as shown in Marinucci, and Robinson (1999,p.119), then $T^{-1}\sum_{t=1}^{T}(u_t - \upsilon_t)^2 \xrightarrow{p} 0$ and therefore $T^{-1}\sum_{t=1}^{T}(u_t - \upsilon_t)\upsilon_t \xrightarrow{p} 0$ by the Cauchy-Schwarz inequality. Therefore, $T^{-1}\sum_{t=1}^{T} u_t^2 \xrightarrow{p} E(\upsilon_t^2)$ also follows. Using the rates of convergence of $\widehat{\alpha} - \alpha$, $\widehat{\beta} - \beta$, $\widehat{\beta}_{\tau}^* - \beta^*$, $\widehat{u}_{t,\tau} = u_t + O(t^{-1/2+\delta})$, so $T^{-1}\sum_{t=1}^{T}\widehat{u}_{t,\tau}{}^2 \xrightarrow{p} E(\upsilon_t^2)$ also follows, i.e. $T^{-1}RSS_R \xrightarrow{p} E(\upsilon_t^2)$. In the same way it is possible to show $T^{-1}RSS_U \xrightarrow{p} E(\upsilon_t^2)$. Under H_0 these results are valid for any T_b (and notice that under H_1 they are at least valid for T_b such that $T_b/T \to \tau_0$). Since it is trivially true that $E(\upsilon_t^2) > 0$, $J(T_a) \geq 0$ and $\inf_{T_a \in \Lambda^*}(J(T_a)) \leq J(T_b)$, then $\inf_{T_a \in \Lambda^*}(J(T_a)) \xrightarrow{p} 0$. The result in part (i) then follows by the CMT.

To complete the proof of part (ii), use (7) and the CMT to establish the results that $T^{-2\delta}RSS_U \xrightarrow{d} \int_0^1 \hat{W}_{\delta}(r, \tau)^2\, dr$ and $T^{-2\delta}RSS_R \xrightarrow{d} \int_0^1 \check{W}_{\delta}(r, \tau)^2\, dr$. It then follows using the CMT that $J(T_a) \xrightarrow{d} J_{\infty}(\tau)$ and finally, again using the CMT, that $J^* \xrightarrow{d} \inf_{\tau \in [\tau_L, \tau_U]} J_{\infty}(\tau)$. The result in part (ii) then follows by the CMT.

To complete the proof, the additional terms arising from the local trend break magnitude term, $\beta^* = \kappa T^{-3/2+\delta}$, are accounted for in similar fashion to Theorem 2 of SV. We therefore omit these details in the interests of brevity.

Table 1 (a). Null rejection frequencies of tests at the 0.05 level. $\rho = 0$.

		$\mathcal{SV}$	$\mathcal{SV}^A$	$\mathcal{SV}_0^A$	$\mathcal{MW}^A$					$\mathcal{MW}_0^A$				
	b				0.025	0.050	0.100	0.200	0.300	0.025	0.050	0.100	0.200	0.300
δ	T													
0	128	0.019	0.019	0.009	0.043	0.046	0.043	0.041	0.039	0.023	0.036	0.037	0.039	0.036
	256	0.030	0.030	0.024	0.039	0.043	0.046	0.043	0.046	0.030	0.036	0.044	0.043	0.046
	512	0.040	0.039	0.026	0.043	0.044	0.043	0.045	0.045	0.030	0.041	0.040	0.045	0.044
	1024	0.043	0.043	0.037	0.043	0.041	0.043	0.042	0.052	0.038	0.040	0.042	0.041	0.051
0.33	128	0.100	0.069	0.019	0.149	0.118	0.079	0.037	0.047	0.050	0.052	0.044	0.029	0.032
	256	0.143	0.090	0.041	0.138	0.119	0.086	0.064	0.064	0.068	0.069	0.061	0.057	0.056
	512	0.179	0.067	0.039	0.099	0.079	0.068	0.062	0.065	0.049	0.052	0.053	0.055	0.055
	1024	0.199	0.060	0.045	0.079	0.065	0.051	0.044	0.048	0.050	0.042	0.045	0.043	0.045
0.66	128	0.082	0.046	0.036	0.213	0.173	0.114	0.059	0.064	0.066	0.062	0.049	0.043	0.046
	256	0.127	0.043	0.040	0.159	0.137	0.104	0.082	0.070	0.082	0.081	0.069	0.064	0.061
	512	0.137	0.038	0.040	0.121	0.097	0.085	0.061	0.050	0.059	0.059	0.050	0.050	0.044
	1024	0.149	0.039	0.042	0.079	0.072	0.058	0.055	0.054	0.056	0.056	0.049	0.047	0.053
1	128	0.045	0.036	0.045	0.195	0.175	0.137	0.076	0.061	0.067	0.063	0.058	0.049	0.043
	256	0.056	0.044	0.051	0.139	0.122	0.102	0.076	0.071	0.074	0.072	0.060	0.058	0.058
	512	0.062	0.055	0.057	0.115	0.100	0.093	0.073	0.062	0.071	0.068	0.065	0.062	0.057
	1024	0.047	0.041	0.046	0.084	0.076	0.066	0.049	0.050	0.060	0.055	0.047	0.042	0.042
1.33	128	0.031	0.037	0.049	0.158	0.149	0.120	0.073	0.065	0.063	0.061	0.052	0.046	0.044
	256	0.038	0.047	0.055	0.119	0.110	0.092	0.074	0.069	0.058	0.057	0.050	0.063	0.054
	512	0.033	0.048	0.056	0.090	0.080	0.079	0.057	0.057	0.061	0.056	0.053	0.049	0.043
	1024	0.041	0.058	0.063	0.078	0.072	0.064	0.064	0.058	0.063	0.058	0.055	0.055	0.054

Table 1 (b). Null rejection frequencies of tests at the 0.05 level. $\rho = 0.5$.

		$\mathcal{SV}$	$\mathcal{SV}^A$	$\mathcal{SV}_0^A$	$\mathcal{MW}^A$					$\mathcal{MW}_0^A$				
	b				0.025	0.050	0.100	0.200	0.300	0.025	0.050	0.100	0.200	0.300
δ	T													
0	128	0.026	0.020	0.005	0.054	0.041	0.034	0.038	0.037	0.013	0.019	0.023	0.033	0.033
	256	0.030	0.022	0.008	0.041	0.036	0.043	0.041	0.044	0.017	0.023	0.031	0.038	0.041
	512	0.029	0.022	0.011	0.034	0.038	0.038	0.048	0.047	0.016	0.026	0.028	0.045	0.044
	1024	0.037	0.034	0.020	0.038	0.038	0.044	0.042	0.047	0.024	0.029	0.040	0.040	0.045
0.33	128	0.031	0.028	0.016	0.097	0.065	0.046	0.029	0.039	0.019	0.020	0.025	0.025	0.028
	256	0.076	0.021	0.010	0.084	0.073	0.059	0.054	0.051	0.024	0.030	0.037	0.041	0.040
	512	0.100	0.021	0.015	0.051	0.049	0.046	0.050	0.051	0.020	0.025	0.029	0.047	0.043
	1024	0.132	0.029	0.024	0.046	0.043	0.045	0.040	0.043	0.033	0.033	0.043	0.040	0.041
0.66	128	0.036	0.028	0.031	0.099	0.081	0.059	0.049	0.048	0.025	0.024	0.025	0.032	0.034
	256	0.059	0.046	0.046	0.083	0.077	0.067	0.061	0.057	0.033	0.038	0.038	0.039	0.052
	512	0.060	0.038	0.038	0.056	0.057	0.047	0.047	0.036	0.030	0.031	0.033	0.042	0.030
	1024	0.078	0.036	0.038	0.054	0.052	0.045	0.042	0.049	0.036	0.035	0.038	0.034	0.048
1	128	0.030	0.030	0.043	0.083	0.077	0.061	0.046	0.044	0.020	0.020	0.024	0.023	0.030
	256	0.037	0.039	0.042	0.071	0.068	0.059	0.054	0.053	0.038	0.038	0.035	0.042	0.046
	512	0.056	0.056	0.060	0.071	0.068	0.061	0.059	0.049	0.041	0.041	0.040	0.053	0.043
	1024	0.040	0.041	0.041	0.058	0.054	0.048	0.041	0.042	0.038	0.037	0.036	0.034	0.037
1.33	128	0.021	0.036	0.041	0.079	0.072	0.060	0.044	0.041	0.043	0.042	0.034	0.036	0.034
	256	0.033	0.053	0.057	0.062	0.059	0.053	0.057	0.049	0.040	0.041	0.035	0.052	0.042
	512	0.031	0.054	0.060	0.063	0.057	0.054	0.048	0.039	0.047	0.043	0.046	0.044	0.037
	1024	0.041	0.065	0.065	0.063	0.059	0.056	0.053	0.054	0.049	0.045	0.046	0.051	0.048

Table 1 (c). Null rejection frequencies of tests at the 0.05 level. $\rho = -0.5$.

		$\mathcal{SV}$	$\mathcal{SV}^A$	$\mathcal{SV}_0^A$	$\mathcal{MW}^A$					$\mathcal{MW}_0^A$				
	b				0.025	0.050	0.100	0.200	0.300	0.025	0.050	0.100	0.200	0.300
δ	T													
0	128	0.017	0.017	0.007	0.038	0.048	0.048	0.042	0.037	0.022	0.038	0.044	0.039	0.035
	256	0.031	0.032	0.028	0.035	0.040	0.049	0.049	0.044	0.031	0.038	0.048	0.048	0.044
	512	0.042	0.042	0.031	0.042	0.043	0.044	0.043	0.048	0.030	0.039	0.042	0.043	0.047
	1024	0.042	0.042	0.038	0.043	0.040	0.044	0.042	0.054	0.040	0.040	0.042	0.041	0.054
0.33	128	0.167	0.120	0.045	0.170	0.131	0.085	0.041	0.045	0.056	0.059	0.054	0.034	0.031
	256	0.229	0.134	0.076	0.150	0.125	0.094	0.065	0.064	0.073	0.076	0.062	0.059	0.055
	512	0.256	0.098	0.066	0.108	0.087	0.074	0.060	0.067	0.054	0.054	0.053	0.056	0.056
	1024	0.254	0.088	0.071	0.081	0.069	0.051	0.043	0.050	0.058	0.046	0.046	0.041	0.047
0.66	128	0.245	0.113	0.078	0.243	0.201	0.129	0.067	0.072	0.097	0.081	0.061	0.049	0.053
	256	0.232	0.087	0.083	0.177	0.150	0.108	0.083	0.074	0.092	0.084	0.071	0.066	0.066
	512	0.221	0.079	0.078	0.132	0.106	0.089	0.061	0.052	0.072	0.066	0.058	0.052	0.048
	1024	0.221	0.064	0.066	0.086	0.074	0.061	0.055	0.055	0.061	0.058	0.050	0.047	0.055
1	128	0.100	0.045	0.075	0.225	0.210	0.154	0.087	0.075	0.081	0.080	0.066	0.052	0.043
	256	0.077	0.041	0.060	0.153	0.133	0.112	0.084	0.075	0.077	0.076	0.064	0.059	0.060
	512	0.071	0.052	0.064	0.121	0.103	0.093	0.075	0.065	0.077	0.075	0.068	0.064	0.057
	1024	0.044	0.025	0.037	0.088	0.081	0.070	0.050	0.052	0.062	0.056	0.049	0.043	0.044
1.33	128	0.039	0.039	0.054	0.187	0.169	0.138	0.079	0.072	0.077	0.073	0.063	0.049	0.045
	256	0.040	0.050	0.058	0.132	0.119	0.098	0.076	0.071	0.064	0.063	0.054	0.063	0.056
	512	0.034	0.047	0.053	0.097	0.087	0.080	0.060	0.060	0.063	0.057	0.057	0.051	0.043
	1024	0.041	0.057	0.063	0.080	0.074	0.068	0.064	0.058	0.065	0.060	0.055	0.059	0.054

Table 1 (d). Null rejection frequencies of tests at the 0.05 level. $\rho = 1 - c^*/T$.

		$\mathcal{SV}$	$\mathcal{SV}^A$	$\mathcal{SV}_0^A$	$\mathcal{MW}^A$					$\mathcal{MW}_0^A$				
	b				0.025	0.050	0.100	0.200	0.300	0.025	0.050	0.100	0.200	0.300
c^*	T													
5	128	0.025	0.019	0.022	0.088	0.076	0.057	0.034	0.041	0.026	0.024	0.024	0.017	0.027
	256	0.037	0.030	0.033	0.064	0.062	0.041	0.031	0.033	0.025	0.025	0.020	0.020	0.030
	512	0.035	0.033	0.031	0.026	0.025	0.025	0.023	0.026	0.013	0.013	0.018	0.020	0.023
	1024	0.039	0.035	0.035	0.019	0.018	0.016	0.028	0.031	0.012	0.013	0.015	0.027	0.029
10	128	0.021	0.014	0.013	0.058	0.047	0.039	0.020	0.025	0.018	0.018	0.015	0.013	0.018
	256	0.020	0.018	0.019	0.030	0.025	0.022	0.026	0.019	0.008	0.008	0.010	0.019	0.017
	512	0.029	0.020	0.022	0.008	0.006	0.008	0.014	0.017	0.003	0.003	0.005	0.013	0.016
	1024	0.030	0.026	0.027	0.005	0.005	0.008	0.016	0.015	0.001	0.002	0.006	0.015	0.014
20	128	0.018	0.014	0.010	0.045	0.035	0.024	0.023	0.026	0.010	0.010	0.008	0.012	0.018
	256	0.023	0.012	0.012	0.018	0.014	0.014	0.017	0.025	0.007	0.006	0.007	0.012	0.020
	512	0.023	0.015	0.016	0.003	0.003	0.004	0.015	0.012	0.001	0.001	0.001	0.013	0.012
	1024	0.027	0.017	0.017	0.001	0.001	0.003	0.009	0.017	0.000	0.000	0.002	0.009	0.014

Table 2(a). δ-adaptive critical values for $\mathcal{SV}$.

0.01 level								
b	δ=0	δ=0.1	δ=0.2	δ=0.3	δ=0.4	δ=0.5	δ=0.6	δ=0.7
0.02	4.49	6.29	8.29	10.34	11.25	11.31	11.17	10.44
0.06	6.31	7.41	8.57	9.89	10.82	10.77	10.45	10.13
0.08	7.31	8.27	9.75	10.76	11.69	11.74	11.56	11.42
0.28	80.50	87.13	86.27	96.39	100.66	96.35	94.61	95.54
0.30	102.20	109.12	113.72	121.69	127.23	122.35	125.28	124.16
b	δ=0.8	δ=0.9	δ=1	δ=1.1	δ=1.2	δ=1.3	δ=1.4	
0.02	8.93	7.54	5.00	3.01	1.68	0.99	0.50	
0.06	9.54	8.81	7.40	6.02	4.07	3.08	2.09	
0.08	10.89	10.14	8.72	6.99	5.12	3.89	2.92	
0.28	103.92	92.18	84.55	73.01	64.89	59.45	49.53	
0.30	129.11	117.60	109.94	98.48	83.48	76.22	62.44	

0.05 level								
b	δ=0	δ=0.1	δ=0.2	δ=0.3	δ=0.4	δ=0.5	δ=0.6	δ=0.7
0.02	2.82	3.89	5.07	6.45	7.58	8.12	8.06	7.19
0.06	3.49	4.17	4.94	5.70	6.41	6.75	6.84	6.47
0.08	4.17	4.75	5.44	6.15	6.76	7.11	7.20	6.83
0.28	38.78	41.01	42.56	44.13	46.07	47.27	46.85	46.73
0.30	49.35	52.22	54.38	56.38	59.14	60.54	60.31	59.83
b	δ=0.8	δ=0.9	δ=1	δ=1.1	δ=1.2	δ=1.3	δ=1.4	
0.02	5.95	4.37	2.84	1.70	0.96	0.45	0.19	
0.06	5.73	4.82	3.79	2.84	1.96	1.31	0.80	
0.08	6.24	5.42	4.38	3.45	2.46	1.73	1.16	
0.28	45.60	43.90	40.35	36.39	31.47	26.73	21.85	
0.30	58.09	56.46	51.58	46.60	40.24	34.06	27.11	

0.10 level								
b	δ=0	δ=0.1	δ=0.2	δ=0.3	δ=0.4	δ=0.5	δ=0.6	δ=0.7
0.02	2.03	2.85	3.87	4.87	5.89	6.48	6.34	5.68
0.06	2.62	3.10	3.62	4.11	4.61	5.07	5.11	4.87
0.08	2.97	3.40	3.86	4.31	4.73	5.19	5.22	5.04
0.28	25.09	25.98	27.40	28.42	29.84	30.32	30.58	30.31
0.30	32.01	33.05	34.89	36.36	38.21	38.92	39.27	38.97
b	δ=0.8	δ=0.9	δ=1	δ=1.1	δ=1.2	δ=1.3	δ=1.4	
0.02	4.56	3.30	2.17	1.28	0.68	0.28	0.10	
0.06	4.18	3.41	2.70	2.04	1.41	0.83	0.44	
0.08	4.47	3.79	3.07	2.38	1.77	1.13	0.66	
0.28	29.08	27.10	24.29	21.83	18.52	15.31	11.93	
0.30	37.63	34.54	31.24	28.03	24.07	20.05	15.35	

Table 2(b). δ-adaptive critical values for $\mathcal{MW}$.

0.01 level

b	δ=0	δ=0.1	δ=0.2	δ=0.3	δ=0.4	δ=0.5	δ=0.6	δ=0.7
0.025	4.72	6.67	9.42	13.41	18.64	24.96	33.46	44.36
0.05	5.91	7.23	8.99	11.23	14.51	17.88	20.66	27.74
0.10	8.94	10.18	11.52	13.32	15.36	17.36	20.60	24.07
0.20	27.30	30.48	33.59	35.75	38.76	41.45	47.70	51.26
0.30	102.10	110.26	115.87	127.83	141.13	142.66	161.89	182.12
b	δ=0.8	δ=0.9	δ=1	δ=1.1	δ=1.2	δ=1.3	δ=1.4	
0.025	57.46	75.32	96.07	119.22	143.80	174.46	207.71	
0.05	35.02	44.99	55.02	65.99	78.29	93.18	108.91	
0.10	28.46	33.68	39.61	45.74	52.35	59.05	66.23	
0.20	57.19	65.42	72.77	77.27	83.62	92.31	103.91	
0.30	210.49	230.69	255.71	267.44	292.59	317.52	356.21	

0.05 level

b	δ=0	δ=0.1	δ=0.2	δ=0.3	δ=0.4	δ=0.5	δ=0.6	δ=0.7
0.025	2.93	4.11	5.64	7.70	10.64	14.65	20.18	27.24
0.05	3.36	4.20	5.26	6.59	8.35	10.69	12.81	17.55
0.10	5.03	5.62	6.40	7.40	8.51	10.31	12.28	14.66
0.20	14.65	15.73	16.59	17.49	19.21	20.86	23.13	25.86
0.30	49.43	52.72	55.47	58.74	63.94	69.17	73.39	83.29
b	δ=0.8	δ=0.9	δ=1	δ=1.1	δ=1.2	δ=1.3	δ=1.4	
0.025	36.45	47.43	61.39	77.95	95.89	116.37	139.69	
0.05	22.25	27.78	34.55	43.05	52.17	61.98	73.19	
0.10	17.36	20.40	24.17	28.57	33.60	38.82	44.49	
0.20	29.62	32.77	36.05	40.75	43.72	48.43	53.07	
0.30	92.79	103.51	114.63	126.13	140.36	154.19	173.39	

0.10 level

b	δ=0	δ=0.1	δ=0.2	δ=0.3	δ=0.4	δ=0.5	δ=0.6	δ=0.7
0.025	2.13	2.96	4.10	5.59	7.78	10.84	15.05	20.38
0.05	2.45	3.05	3.77	4.71	6.02	7.80	9.53	13.20
0.10	3.46	3.89	4.51	5.18	6.04	7.18	8.68	10.43
0.20	9.31	9.82	10.68	11.47	12.49	13.67	15.31	17.49
0.30	32.01	33.51	35.57	37.67	41.02	44.07	48.41	53.50
b	δ=0.8	δ=0.9	δ=1	δ=1.1	δ=1.2	δ=1.3	δ=1.4	
0.025	27.70	36.61	46.87	58.70	72.37	88.87	108.10	
0.05	16.85	21.15	26.50	32.31	38.95	47.34	56.70	
0.10	12.62	15.12	18.08	21.01	24.54	29.15	33.62	
0.20	19.48	21.79	24.59	27.70	30.72	34.10	37.76	
0.30	59.55	65.63	73.91	81.14	89.84	100.66	111.42	

Figure 1 (a). Local power of tests at the 0.05 level. T=256, δ=0.00

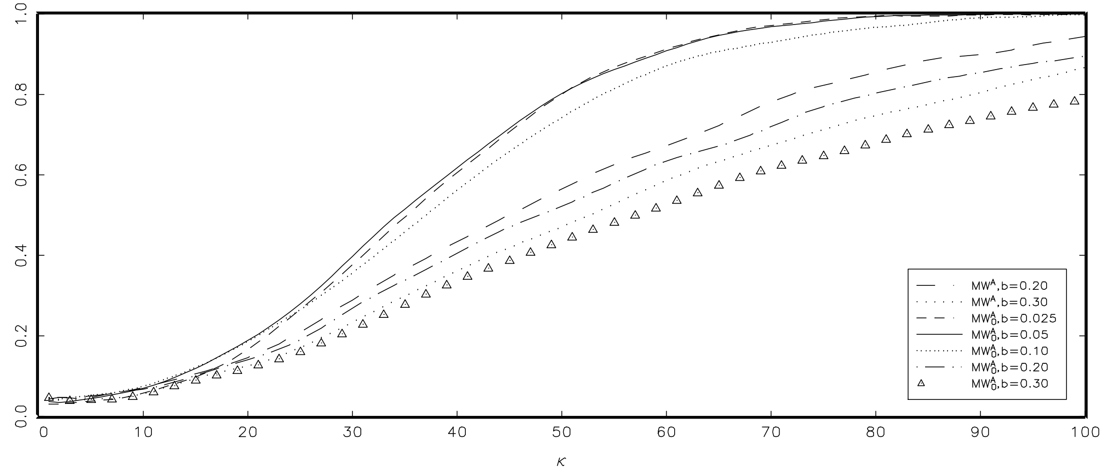

Figure 1 (b). Local power of tests at the 0.05 level. T=256, δ=0.33

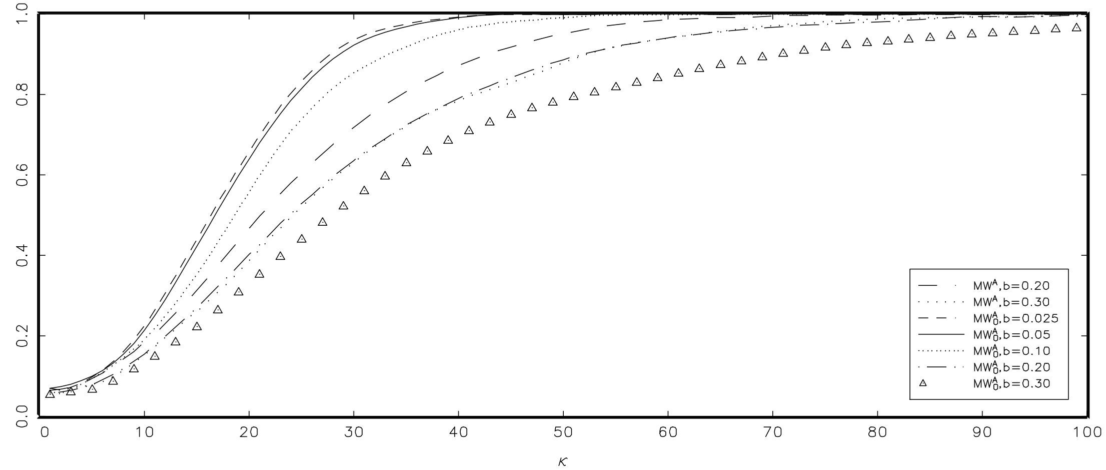

Figure 1 (c). Local power of tests at the 0.05 level. T=256, δ=0.66

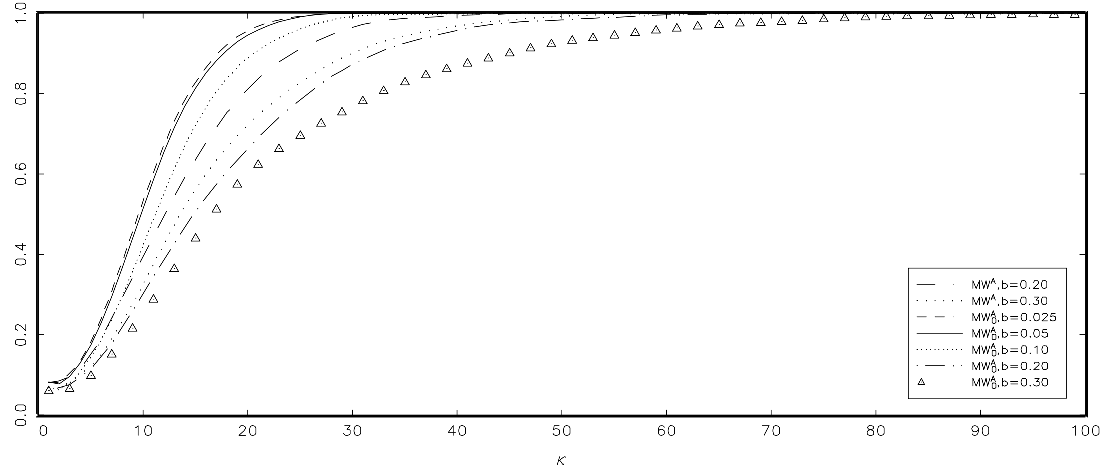

Figure 1 (d). Local power of tests at the 0.05 level. T=256, δ=1.00

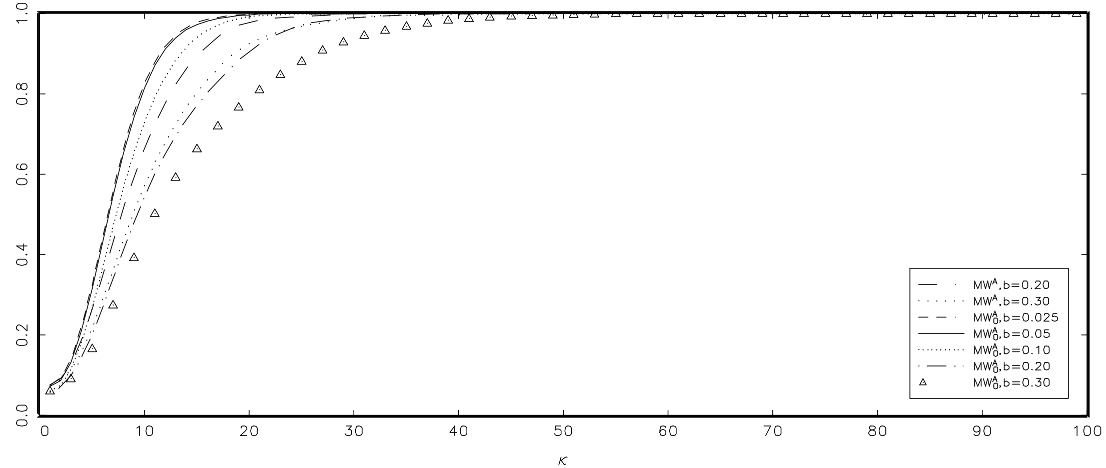

Figure 1 (e). Local power of tests at the 0.05 level. T=256, δ=1.33

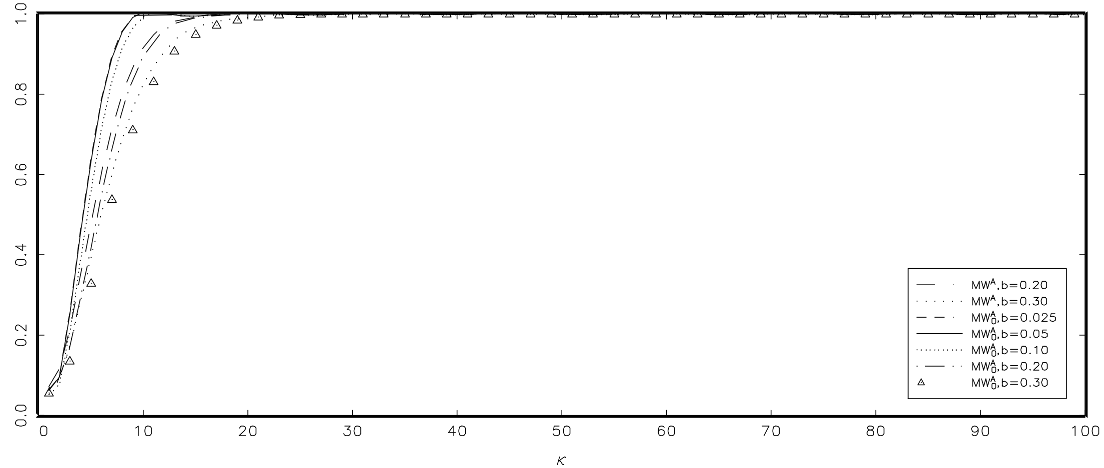

Figure 2 (a). Local power of tests at the 0.05 level. T=256, δ=0.00

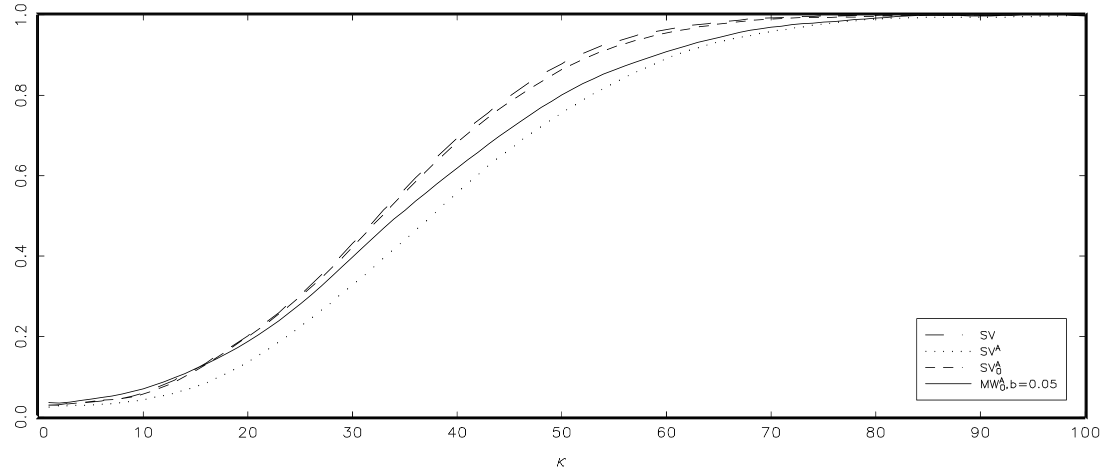

Figure 2 (b). Local power of tests at the 0.05 level. T=256, δ=0.33

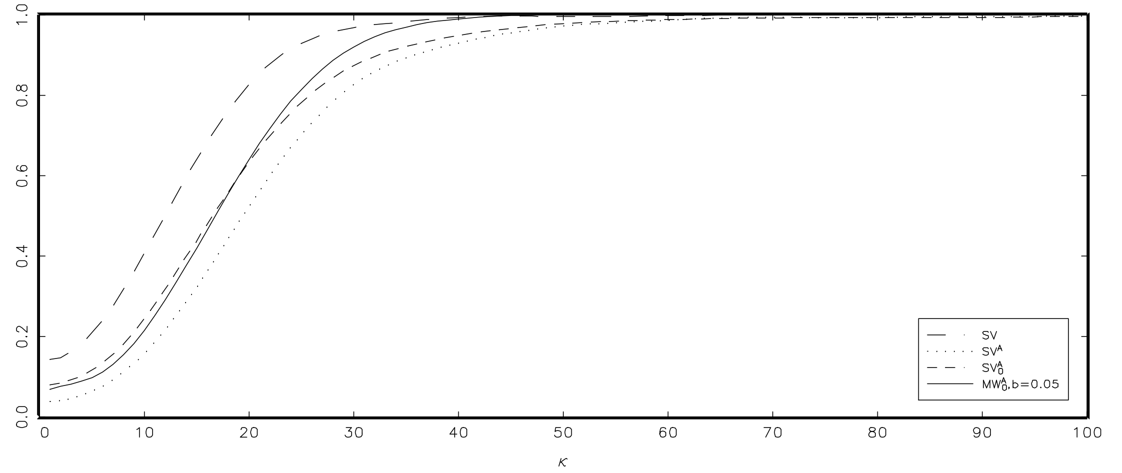

Figure 2 (c). Local power of tests at the 0.05 level. T=256, δ=0.66

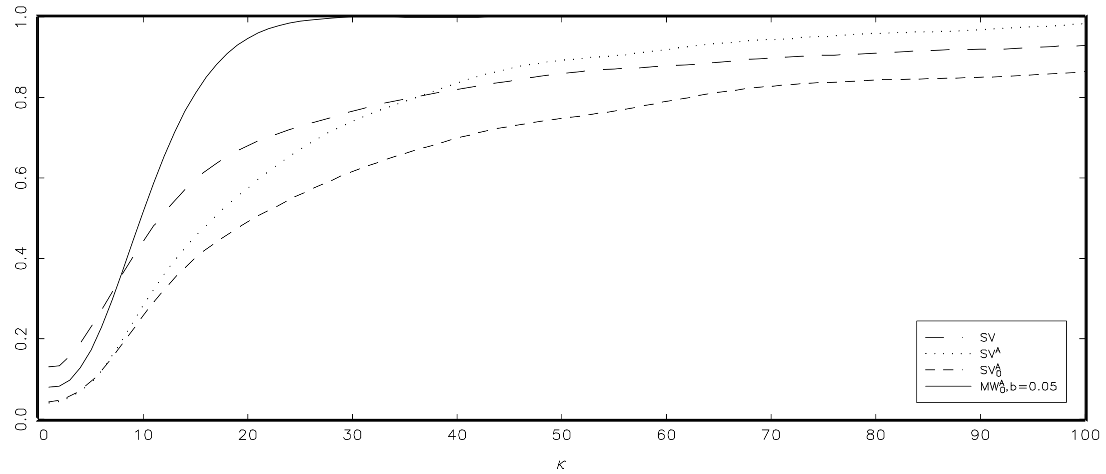

Figure 2 (d). Local power of tests at the 0.05 level. T=256, δ=1.00

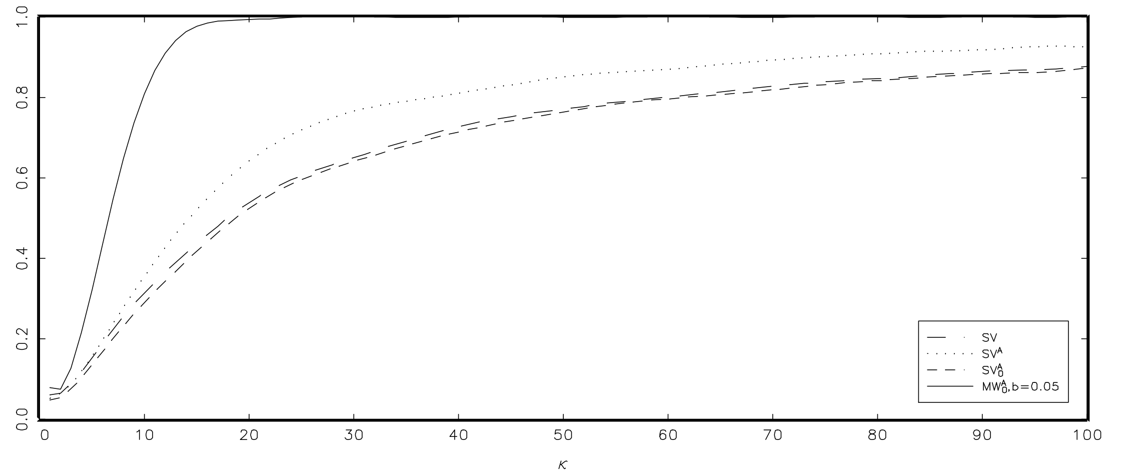

Figure 2 (e). Local power of tests at the 0.05 level. T=256, δ=1.33

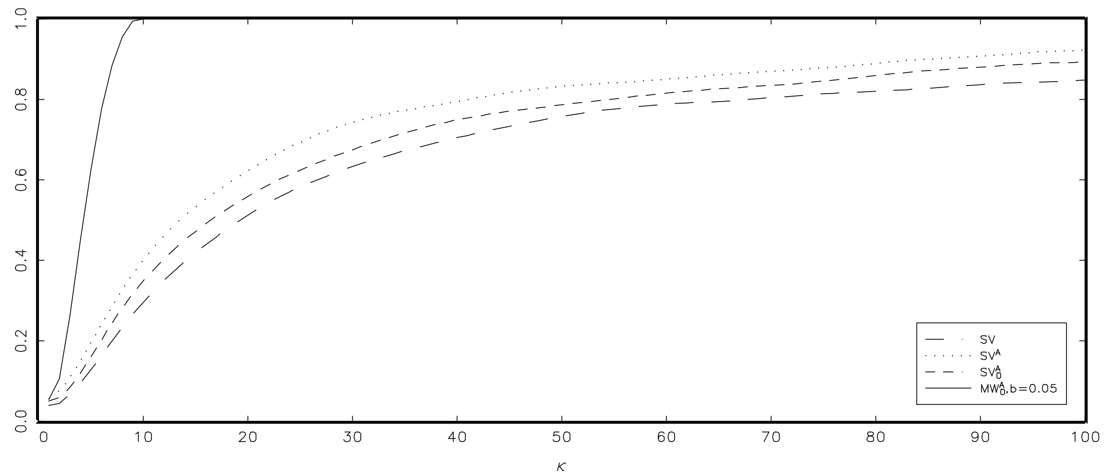

Figure 3 (a). Local power of tests at the 0.05 level. T=256, c*=5

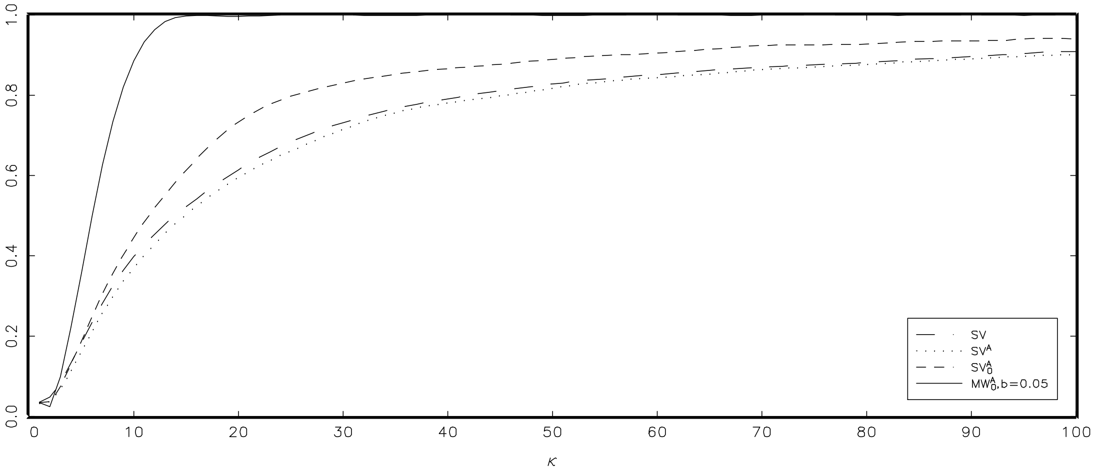

Figure 3 (b). Local power of tests at the 0.05 level. T=256, c*=10

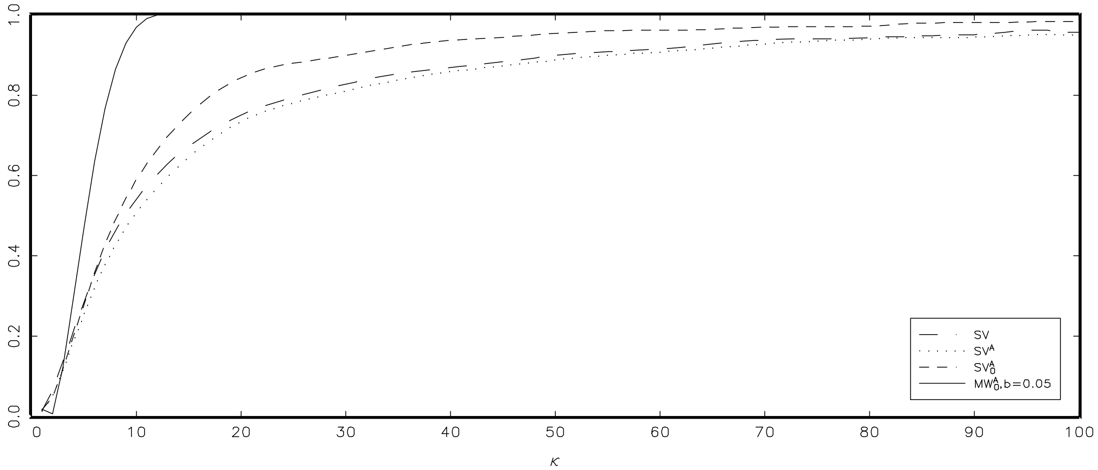

Figure 3 (c). Local power of tests at the 0.05 level. T=256, c*=20