Granger Centre Discussion Paper Series

Unit root testing under a local break in trend

by

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

Granger Centre Discussion Paper No. 11/02

Unit Root Testing under a Local Break in Trend

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

Granger Centre for Time Series Econometrics, University of Nottingham

October 2011

Abstract

Recent approaches to testing for a unit root when uncertainty exists over the presence and timing of a trend break employ break detection methods, so that a with-break unit root test is used only if a break is detected by some auxiliary statistic. While these methods achieve near asymptotic efficiency in both fixed trend break and no trend break environments, in finite samples pronounced "valleys" in the power functions of the tests (when mapped as functions of the break magnitude) are observed, with power initially high for very small breaks, then decreasing as the break magnitude increases, before increasing again. In response to this problem we propose two practical solutions, based either on the use of a with-break unit root test but with adaptive critical values, or on a union of rejections principle taken across with-break and without break unit root tests. These new procedures are shown to offer improved reliability in terms of finite sample power. We also develop local limiting distribution theory for both the extant and the newly proposed unit root statistics, treating the trend break magnitude as local-to-zero. We show that this framework allows the asymptotic analysis to closely approximate the finite sample power valley phenomenon, thereby providing useful analytical insights.

Keywords: Unit root test; local trend break; union of rejections; adaptive critical values; asymptotic local power.

JEL Classification: C22.

*Corresponding author: Robert Taylor, School of Economics, University of Nottingham, Nottingham NG7 2RD, UK. E-mail: Robert.Taylor@nottingham.ac.uk

1 Introduction

Macroeconomic series appear to often be characterized by broken trend functions; see, *inter alia*, Stock and Watson (1996,1999,2005) and Perron and Zhu (2005). Consequently, following the seminal paper by Perron (1989), when testing for a unit root it has become a matter of regular practice to allow for the possibility of this kind of deterministic structural change. While Perron (1989) treated the location of the potential trend break as known, most recent approaches have focused on the case where the possible break occurs at an unknown point in the sample; see, *inter alia*, Zivot and Andrews (1992), Banerjee *et al.* (1992), Perron (1997) and Perron and Rodríguez (2003) [PR].

Taking the presence of a linear trend in the data generation process [DGP] as given, among augmented Dickey-Fuller [ADF] style unit root tests it is the Elliott *et al.* (1996) [ERS] test based on GLS detrending that is near asymptotically efficient[1] in terms of local power when no additional trend break is present. When a trend break is *known* to be present, it is now a test based on PR's GLS detrended ADF statistic which allows for a trend break in the deterministic function that is asymptotically efficient, provided the break point is known. This efficiency carries over to the case where a break occurs at an unknown point, the case we consider here, provided the unknown break point can be detected (i.e. dated) precisely enough, as this allows the critical values for the known break point case of to be applied in the limit.

However, when a trend break does *not* occur the PR test is no longer asymptotically efficient since an irrelevant trend break regressor is included in its deterministic specification which compromises power. There is also a second consideration arising from the fact that the unit root null asymptotic critical values for the PR test based on estimating the break point differ markedly according to whether a break occurs or not. Estimation is typically carried out by minimizing the OLS or GLS residual sum of squares across an interval of candidate trend break points. If a trend break exists and is of sufficiently large magnitude, then it will be correctly identified by this procedure. However, when no break exists, we find that the no break case critical values are substantially left-shifted relative to their break case counterparts. Since the PR test is a left-tailed unit root test, it is then necessary to always employ the no break case critical values in order to avoid over-sizing problems in the no break case. Of course, this implies that when a break does occur (and can be dated with sufficient precision), the PR test is rendered conservative (under-sized) which obviously reduces the level of power from that obtainable if the non-conservative break critical values had been employed. The underlying problem is then essentially one of uncertainty over whether or not a trend break occurs.

Two recent papers by Carrion-i-Silvestre *et al.* (2009) [CKP] and Harris *et al.* (2009) [HHLT] have proposed solutions to the issues raised in the last paragraph.

[1] Although not formally asymptotically efficient, in the limit these tests lie arbitrarily close to the asymptotic Gaussian local power envelopes for these testing problems and, hence, with a small abuse of language we shall refer to tests with this property as 'asymptotically efficient'.

Although the two procedures differ, essentially, both utilize auxiliary statistics which are employed to detect the presence of a trend break occurring at an unknown point and then use the outcome of the detection step to indicate whether the unit root test should include a trend break in the deterministic specification. HHLT employ their detection methods, either a modified break fraction estimator or the trend break test of Harvey *et al.* (2009b), to choose between the ERS and PR tests, while CKP use their detection method, the trend break test of Perron and Yabu (2009b) [PY], to choose between likelihood ratio [LR] test variants of ERS and PR. When there is no break all the procedures are asymptotically efficient, and when a break exists both procedures are also asymptotically efficient since they can employ non-conservative critical values in this case by virtue of their respective break point estimators converging to the unknown break fraction at a sufficiently fast rate. The only real difference between the procedures lies in the way in which the break detection stage is carried out.[2]

Both CKP and HHLT provide finite sample simulation evidence on the power performance of their respective unit root test procedures and, since the simulation DGPs are similar, the finite sample powers, shown as functions of the trend break magnitude, appear similar also. What is very much apparent from these results is that while the finite sample power levels are acceptably high when the trend break magnitude is either zero or large, there is an intermediate range of values of the trend break magnitude where the power falls off to an alarming extent, giving rise to a pronounced "valley" in the power profiles when graphed across trend break magnitudes. This behaviour arises since a range of break magnitudes exists where the breaks are simply too small to be reliably detected, so that unit root tests which exclude the trend break regressors will tend to be conducted here. Whilst these breaks are not readily detectable, they are still of sufficient magnitude to effect a severe decrease in the power of the without-break unit root tests. This phenomenon is not predicted by the preceding asymptotic analyses in either CKP or HHLT and therefore comes as something of an unpleasant surprise. The reason for this discrepancy is that the asymptotics of CKP and HHLT assume that the trend breaks have a fixed (independent of the sample size) magnitude. By virtue of consistency of the break detection procedures it follows that in the limit trend breaks are detected (and dated) with complete certainty - no matter how small their magnitude. The finite sample effects of breaks with magnitudes lying in an intermediate region where, as in practice, detection is not a certain event (indeed, the probability of detection may be very low), are thereby entirely obscured under the fixed magnitude break asymptotic framework.

These observations throw up two natural questions. The first is whether or not a solution to the power valley problem can be found, and attempting to address this constitutes the first contribution of this paper. If we consider only procedures that use some auxiliary statistic to select between application of unit root tests which exclude

[2]It should be noted that the procedure outlined in CKP also allows for the possibility of multiple breaks in trend. Moreover, CKP also consider procedures based on the GLS de-trended M-type unit root tests of Ng and Perron (2001) and PR, but only report simulation results in the unknown break date case for the approach based on LR-type tests.

or include trend break regressors, then the answer is almost certainly no, since the discriminatory ability of any statistic in finite samples must necessarily be poor for some range of break magnitudes, leading to application of unit root tests that neglect the break, which is where the problem arises. However, if we are prepared to sacrifice all, or at least some, of the (potential) power available from those unit root tests which exclude trend break regressors then progress can be made. The simplest approach available would be to just apply a PR-type test employing conservative critical values. However, we show that one can do rather better than this. We explore two possible procedures, both of which employ auxiliary statistics in a discriminatory role, but at a somewhat less explicit level than CKP and HHLT. The first involves again using only the PR test but with adaptive critical values, whereby the conservative critical value is used only when no break is detected; otherwise the known break date critical value is used. This approach avoids the potential for very low power in the presence of small breaks seen with the HHLT and CKP tests, but can never achieve the additional power available when no break is present. The second approach, based on a union of rejections approach (again with adaptive critical values), rejects if either the with-break or the without-break version of the unit root test rejects, which thereby allows us to capture some of the additional power available under the no break case.

The second question raised is can we find an alternative asymptotic framework which, unlike the prior approaches based on the assumption of a fixed trend break magnitude, is capable of reproducing the power valley phenomenon which we observe in finite samples? Addressing this question is the second contribution of this paper. To this end, we examine the local (to unit root) asymptotic power of the tests when the trend break is also made local (to zero) in magnitude, by means of shrinking its magnitude with the sample size. By employing the relevant Pitman drift for a local trend break we find that resulting local asymptotic theory can indeed yield very good predictions of finite sample behaviour. The local trend break model is therefore important because it retains in the asymptotic framework the genuine uncertainty that will exist in finite samples as to whether or not a trend break is present in the data. This kind of uncertainty, of course, is exactly that which the CKP and HHLT procedures were designed to address and yet, ironically, it is precisely in these situations where they perform most poorly.

The paper is organised as follows. Our reference trend break model is outlined in section 2. The unit root test procedures of HHLT and CKP are presented in section 3, along with an analysis of the finite sample behaviour of these tests. In section 4 we introduce the alternative procedures discussed above based on adaptive critical values and/or a union of rejections principle, and examine their finite sample behaviour relative to the HHLT and CKP tests. Section 5 details the large sample distributions of all the procedures under a local-to-zero trend break, and local asymptotic power simulations confirm that the limit representations can closely predict the finite sample power functions of the tests. Section 6 offers some concluding remarks. Proofs are collected in an Appendix.

In what follows the notation: '$\lfloor \cdot \rfloor$' denotes the integer part; '$\xrightarrow{d}$' denotes weak

convergence, and '$\xrightarrow{p}$' convergence in probability; '1(.)' denotes the indicator function and '$x := y$' ('$x =: y$') indicates that x is defined by y (y is defined by x).

2 The Trend Break Model

In keeping with the analysis of HHLT and CKP in the case of a single break in trend, we consider a time series $\{y_t\}$ to be generated according to the following model,

$$y_t = \mu + \beta t + \gamma_T DT_t(\tau_0) + u_t, \quad t = 1, ..., T \tag{1}$$
$$u_t = \rho_T u_{t-1} + \varepsilon_t, \quad t = 2, ..., T \tag{2}$$

where $DT_t(\tau_0) := 1(t > \lfloor \tau_0 T \rfloor)(t - \lfloor \tau_0 T \rfloor)$, with $\lfloor \tau_0 T \rfloor$ the potential trend break point with associated break fraction τ_0, and break magnitude γ_T. A break in trend occurs in $\{y_t\}$ at time $\lfloor \tau_0 T \rfloor$ when $\gamma_T \neq 0$. The true break fraction τ_0 is treated as unknown, but is assumed to satisfy $\tau_0 \in \Lambda$, where $\Lambda := [\tau_L, \tau_U]$ with $0 < \tau_L < \tau_U < 1$; the fractions τ_L and τ_U representing trimming parameters, below and above which, respectively, no break is deemed allowable to occur.

We assume the initialization of $\{u_t\}$ is such that $u_1 = o_p(T^{1/2})$, while $\{\varepsilon_t\}$ is assumed to satisfy the following linear process assumption:

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

$$\varepsilon_t = C(L)\eta_t, \; C(L) := \sum_{j=0}^{\infty} C_j L^j$$

with $C(1)^2 > 0$ and $\sum_{i=0}^{\infty} i|C_i| < \infty$, and where $\{\eta_t\}$ is an IID sequence with mean zero, variance σ_η^2 and finite fourth moment. The long-run variance of ε_t is defined as $\omega_\varepsilon^2 := \lim_{T\to\infty} T^{-1} E(\sum_{t=1}^{T} \varepsilon_t)^2 = \sigma_\eta^2 C(1)^2$. Finally, let σ_ε^2 denote the (short-run) variance of ε_t.

Our focus is on testing the unit root null hypothesis $H_0 : \rho_T = 1$, against the local alternative, $H_1 : \rho_T = 1 - c/T$, $0 < c < \infty$, without assuming knowledge of whether or not a trend break is actually present. As discussed in section 1, two alternative assumptions can be made regarding the trend break magnitude. The first, employed by HHLT and CKP, is that the magnitude is fixed (independent of the sample size, T); i.e., $\gamma_T = \gamma$. Alternatively, one can let the break magnitude be *local-to-zero* by setting $\gamma_T = \kappa\omega_\varepsilon T^{-1/2}$, thereby adopting the appropriate Pitman drift for a trend break in a local-to-unity process; cf. Vogelsang (1998).[3] For a given sample size, these two assumptions are, of course, observationally equivalent, but as we shall see in section 5, they deliver quite different asymptotic distribution theory. In particular the local-to-zero assumption allows the asymptotic theory to accurately reproduce the power valley phenomenon observed in finite samples.

[3]Scaling the trend break by ω_ε is merely a convenience device allowing it to be factored out of the local limit distributions that arise in section 5.

3 The HHLT and CKP Tests

HHLT and CKP both propose testing strategies for the unit root null when there is uncertainty regarding the presence of a break in trend. The two procedures centre on implementing either a GLS-detrended unit root test allowing for simply a constant and trend, or a corresponding test that additionally allows for a break in trend; in the latter case, critical values associated with a known break fraction are used. The decision as to whether the with-break or without-break version of the unit root test is employed is governed by auxiliary statistics designed to detect the presence of a trend break occurring at an unknown point. The HHLT and CKP testing approaches differ in the form of unit root tests employed, the break fraction estimator adopted, and the auxiliary method proposed for determining whether or not a trend break is incorporated into the unit root test. We now give a description of the HHLT and CKP procedures; in what follows, it is convenient to define $\mathbf{r}_\theta$ generically as the residuals from a regression of $\mathbf{y}_\theta$ on $\mathbf{Z}_\theta$, and $\mathbf{r}_{\theta,\tau}$ as the residuals from a regression of $\mathbf{y}_\theta$ on $\mathbf{Z}_{\theta,\tau}$, where

$$\begin{aligned}
\mathbf{y}_\theta &:= [y_1, y_2 - \theta y_1, ..., y_T - \theta y_{T-1}]' \\
\mathbf{Z}_\theta &:= [z_1, z_2 - \theta z_1, ..., z_T - \theta z_{T-1}]' \text{ with } z_t := [1, t]' \\
\mathbf{Z}_{\theta,\tau} &:= [z_1, z_2 - \theta z_1, ..., z_T - \theta z_{T-1}]' \text{ with } z_t := [1, t, DT_t(\tau)]'.
\end{aligned}$$

and with the corresponding residual sums of squares denoted by $S(\theta) := \mathbf{r}_\theta'\mathbf{r}_\theta$ and $S(\theta,\tau) := \mathbf{r}_{\theta,\tau}'\mathbf{r}_{\theta,\tau}$.

3.1 The HHLT Tests

HHLT propose two approaches to testing. The starting point for their preferred procedure is to compute an initial estimator of the break fraction, based on a first differenced version of the regression in (1), that is,

$$\tilde{\tau} := \arg\min_{\tau\in\Lambda} S(1,\tau).$$

Next, conditional on this break fraction estimator, they construct a modified break date estimator which also assumes the role of break detection. Specifically, following Vogelsang (1997,1998), they first compute the Wald statistic

$$W_T(\tilde{\tau}) := \frac{S_R}{S_U(\tilde{\tau})} - 1$$

with S_R the RSS from the fitted (cumulated) restricted OLS regression

$$\textstyle\sum_{i=1}^t y_i = \hat{\mu}t + \hat{\beta}\sum_{i=1}^t i + \hat{s}_{Rt}$$

and $S_U(\tilde{\tau})$ the RSS from the fitted (cumulated) unrestricted OLS regression[4]

$$\textstyle\sum_{i=1}^t y_i = \hat{\mu}t + \hat{\beta}\sum_{i=1}^t i + \hat{\gamma}\sum_{i=1}^t DT_i(\tilde{\tau}) + \hat{s}_{Ut}.$$

[4]We suppress the dependence of the estimates on τ for notational brevity, unless it becomes essential to the argument. Similarly, we also use the "hat" notation for estimates in a generic sense.

Following this, $W_T(\tilde{\tau})$ is used in a weight function of the form

$$\bar{\lambda} := \exp\{-g_W T^{-1/2} W_T(\tilde{\tau})\} \tag{3}$$

where g_W is some positive constant, to yield the final modified estimator of τ_0, given by

$$\bar{\tau} := (1 - \bar{\lambda})\tilde{\tau}.$$

The modified estimator has the properties that: (i) when no break occurs, $\bar{\lambda}$ converges to unity in such a way that $\bar{\tau}$ converges to zero (at rate $O_p(T^{-1/2})$), thereby signalling the lack of a trend break in the series, and (ii) when a break in trend of *fixed* magnitude occurs, $\bar{\lambda}$ converges to zero in such a way that $\bar{\tau}$ converges to τ_0 (at the rate $O_p(T^{-1})$), achieving the same rate of consistency as the first differenced-based estimator, $\tilde{\tau}$. It is therefore evident that HHLT use the modified estimator $\bar{\tau}$ as an auxiliary statistic for detecting whether or not a break in trend is present: if $\bar{\tau}$ lies within the range of allowable break fractions, that is if $\tau_L \leq \bar{\tau} \leq \tau_U$, a unit root test allowing for a break in trend is applied; alternatively if $\bar{\tau} < \tau_L$, a standard unit root test with unbroken trend is used.

The first HHLT test statistic is then given by

$$HHLT_{\bar{\tau}} := \begin{cases} DF_t^{GLS} & \text{if } \bar{\tau} < \tau_L \\ DF_{tb}^{GLS}(\bar{\tau}) & \text{if } \bar{\tau} \geq \tau_L \end{cases}. \tag{4}$$

Here DF_t^{GLS} is the ERS unit root statistic which allows for a constant and linear time trend, *viz.*, the t-ratio associated with $\hat{\phi}$ in the fitted ADF regression

$$\Delta\tilde{u}_t = \hat{\phi}\tilde{u}_{t-1} + \sum_{j=1}^{p} \hat{\delta}_j \Delta\tilde{u}_{t-j} + \hat{\eta}_t, \quad t = p+2, ..., T \tag{5}$$

where $\tilde{u}_t := y_t - \tilde{\mu} - \tilde{\beta}t$, with $\tilde{\mu}$ and $\tilde{\beta}$ obtained from the regression of $\mathbf{y}_{\bar{\rho}}$ on $\mathbf{Z}_{\bar{\rho}}$ with $\bar{\rho} = 1 - \bar{c}/T$ (with $\bar{c} = 13.5$). For a generic break fraction τ, $DF_{tb}^{GLS}(\tau)$ is the PR statistic obtained as the t-ratio for $\hat{\phi}$ in the fitted ADF regression

$$\Delta\tilde{u}_{\tau,t} = \hat{\phi}\tilde{u}_{\tau,t-1} + \sum_{j=1}^{p} \hat{\delta}_j \Delta\tilde{u}_{\tau,t-j} + \hat{\eta}_t, \quad t = p+2, ..., T \tag{6}$$

where $\tilde{u}_{\tau,t} := y_t - \tilde{\mu}_\tau - \tilde{\beta}_\tau t - \tilde{\gamma}_\tau DT_t(\tau)$, with $\tilde{\mu}_\tau$, $\tilde{\beta}_\tau$ and $\tilde{\gamma}_\tau$ obtained from the regression of $\mathbf{y}_{\bar{\rho}_\tau}$ on $\mathbf{Z}_{\bar{\rho}_\tau,\tau}$ with $\bar{\rho}_\tau = 1 - \bar{c}_\tau/T$. Here, $\bar{c}_\tau$ is chosen according to τ using either Table 1 of HHLT or Table 1 of CKP. All these expressions are evaluated at $\tau = \bar{\tau}$. Standard critical values pertaining to DF_t^{GLS} are used, while for $DF_{tb}^{GLS}(\bar{\tau})$, the statistic is compared with critical values associated with a known break fraction, as given in HHLT. In both (5) and (6), and indeed in all subsequent ADF-type regressions in this paper, the lag truncation parameter, p, is taken to have been chosen according to an appropriate model selection procedure, such as the modified Akaike information criterion (MAIC) procedure of Ng and Perron (2001) and Perron and Qu (2007).

The second procedure considered by HHLT selects between DF_t^{GLS} and $DF_{tb}^{GLS}(\tilde{\tau})$ on the basis of the outcome of a robust pre-test for a break in trend. Specifically, they first apply the trend break test of Harvey *et al.* (2009b), given by

$$t_\lambda := \lambda |t_0(\hat{\tau})| + m_\xi(1-\lambda)|t_1(\mathring{\tau})| \tag{7}$$

with m_ξ a positive finite constant as detailed in HHLT. For a generic break fraction τ, $t_0(\tau)$ and $t_1(\tau)$ are the autocorrelation-adjusted t-ratios

$$t_0(\tau) := \frac{\hat{\gamma}_\tau}{\sqrt{\hat{\omega}^2(\hat{u}_{\tau,t})(\sum_{t=1}^T \mathbf{x}_{0,\tau,t}\mathbf{x}'_{0,\tau,t})^{-1}_{33}}}, \qquad t_1(\tau) := \frac{\check{\gamma}_\tau}{\sqrt{\hat{\omega}^2(\check{v}_{\tau,t})(\sum_{t=2}^T \mathbf{x}_{1,\tau,t}\mathbf{x}'_{1,\tau,t})^{-1}_{22}}}$$

where $\mathbf{x}_{0,\tau,t} := [1, t, DT_t(\tau)]'$ and $\mathbf{x}_{1,\tau,t} := [1, DU_t(\tau)]'$ with $DU_t(\tau) := 1(t > \lfloor \tau T \rfloor)$, and where $\hat{\gamma}_\tau$ and $\hat{u}_{\tau,t}$ obtained from the estimated OLS regression

$$y_t = \hat{\mu}_\tau + \hat{\beta}_\tau t + \hat{\gamma}_\tau DT_t(\tau) + \hat{u}_{\tau,t}$$

and $\check{\gamma}_\tau$ and $\check{v}_{\tau,t}$ from the estimated regression

$$\Delta y_t = \check{\beta}_\tau + \check{\gamma}_\tau DU_t(\tau) + \check{v}_{\tau,t}.$$

Here, for a generic e_t, $\hat{\omega}^2(e_t)$ denotes the long-run variance estimator using a Bartlett kernel, that is

$$\hat{\omega}^2(e_t) := \hat{\psi}_0 + 2\sum_{j=1}^{l}(1 - \frac{j}{l+1})\hat{\psi}_j, \qquad \hat{\psi}_j = T^{-1}\sum_{t=j+1}^{T} e_t e_{t-j} \tag{8}$$

where the bandwidth, l, is required to satisfy the usual condition that $1/l + l^2/T \to 0$ as $T \to \infty$. We follow HLT and set $l = O(T^{1/4})$ in what follows. In (7) the $t_0(\tau)$ and $t_1(\tau)$ statistics are evaluated at $\hat{\tau} := \arg\max_{\tau\in\Lambda}|t_0(\tau)|$ and $\mathring{\tau} := \arg\max_{\tau\in\Lambda}|t_1(\tau)|$, respectively. Finally, λ is a weight function given by

$$\lambda := \exp\{-(g_S S_0 S_1)^2\} \tag{9}$$

with g_S a positive finite constant and

$$S_0 := \frac{\sum_{t=1}^T(\sum_{i=1}^t \hat{u}_{\hat{\tau},i})^2}{T^2\hat{\omega}^2(\hat{u}_{\hat{\tau},t})}, \qquad S_1 := \frac{\sum_{t=2}^T(\sum_{i=2}^t \check{v}_{\mathring{\tau},i})^2}{(T-1)^2\hat{\omega}^2(\check{v}_{\mathring{\tau},t})}.$$

The t_λ pre-test is conducted using a sample size-dependent critical value $cv_{t_\lambda,T}$, which shrinks the Type 1 error of t_λ towards zero with increasing sample size (i.e. $cv_{t_\lambda,T} \to \infty$ as $T \to \infty$), but also chosen such that t_λ remains a consistent test. The pre-test-based HHLT procedure is then

$$HHLT_{t_\lambda} := \begin{cases} DF_t^{GLS} & \text{if } t_\lambda < cv_{t_\lambda,T} \\ DF_{tb}^{GLS}(\tilde{\tau}) & \text{if } t_\lambda \geq cv_{t_\lambda,T} \end{cases}. \tag{10}$$

In keeping with the $HHLT_{\bar{\tau}}$ approach, standard critical values are used for DF_t^{GLS} (along with $\bar{c} = 13.5$), and for $DF_{tb}^{GLS}(\tilde{\tau})$, critical values and $\bar{c}_\tau$ values associated with a known break fraction are employed.

3.2 The CKP Test

In the CKP procedure, the auxiliary statistic used for trend break detection is the pre-test of Kejriwal and Perron (2010). Since here we are considering only a single trend break, this reduces to performing the PY test for one break. The PY statistic, which builds on the trend testing approach of Perron and Yabu (2009a), has the form

$$PY := \log[T^{-1}\sum_{\tau\in\Lambda}\exp\{\tfrac{1}{2}W_{RQF}(\tau)\}]$$

where

$$W_{RQF}(\tau) := \frac{S(\tilde{\rho}_{MS}) - S(\tilde{\rho}_{MS},\tau)}{\hat{h}_{\varepsilon}}$$

with

$$\tilde{\rho}_{MS} := \begin{cases} \tilde{\rho}_M & \text{if } |\tilde{\rho}_M - 1| > T^{-1/2} \\ 1 & \text{if } |\tilde{\rho}_M - 1| \leq T^{-1/2} \end{cases} \tag{11}$$

where $\hat{h}_{\varepsilon}$ denotes the "AN" long-run variance estimator detailed in section 4 of PY (the exact form of which depends on whether $\tilde{\rho}_{MS} = \tilde{\rho}_M$ or $\tilde{\rho}_{MS} = 1$), and $\tilde{\rho}_M$ is the Roy and Fuller (2001) bias-corrected estimator of ρ_T.[5] The latter has the form

$$\tilde{\rho}_M := \tilde{\rho} + C(t_{\hat{\phi}})\hat{\sigma}_{\hat{\phi}}$$

with $\tilde{\rho} := 1 + \hat{\phi}$ obtained from the fitted ADF regression

$$\Delta\hat{u}_{\tau,t} = \hat{\phi}\hat{u}_{\tau,t-1} + \sum_{j=1}^{p}\hat{\delta}_j\Delta\hat{u}_{\tau,t-j} + \hat{\eta}_t, \quad t = p+2,...,T \tag{12}$$

and $\hat{\sigma}_{\hat{\phi}}$ and $t_{\hat{\phi}}$ denote the standard error of $\hat{\phi}$ and the t-ratio on $\hat{\phi}$, respectively. Here, $C(t_{\hat{\phi}})$ is a four-regime step function, the precise detail of which can be found in section 2.5 of PY. As with $HHLT_{t_\lambda}$, the PY pre-test is conducted using a critical value $cv_{PY,T}$ which shrinks the Type 1 error of PY towards zero with T, but again chosen such that the test remain consistent. Depending on the outcome of this pre-test for a break in trend, a corresponding GLS de-trended unit root test[6] is then applied which either includes or excludes a broken trend.

The LR-based CKP test statistic is then given by

$$CKP := \begin{cases} P_t^{GLS} & \text{if } PY < cv_{PY,T} \\ P_{tb}^{GLS}(\breve{\tau}) & \text{if } PY \geq cv_{PY,T} \end{cases} \tag{13}$$

where P_t^{GLS} is the feasible point optimal unit root test considered in ERS, i.e.

$$P_t^{GLS} := \frac{S(\bar{\rho}) - \bar{\rho}S(1)}{\tilde{\omega}^2}$$

[5] Again, we suppress the dependence of the estimators on τ.

[6] A corresponding approach based on ADF unit root tests formed from OLS de-trended data is considered in Kim and Perron (2009).

with $\tilde{\omega}^2$, based on the estimated regression (5), given by

$$\tilde{\omega}^2 := \frac{(T-p-1)^{-1}\sum_{t=p+2}^{T}\hat{\eta}_t^2}{(1-\sum_{j=1}^{p}\hat{\delta}_j)^2}$$

and $\bar{c} = 13.5$. For a generic τ, $P_{tb}^{GLS}(\tau)$ is given by

$$P_{tb}^{GLS}(\tau) := \frac{S(\bar{\rho}_\tau, \tau) - \bar{\rho}_\tau S(1,\tau)}{\tilde{\omega}^2(\tau)}$$

where $\tilde{\omega}^2(\tau)$ is defined as for $\tilde{\omega}^2$ above, but based on the estimated regression (6).[7] As with the HHLT approach, values of $\bar{c}_\tau$ are chosen according to τ, and can again be obtained from either Table 1 of HHLT or Table 1 of CKP. These expressions are evaluated at $\tau = \breve{\tau}$, where $\breve{\tau}$ is an estimate of the true break fraction τ_0 based on a GLS regression, i.e.

$$\breve{\tau} := \arg\min_{\tau\in\Lambda} S(\bar{\rho}_\tau, \tau).$$

Standard critical values are used for P_t^{GLS}, and critical values associated with a known break fraction are used for $P_{tb}^{GLS}(\breve{\tau})$, as provided by CKP.

3.3 Finite Sample Size and Power

We now consider the finite sample behaviour of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and *CKP*, using a set of size and power simulations based on the DGP (1)-(2). For a given sample size T and break fraction τ_0, we compute the size ($\rho_T = 1$) and size-adjusted power ($\rho_T = 1 - c/T$, $c > 0$) of nominal 0.05-level tests across break magnitudes ranging from $\gamma_T = 0$ to settings that ensure the break is obvious and very easily detected by the relevant procedures ($\gamma_T \approx 1.2$ for $T = 150$ and $\gamma_T \approx 0.9$ for $T = 300$ with IID errors). Specifically, and in line with the Pitman drift rate detailed in Section 2, we let $\gamma_T = \kappa\omega_\varepsilon T^{-1/2}$ with $\kappa = \{0, ..., 15\}$. We conduct simulations for the sample sizes $T = 150$ and $T = 300$, basing the results on 5,000 Monte Carlo replications.

The tests are conducted using 15% trimming, i.e. $\Lambda = [0.15, 0.85]$, and are implemented using the asymptotic critical values and $\bar{c}_\tau$ values reported in Table 1 of HHLT, along with similarly simulated asymptotic critical values based on the limit distributions of P_t^{GLS} (see ERS) and $P_{tb}^{GLS}(\tau)$ (see CKP) with a known date (i.e. $\tau = \tau_0$), as required for the *CKP* procedure.[8] The tests are computed as described above, with $g_W = 1.5$ in (3) - this setting delivered the test with the best size control in HHLT,

[7] It is worth noting from the results reported in ERS and PR that the asymptotic local power functions of P_t^{GLS} and DF_t^{GLS}, and indeed those of the corresponding M-type unit root tests, are virtually indistinguishable from one another, as are those of $P_{tb}^{GLS}(\tau)$, $DF_{tb}^{GLS}(\tau)$ and the analogous M-type tests, and so in practice it does not matter to any great extent which of these unit root tests we use in computing the *HHLT* and *CKP* procedures.

[8] Given that the procedures use estimated break fractions, in each replication we obtained the $\bar{c}_\tau$ values and critical values by linear interpolation between the two nearest grid points in τ.

along with the HHLT settings of $l = \lfloor 4(T/100)^{1/4} \rfloor$ in (8) and $g_S = 500$ in (9); the break detection procedures t_λ and PY implicit in $HHLT_{t_\lambda}$ and CKP, respectively, are implemented at the nominal 0.05-level for both sample sizes (i.e. the significance level is not shrunk towards zero in T here). For each test and combination of DGP settings, we report the maximum size observed across $\kappa = \{0, ..., 15\}$, along with the value of κ for which the maximum size is obtained, denoted κ^*. Power curves across κ are then reported, and are size-adjusted by scaling the with-break and without-break unit root test critical values involved in the relevant procedure by a common factor, such that the size of the overall procedure is 0.05 when $\kappa = \kappa^*$; this same scaling is used for all values of κ.

First, we abstract from the effects of serial correlation and lag selection by investigating the behaviour of the tests for $\varepsilon_t \sim NIID(0, 1)$, with $p = 0$ assumed in the ADF-type regressions. We set $\mu = \beta = 0$ and consider $\tau_0 = \{0.3, 0.5, 0.7\}$. Table 1 reports the maximum sizes of the tests (and the relevant κ^* values) across $\kappa = \{0, ..., 15\}$ for these settings, and Figures 1(a)-(c) through 4(a)-(c) report the corresponding size-adjusted powers for the two sample sizes considered and $c = \{20, 30\}$. The $HHLT_{\bar{\tau}}$ and $HHLT_{t_\lambda}$ procedures are subject to noticeable over-size in small samples (for both tests the maximum size is 0.11 for $T = 150$), although these distortions do fall as the sample size increases. For CKP, we observe very little upward size distortion for $T = 150$ where it displays the best size control of all of the tests considered (including the new test procedures proposed in section 4 below), and although the distortions worsen for $T = 300$, the maximum size is at most 0.07 here; in this case $DF_{tb}^{GLS}(\tilde{\tau})$ shows the best overall size control.

Our primary focus here concerns the size-adjusted powers. When $\kappa = 0$, the powers of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP are fairly similar to one another, and capture much of the superior power available from the without-break tests when no break in trend occurs. At the other extreme, when κ is large, the power of the procedures is seen to change little in κ, since here the breaks are readily detected, and the break fraction estimators are close to the true τ_0; in this region all the tests have lower powers than for $\kappa = 0$ due to the fact that it is the with-break test that is now applied. For these large κ values, the power curves of the tests essentially only differ due to the level of size-corrections applied, with the resultant power of $HHLT_{\bar{\tau}}$ and $HHLT_{t_\lambda}$ being lower than that of CKP. The key finding in Figures 1(a)-(c) through 4(a)-(c), however, is the finite sample power behaviour of the procedures for intermediate trend break magnitudes. Here, we see clearly the power valley phenomenon, whereby power for all tests falls well below the levels associated with very small or very large break magnitudes. The relative power rankings in this valley region depend on the values of τ_0 and c: for $\tau_0 = 0.3$ and $\tau_0 = 0.5$, the power drop-off is most pronounced for CKP when $c = 20$, and for CKP and $HHLT_{t_\lambda}$ when $c = 30$, while for $\tau_0 = 0.7$ it is the $HHLT_{\bar{\tau}}$ test that displays the greatest valley. In general, we observe that the power valleys appear deeper for $c = 30$ than for $c = 20$, a feature that arises because the increase in the power of the procedures is greater for large κ than elsewhere.

Figures 1(a)-(c) through 4(a)-(c) also report powers for the without-break unit

root test DF_t^{GLS} and comparison of this test's power profile with that of (say) the HHLT tests provides some insight into the cause of the power valleys. We observe that the $HHLT_{\bar{\tau}}$ and $HHLT_{t_\lambda}$ powers closely follow that of DF_t^{GLS} for small values of κ; this occurs, of course, because the implicit break detection procedures fail to identify a break for these magnitudes. For very small κ, failure to detect a break does not significantly compromise the power of DF_t^{GLS} nor, consequently, the power of the HHLT tests; however, for intermediate magnitudes of κ we observe that while the break magnitudes are too small to be reliably detected, they can be sufficiently large to drastically reduce the power of DF_t^{GLS} and therefore the HHLT tests. For larger κ, the power profiles of the latter tests now begin to deviate upwards from that of DF_t^{GLS} since here the breaks become increasingly detectable (and dateable) and consequently the more powerful with-break unit root test is being applied. Similar comments also apply to the valley associated with *CKP*. The power valley phenomenon is therefore seen to arise from the interplay of the performance of the break detection procedures and the impact of breaks on without-break unit root tests.

To investigate the behaviour of the tests in the presence of serially correlated errors, we also conduct simulations with autoregressive and moving average specifications for ε_t in (2). To that end, we now let the ε_t be generated according to either an AR(1) process: $\varepsilon_t = 0.5\varepsilon_{t-1} + \eta_t$ (with $\eta_t \sim NIID(0,1)$ and $\varepsilon_1 = \eta_1$), or an MA(1) process: $\varepsilon_t = \eta_t - 0.5\eta_{t-1}$ (again with $\eta_t \sim NIID(0,1)$ and $\varepsilon_1 = \eta_1$). As noted in section 2, when a break in trend is present the break magnitude is standardised by ω_ε in each case. In the ADF regressions, we now set p to the value selected by the MAIC procedure of Ng and Perron (2001), as modified by Perron and Qu (2007), with $p_{\max} = \left\lfloor 12(T/100)^{1/4} \right\rfloor$. Focusing on the representative case of $\tau_0 = 0.5$ and $c = 30$, Table 2 reports the maximum sizes of the tests across κ (and the corresponding κ^* values), and Figures 5(a)-(c) and 6(a)-(c) present the power curves (size-adjusted for all cases where the corresponding maximum size exceeds 0.05); also reported as a benchmark case are results for IID errors, but where p is selected according to the MAIC criterion rather than set to zero. Comparing first the IID results with the corresponding $p = 0$ versions in Table 1 and Figures 3(b) and 4(b), we find that the use of lag selection lowers the maximum sizes observed for each procedure, and also reduces overall power. However, the central finding of the power valley phenomenon, and also the rankings of the different tests across κ, are unaffected by this change. Relative to the IID case, results for the AR(1) and MA(1) specifications show higher maximum sizes (although all tests have maximum empirical size less than 0.10), with the best overall size properties apparently shown by $DF_{tb}^{GLS}(\tilde{\tau})$, and while the overall picture of the power valley phenomenon remains, we observe that the valleys are relatively more exaggerated in the AR(1) case, and relatively less pronounced for the MA(1) case considered.

To provide some perspective of the potential seriousness of the power valleys under intermediate magnitude trend breaks at a practical level, Figure 7(a) presents a realisation from the DGP (1)-(2) for $T = 300$ when $\varepsilon_t \sim NIID(0,1)$. Here we set when $\tau_0 = 0.5$ a break magnitude of $\kappa = 2$ (i.e. $\gamma_T \approx 0.12$), with $\mu = \beta = 0$ and $c = 20$ (i.e.

$\rho_T = 1 - c/T \approx 0.93$). From visual inspection, it is certainly debatable as to whether there is a trend break present in y_t; this series is fairly typical of the kind of empirical time series data for which a priori we might be unsure whether or not to allow for a trend break when conducting a unit root test. However, it is also just this sort of specification for which $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP have very low power, according to Figure 2(b). Figure 7(b) shows the same realisation but with a larger local trend break of $\kappa = 6$ (i.e. $\gamma_T \approx 0.35$). Here, from Figure 2(b), $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP have essentially emerged from the power valley region. However, we observe that the trend break now clearly dominates the behaviour of y_t, and there is no real uncertainty as to whether a trend break should be incorporated when testing for a unit root. The point being made here is that these procedures are designed specifically for cases where there is genuine uncertainty over the presence of a break; unfortunately, however, it is in precisely this range of break magnitudes that the procedures lack power. The region of trend break magnitudes where this arises is far from a theoretical irrelevance; on the contrary, it represents a region that is likely to be really rather important to those analyzing empirical data.

4 Alternative Procedures

It is difficult to envisage alternative procedures that retain the (asymptotic) efficiency properties of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP in the cases of no break or a fixed magnitude break, but are not susceptible to the finite sample power valley problem observed in Section 3, since all such procedures must ultimately rely on an imperfect form of break detection in deciding whether to apply a without-break or with-break unit root test. One option would be to always implement $DF_{tb}^{GLS}(\tilde{\tau})$ or $P_{tb}^{GLS}(\check{\tau})$, cf. PR, thereby removing completely the potential for low power that can arise from applying a without-break unit root test. Such an approach is essentially the same as implementing HHLT- or CKP-type procedures but with the break detection step biased entirely in favour of "detecting" a break. The inherent disadvantage of this approach is that in order to control size across all possible break magnitudes (including $\gamma_T = 0$) the with-break unit root test must be applied with a conservative critical value obtained under the no-break DGP (where $\tilde{\tau}$ and $\check{\tau}$ are $O_p(1)$ variates with support on Λ).

To illustrate the performance of this procedure, Tables 1-2 and Figures 1(d)-(f) through 6(d)-(f) report the maximum sizes, and size-adjusted powers, respectively, of $DF_{tb}^{GLS}(\tilde{\tau})$, respectively, when implemented using an asymptotic conservative critical value (denoted cv_{tb}^{consv}); this test is essentially that proposed by PR, but adopting the $\tilde{\tau}$ break fraction estimator.[9] As expected, this procedure delivers decent finite sample size control and clearly avoids the power valley problem since a break is always fitted (i.e. no decision is made whether or not to fit a break on the basis of pre-testing). However, its power is considerably lower than that of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP when the break in trend is absent, since there is now no option to apply a without-break unit root test,

[9] At the nominal 0.10-, 0.05- and 0.01-levels, cv_{tb}^{consv} is -3.44, -3.72 and -4.26.

and is also lower when the break is large due to the use of the conservative critical value.

We now examine two alternative procedures which are designed to help mitigate the power valley phenomenon, but without taking recourse to simply applying only $DF_{tb}^{GLS}(\tilde{\tau})$ or $P_{tb}^{GLS}(\check{\tau})$ with conservative critical values. Our procedures are motivated more by pragmatic considerations, rather than by those of achieving asymptotic efficiency in all possible circumstances, given the inherent practical disadvantages associated with the latter approach. Additionally, they do not involve the computation of any statistics not already required as part of the $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP procedures.

4.1 $DF_{tb}^{GLS}(\tilde{\tau})$ with Adaptive Critical Values

As noted above, a disadvantage of using $DF_{tb}^{GLS}(\tilde{\tau})$ is that to control size in the event that no break occurs, conservative critical values must be used. These have the inherent cost that when a break of reasonable magnitude occurs, power is lost in comparison to using critical values associated with a known break fraction. A possibility we consider here then is to adapt the critical values of $DF_{tb}^{GLS}(\tilde{\tau})$ according to the outcome of a break detection pre-test, generically labelled B, with an associated critical value cv_B. Denoting the asymptotic known break fraction critical value corresponding to $\tilde{\tau}$ by $cv_{tb}^{\tilde{\tau}}$, we define the following adaptive procedure:

$$A(B) := DF_{tb}^{GLS}(\tilde{\tau}) \text{ with critical value } \begin{cases} cv_{tb}^{\text{consv}} & \text{if } B < cv_B \\ cv_{tb}^{\tilde{\tau}} & \text{if } B \geq cv_B \end{cases}. \tag{14}$$

Here, then, conservative critical values are applied to $DF_{tb}^{GLS}(\tilde{\tau})$ only in those cases where no trend break is detected by the pre-test B, with known date critical values used when a break has been identified. In keeping with the behaviour of $DF_{tb}^{GLS}(\tilde{\tau})$ when employing cv_{tb}^{consv} alone, this approach avoids the potential for very low power in the presence of small breaks, since the pre-test is only used to determine critical values for a with-break unit root test, and does not allow for the possibility of inappropriately applying a without-break unit root test procedure. Given that both $HHLT_{t_\lambda}$ and CKP employ robust (to the order of integration) pre-tests for a break in trend, it makes sense to also use these procedures for B in (14); in what follows, we therefore consider two versions of $A(B)$, namely $A(t_\lambda)$ using the Harvey et al. (2009b) test, and $A(PY)$ using the PY test, with corresponding critical values cv_{t_λ} and cv_{PY}.[10]

4.2 An Adaptive Union of Rejections of DF_t^{GLS} and $DF_{tb}^{GLS}(\tilde{\tau})$

The second procedure we propose is in the spirit of work of Harvey *et al.* (2009a), and attempts to capture some of the additional power associated with DF_t^{GLS} when there

[10]Note that we use fixed critical values for the pre-tests, rather than the sample size-dependent versions used in in $HHLT_{t_\lambda}$ and CKP.

is no break in trend, while excluding the possibility of only implementing DF_t^{GLS} when a break is present. Denoting the asymptotic critical value associated with DF_t^{GLS} by cv_t[11], the starting point is a "union of rejections" decision rule:

$$U := \text{Reject } H_0 \text{ if } \left\{ DF_t^{GLS} < cv_t \text{ or } DF_{tb}^{GLS}(\tilde{\tau}) < cv_{tb}^{\text{consv}} \right\}$$

whereby the unit root null is rejected if either DF_t^{GLS} or $DF_{tb}^{GLS}(\tilde{\tau})$ rejects. The overall decision rule U will be asymptotically over-sized (for zero and non-zero breaks) due to the combination of two tests. To control size under zero and fixed magnitude breaks, we apply a common scaling constant to cv_t and cv_{tb}^{consv}, i.e.

$$U^{\delta} := \text{Reject } H_0 \text{ if } \left\{ DF_t^{GLS} < \delta cv_t \text{ or } DF_{tb}^{GLS}(\tilde{\tau}) < \delta cv_{tb}^{\text{consv}} \right\}$$

where δ (> 1) is evaluated under H_0 in the no trend break case (where the size of U is at a maximum). For a given significance level for U^{δ}, the value of δ can be derived as follows, using the procedure in the rejoinder to the commentaries in Harvey *et al.* (2009a). First note that U^{δ} can alternatively be expressed as

$$U^{\delta} := \text{Reject } H_0 \text{ if } \left\{ DF_{UR}^{GLS} := \min\left(DF_t^{GLS}, \frac{cv_t}{cv_{tb}^{\text{consv}}} DF_{tb}^{GLS}(\tilde{\tau}) \right) < \delta cv_t \right\}.$$

The appropriate constant δ can then be determined by simulating the limit distribution of DF_{UR}^{GLS}, calculating the asymptotic critical value for this empirical distribution, say cv_{UR}, and then computing $\delta := cv_{UR}/cv_t$. We obtained constants in this way at the 0.10, 0.05 and 0.01 nominal significance levels, simulating the limit distributions of DF_t^{GLS} and $DF_{tb}^{GLS}(\tau)$ given in ERS and PR, respectively, with τ evaluated using the limit representation of $\tilde{\tau}$

$$\tilde{\tau} \xrightarrow{d} \arg\sup_{\tau \in \Lambda} \begin{bmatrix} W_c(1) \\ W_c(1) - W_c(\tau) \end{bmatrix}' \begin{bmatrix} 1 & (1-\tau) \\ (1-\tau) & (1-\tau) \end{bmatrix}^{-1} \begin{bmatrix} W_c(1) \\ W_c(1) - W_c(\tau) \end{bmatrix}$$

where $W_c(r) := \int_0^r e^{-(r-s)c} dW(s)$, with $W(r)$ a standard Wiener process (see the proof of Theorem 3(i) below with $\kappa = 0$). Here and throughout the paper, all asymptotic simulations were conducted using 50000 Monte Carlo replications, and by approximating the Wiener processes using $NIID(0,1)$ random variates, with the integrals approximated by normalized sums of 2000 steps. At the nominal 0.10-, 0.05- and 0.01-levels, the appropriate values are $\delta = 1.092$, $\delta = 1.065$ and $\delta = 1.029$.

It is possible, however, to further modify this simple union of rejections to capture additional power when a break is present. Using a related strategy to that adopted in the case of the adaptive critical values procedure in A of (14) above, if there is clear evidence of a break in trend, it is unnecessary to include DF_t^{GLS} in the union, and also to use the conservative critical value for $DF_{tb}^{GLS}(\tilde{\tau})$. Instead, as with $A(B)$, in these cases we can simply apply $DF_{tb}^{GLS}(\tilde{\tau})$ with a critical value associated with a known

[11] At the nominal 0.10-, 0.05- and 0.01-levels, cv_t is -2.56, -2.84 and -3.41 respectively

break fraction. Using the break pre-test B to decide between these possibilities, we can specify an adaptive union of rejections decision rule as follows:

$$U_B^{\delta} := \begin{cases} \text{Reject } H_0 \text{ if } \left\{ DF_t^{GLS} < \delta cv_t \text{ or } DF_{tb}^{GLS}(\tilde{\tau}) < \delta cv_{tb}^{\text{consv}} \right\} & \text{if } B < cv_B \\ \text{Reject } H_0 \text{ if } \left\{ DF_{tb}^{GLS}(\tilde{\tau}) < cv_{tb}^{\tilde{\tau}} \right\} & \text{if } B \geq cv_B \end{cases} \quad (15)$$

where the δ are those values calculated for U^{δ}. As with $A(B)$ we consider $B = t_{\lambda}$ and $B = PY$ as candidate pre-tests for constructing U_B^{δ}, leading to two alternative versions of the procedure.

4.3 Finite Sample Size and Power

We now consider the finite sample behaviour of the alternative procedures $A(t_{\lambda})$, $A(PY)$, $U_{t_{\lambda}}^{\delta}$ and U_{PY}^{δ}, using the same size and power simulations employed in Section 3.3. The same settings for the DGP and the implementation of the tests are employed, and Tables 1-2 and Figures 1(d)-(f) through 6(d)-(f) report the maximum sizes, and powers, of the tests, respectively. The powers are size-adjusted using the same methodology as before, scaling the unit root test critical values (i.e. cv_{tb}^{consv} and $cv_{tb}^{\tilde{\tau}}$ in $A(B)$; cv_t, cv_{tb}^{consv} and $cv_{tb}^{\tilde{\tau}}$ in U_B^{δ}) by a common factor, such that the size of the overall procedure is 0.05 when $\kappa = \kappa^*$ (note that in the case of U_B^{δ}, this is in addition to the critical value scaling δ on cv_t and cv_{tb}^{consv}); as before, this same scaling is then used across all κ.

For the cases of IID errors with $p = 0$, the maximum sizes of the alternative procedures that employ the t_{λ} pre-test are in the region of 0.10 for $T = 150$ and 0.08-0.09 for $T = 300$, while those based on PY are around 0.08 for $T = 150$ and 0.07-0.08 for $T = 300$. The sizes can therefore be somewhat more distorted than the CKP test, but are at least as well behaved as the HHLT procedures.

As expected, and in line with the results for (conservative) $DF_{tb}^{GLS}(\tilde{\tau})$, the power curves for the adaptive procedures $A(t_{\lambda})$ and $A(PY)$ do not display valleys, and have minimum powers across κ that are in many cases considerably higher than those for $HHLT_{\bar{\tau}}$, $HHLT_{t_{\lambda}}$ and CKP. For large κ, $A(PY)$ has size-adjusted power similar to that of CKP, and therefore higher than $HHLT_{\bar{\tau}}$ and $HHLT_{t_{\lambda}}$, while the power of $A(t_{\lambda})$ is closer to the $HHLT_{\bar{\tau}}$ and $HHLT_{t_{\lambda}}$ powers in these cases, due to the greater degree of size correction required for this procedure. It can also be seen that $A(t_{\lambda})$ and $A(PY)$ have generally more appealing power profiles than $DF_{tb}^{GLS}(\tilde{\tau})$. As expected, the power gains are most marked for the larger values of κ where the pre-tests reject in favour of a break, although some power losses do exist for small κ due to the fact that $A(t_{\lambda})$ and $A(PY)$ require greater size-adjustment. Comparing $A(t_{\lambda})$ with $A(PY)$, we observe that the $A(PY)$ procedure has the greater power for small and large trend break magnitudes, with the ranking reversed for intermediate values of κ. Overall, the picture is one of greater robustness to κ compared to the power profiles of $HHLT_{\bar{\tau}}$, $HHLT_{t_{\lambda}}$ and CKP; the price of this robustness is seen when $\kappa = 0$ or very small, where $A(t_{\lambda})$ and $A(PY)$ lose power relative to the tests of HHLT and CKP, due to their lack of an option to implement a without-break unit root test. Qualitatively similar

comments apply to the results for AR(1) and MA(1) errors, although in the moving average case, $A(t_\lambda)$ and $A(PY)$ have similar power to each other for intermediate and large breaks.

The $U^\delta_{t_\lambda}$ and U^δ_{PY} procedures also display a lesser degree of sensitivity to the magnitude of κ than the original $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP tests. Relative to $A(t_\lambda)$ and $A(PY)$ (and $DF^{GLS}_{tb}(\tilde{\tau})$), these union of rejections-based procedures have much higher power when no break is present, capturing much of the power attained by $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP tests at this point due to the inclusion of DF^{GLS}_t in the decision rule. For large break magnitudes, U^δ_{PY} retains the high power enjoyed by both CKP and $A(PY)$; $U^\delta_{t_\lambda}$ also displays decent power in this region, roughly achieving the (somewhat lower) level of power seen for $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and $A(t_\lambda)$. In the intermediate range of κ values, $U^\delta_{t_\lambda}$ and U^δ_{PY} also display power valleys; for these procedures, the valleys do not arise as a result of an inappropriate use of DF^{GLS}_t, but instead because for intermediate local break magnitudes, they are essentially applying $DF^{GLS}_{tb}(\tilde{\tau})$ with a doubly scaled conservative critical value, unsurprisingly reducing the power in these cases. However, the valleys are less pronounced relative to the tests of HHLT and CKP, and the minimum powers across κ exceed those of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP (with the exception of U^δ_{PY} which can have power lower than the minimum of $HHLT_{\bar{\tau}}$ when $c = 20$ and $\tau_0 = 0.3, 0.5$). Comparing $U^\delta_{t_\lambda}$ and U^δ_{PY}, $U^\delta_{t_\lambda}$ has relatively lower power for small and large κ, but manifests a smaller power valley for intermediate break magnitudes. Qualitatively similar comments apply in the case of AR(1) errors; in the presence of MA(1) dynamics, $U^\delta_{t_\lambda}$ now outperforms U^δ_{PY} for most values of κ, and the benefits relative to CKP are less clear-cut.

Overall, we find that the alternative testing procedures proposed offer a reasonable degree of robustness to the break magnitudes in terms of size-adjusted power, and in doing so provide potential improvements over the HHLT and CKP testing strategies. Of the different $A(B)$ and U^δ_B approaches, the $A(t_\lambda)$ and $A(PY)$ procedures offer the greater robustness to κ combined with the most appealing minimum power values, while $U^\delta_{t_\lambda}$ and U^δ_{PY} deliver superior power gains for small and zero break magnitudes, at the cost of some power losses for intermediate break magnitudes.

5 Local Asymptotic Behaviour

In this section we examine the asymptotic behaviour of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP, together with the newly proposed procedures $A(t_\lambda)$, $A(PY)$, $U^\delta_{t_\lambda}$ and U^δ_{PY}, with the purpose of adequately mimicking the observed finite sample power properties of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP in the limit. To do this, we must clearly dispense with the fixed magnitude trend break assumption of HHLT and CKP, since such an approach only adequately models the cases of either a zero break or a large, essentially perfectly detectable break, and does not provide any asymptotic prediction of the power valleys phenomenon seen in section 3.3. Instead, we consider the asymptotic behaviour of the tests in a doubly-local setting where, in addition to allowing local-to-unity behaviour

in the autoregressive root, we adopt the local trend break magnitude assumption of Section 2 setting $\gamma_T = \kappa\omega_\varepsilon T^{-1/2}$. This framework ensures that the magnitude of the local trend, κ, appears in the limit distributions of the aforementioned statistics such that the asymptotic theory can then form a useful approximation to the finite sample distributions of these statistics (as we shall demonstrate in section 5.6 when we simulate these local limiting distributions). However, the local nature of the autoregressive root and trend raises a number of calibration issues which warrant careful consideration before we proceed with the asymptotic analysis. We discuss these now.

Consider first the $HHLT_{\bar{\tau}}$ statistic. Here the local trend break necessarily renders $W_T(\tilde{\tau})$ of order $O_p(1)$. As a consequence, $T^{-1/2}W_T(\tilde{\tau}) \xrightarrow{p} 0$ and therefore $\bar{\lambda} = \exp\{-g_W T^{-1/2} W_T(\tilde{\tau})\} \xrightarrow{p} 1$. Hence, $\bar{\tau} \xrightarrow{p} 0$ and in the limit $HHLT_{\bar{\tau}}$ will always reduce to the without-break unit root test statistic DF_t^{GLS} under a local break in trend. Therefore, as it stands, $DF_{tb}^{GLS}(\bar{\tau})$ is precluded from influencing the asymptotic local behaviour of $HHLT_{\bar{\tau}}$. This is clearly an unsatisfactory state of affairs since our intention here is to conduct a local asymptotic analysis to proxy finite sample behaviour, where DF_t^{GLS} and $DF_{tb}^{GLS}(\bar{\tau})$ should clearly both come into play. To this end, in what follows we replace the $g_W T^{-1/2}$ scaling on $W_T(\tilde{\tau})$ with a positive constant g'_W, which ensures that $\bar{\lambda}$ is not degenerate at 1 in the limit, but instead has a proper distribution. Notice that in this modified framework, the constant g'_W will now have an asymptotic effect, unlike g_W in the original formulation of $HHLT_{\bar{\tau}}$.

For $HHLT_{t_\lambda}$, HHLT show that when $\kappa = 0$ in a local to unit root setting, $t_0(\hat{\tau}) = O_p\{(T/l)^{1/2}\}$ and $S_0 = O_p(T/l)$. The latter result implies $\lambda \xrightarrow{p} 0$ at an exponentially fast rate, giving $t_\lambda \xrightarrow{p} m_\xi |t_1(\mathring{\tau})|$. The same rates also pertain under a local trend break ($\kappa \neq 0$). There is consequently no flexibility for t_λ to reflect any less persistent autoregressive (i.e. local to unit root rather than exact unit root) behaviour, which again sits rather uncomfortably if we wish our asymptotics to mirror finite sample behaviour. We therefore replace g_S in (9) with $g'_S(T/l)^{-1}$, and then also replace $t_0(\hat{\tau})$ in (7) with $g'_t(T/l)^{-1/2}t_0(\hat{\tau})$, where g'_S and g'_t are positive constants. This ensures that $\lambda = O_p(1)$ and also keeps $t_\lambda = O_p(1)$, such that t_λ then uses a combination of $|t_0(\hat{\tau})|$ and $|t_1(\mathring{\tau})|$ in the limit, with the constants g'_S and g'_t having an asymptotic effect. An additional consideration is the fact that since the t_λ pre-test is found to be $O_p(1)$ under a local trend break, if it is implemented with a sample-size dependent critical value such that $cv_{t_\lambda,T} \to \infty$ as $T \to \infty$, then in the limit $HHLT_{t_\lambda}$ will always reduce to the without-break unit root test statistic DF_t^{GLS}. To ensure that $DF_{tb}^{GLS}(\tilde{\tau})$ also enters the local asymptotic analysis, in what follows we replace $cv_{t_\lambda,T}$ with a fixed critical value, denoted cv_{t_λ}. The same issues also apply to the use of t_λ in $A(t_\lambda)$ and $U_{t_\lambda}^{\delta}$, hence we adopt the aforementioned modifications in establishing the limit behaviour of t_λ in these contexts also.

Turning to *CKP*, we find that the local-to unity-autoregressive results in $T|\tilde{\rho}_M - 1| = O_p(1)$. Since we may write (11) in the equivalent form

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

it is clear that $\tilde{\rho}_{MS} \xrightarrow{p} 1$ and so, asymptotically, PY will always be based on $W_{RQF}(\tau)$ calculated with $\tilde{\rho}_{MS} = 1$. Since we again find that the pre-test statistic is unable to capture any local to unit root behaviour, in parallel with the modification to t_λ, we instead set

$$\tilde{\rho}_{MS} := \begin{cases} \tilde{\rho}_M & \text{if } T\,|\tilde{\rho}_M - 1| > g'_M \\ 1 & \text{if } T\,|\tilde{\rho}_M - 1| \leq g'_M \end{cases} \tag{16}$$

where g'_M is a positive constant. We can then choose g'_M (which will enter the limit distribution of PY) to ensure that both $\tilde{\rho}_M$ and 1 are represented in the limit distribution of $\tilde{\rho}_{MS}$. Finally, since PY is $O_p(1)$ under a local trend break, use of a sample-size dependent critical value such that $cv_{PY,T} \to \infty$ as $T \to \infty$ results in CKP reducing to the without-break unit root test statistic P_t^{GLS} in the limit. To ensure that $P_{tb}^{GLS}(\breve{\tau})$ is also present in the asymptotic analysis, as is suggested for the $HHLT_{t_\lambda}$ statistic above, we replace $cv_{PY,T}$ with a fixed critical value cv_{PY} . As with t_λ, we also adopt these PY modifications when deriving the local limit representations for $A(PY)$ and U_{PY}^{δ}.

We now establish the limiting properties of the constituent components of the tests under our doubly-local framework. We begin with DF_t^{GLS} and P_t^{GLS}, followed by their with-trend break counterparts evaluated at a generic break fraction. Next we consider the break fraction estimators involved in the construction of the statistics. These limit analyses can be conducted without reference to the modifications discussed above. These modifications do become important, however, once we move forward to discuss the limit behaviour of the break detection elements, and subsequently the limits of the $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$, CKP, $A(t_\lambda)$, $A(PY)$, $U_{t_\lambda}^{\delta}$ and U_{PY}^{δ} procedures in their entirety under the doubly-local asymptotic framework.[12]

5.1 Limits of DF_t^{GLS} and P_t^{GLS}

We first establish the limits of the without-break unit root tests DF_t^{GLS} and P_t^{GLS} under a neglected local break in trend. The asymptotic behaviour of these statistics is given in the following theorem.

Theorem 1 *Let y_t be generated according to (1)-(2) under Assumption 1. Then,*

(i)

$$DF_t^{GLS} \xrightarrow{d} \frac{K_{c,\bar{c}}(1,\tau_0,\kappa)^2 - 1}{2\sqrt{\int_0^1 K_{c,\bar{c}}(r,\tau_0,\kappa)^2 dr}} =: \mathcal{D}_{c,\bar{c}}^{DF_t}(\tau_0,\kappa)$$

where

$$K_{c,\bar{c}}(r,\tau_0,\kappa) := W_c(r) + \kappa(r-\tau_0)\mathbb{I}_{\tau_0}^r - \{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\}r/a_{\bar{c}}$$

[12]Notice, however, that these modifications are made only for the purpose of calibrating the resulting local asymptotic distributions to the finite sample distributions observed previously. Practical calculation of the statistics should still be done as outlined in sections 3 and 4.

with

$$\begin{aligned} b_{c,\bar{c}} &:= (1+\bar{c})W_c(1) + \bar{c}^2\textstyle\int_0^1 sW_c(s)dr, \\ f_{c,\bar{c}}(\tau_0) &:= (1-\tau_0)\left\{a_{\bar{c}} - \bar{c}^2\tau_0(1+\tau_0)/6\right\}, \\ a_{\bar{c}} &:= 1+\bar{c}+\bar{c}^2/3 \end{aligned}$$

and $W_c(r) := \int_0^r e^{-(r-s)c}dW(s)$, $W(r)$ *a standard Wiener process.*

(ii)

$$\begin{aligned} P_t^{GLS} \xrightarrow{d} & \left\{b_{c,0} + \kappa f_{c,0}(\tau_0)\right\}^2 - \left\{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\right\}^2/a_{\bar{c}} \\ & + \bar{c}^2\textstyle\int_0^1 W_c(r)^2dr + \bar{c}W_c(1)^2 + \kappa j_{c,\bar{c}}(\tau_0) + \kappa^2 k_{\bar{c}}(\tau_0) =: \mathcal{D}_{c,\bar{c}}^{P_t}(\tau_0,\kappa) \end{aligned}$$

where

$$\begin{aligned} j_{c,\bar{c}}(\tau_0) &:= 2\bar{c}^2\textstyle\int_{\tau_0}^1 (r-\tau_0)W_c(r)dr + 2\bar{c}\int_{\tau_0}^1 (r-\tau_0)dW_c(r) + 2\bar{c}\int_{\tau_0}^1 W_c(r)dr, \\ k_{\bar{c}}(\tau_0) &:= \bar{c}^2(1-3\tau_0+3\tau_0^2-\tau_0^3)/3 + \bar{c}(1-\tau_0)^2. \end{aligned}$$

Remark 1: Notice that for $\kappa = 0$ the representations given in Theorem 1 reduce to those given in, for example, ERS, for the limit distributions of DF_t^{GLS} and P_t^{GLS} for the without-break case, as would be expected. When $\kappa \neq 0$, however, the statistics continue to be of $O_p(1)$, but are now dependent on both the timing and magnitude of the neglected local break in trend.

5.2 Limits of $DF_{tb}^{GLS}(\tau)$ and $P_{tb}^{GLS}(\tau)$

Next we determine the limits of the with-break unit root tests $DF_{tb}^{GLS}(\tau)$ and $P_{tb}^{GLS}(\tau)$, implemented for some generic break fraction, τ, which may be different from the true break fraction, τ_0. In what follows it is also convenient to define $\mathbb{I}_j^i := 1(i > j)$.

Theorem 2 *Let the conditions of Theorem 1 hold. Then, for any* $\tau \in \Lambda$,

(i)

$$DF_{tb}^{GLS}(\tau) \xrightarrow{d} \frac{L_{c,\bar{c}_\tau}(1,\tau_0,\tau,\kappa)^2 - 1}{2\sqrt{\int_0^1 L_{c,\bar{c}_\tau}(r,\tau_0,\tau,\kappa)^2 dr}} =: \mathcal{D}_{c,\bar{c}_\tau}^{DF_{tb}}(\tau_0,\tau,\kappa)$$

where

$$\begin{aligned} L_{c,\bar{c}_\tau}(r,\tau_0,\tau,\kappa) := \; & W_c(r) + \kappa(r-\tau_0)\mathbb{I}_{\tau_0}^r \\ & - \begin{bmatrix} r \\ (r-\tau)\mathbb{I}_\tau^r \end{bmatrix}' \begin{bmatrix} a_{\bar{c}_\tau} & m_{\bar{c}_\tau}(\tau) \\ m_{\bar{c}_\tau}(\tau) & d_{\bar{c}_\tau}(\tau) \end{bmatrix}^{-1} \begin{bmatrix} b_{c,\bar{c}_\tau} + \kappa f_{c,\bar{c}_\tau}(\tau_0) \\ b_{c,\bar{c}_\tau}(\tau) + \kappa f_{c,\bar{c}_\tau}(\tau_0,\tau) \end{bmatrix} \end{aligned}$$

with

$$\begin{aligned}
b_{c,\bar{c}_\tau}(\tau) &:= (1+\bar{c}_\tau-\bar{c}_\tau\tau)W_c(1)-W_c(\tau)+\bar{c}_\tau^2\textstyle\int_\tau^1(s-\tau)W_c(s)ds,\\
f_{c,\bar{c}_\tau}(\tau_0,\tau) &:= (1-\tau_0)\{a_{\bar{c}_\tau}-\bar{c}_\tau\tau-\bar{c}_\tau^2\tau(1-\tau_0)/2-\bar{c}_\tau^2\tau_0(1+\tau_0)/6\}\\
&\quad -(\tau-\tau_0)\{1-\bar{c}_\tau^2(\tau-\tau_0)^2/6\}\mathbb{I}_{\tau_0}^{\tau},\\
m_{\bar{c}_\tau}(\tau) &:= a_{\bar{c}_\tau}-\tau(1+\bar{c}_\tau+\bar{c}_\tau^2/2-\bar{c}_\tau^2\tau^2/6),\\
d_{\bar{c}_\tau}(\tau) &:= a_{\bar{c}_\tau}-\tau(1+2\bar{c}_\tau-\bar{c}_\tau\tau+\bar{c}_\tau^2-\bar{c}_\tau^2\tau+\bar{c}_\tau^2\tau^2/3)
\end{aligned}$$

(ii)

$$\begin{aligned}
P_{tb}^{GLS}(\tau) \xrightarrow{d} &\begin{bmatrix} b_{c,0}+\kappa f_{c,0}(\tau_0) \\ b_{c,0}(\tau)+\kappa f_{c,0}(\tau_0,\tau)\end{bmatrix}'\begin{bmatrix} 1 & m_0(\tau) \\ m_0(\tau) & d_0(\tau)\end{bmatrix}^{-1}\begin{bmatrix} b_{c,0}+\kappa f_{c,0}(\tau_0) \\ b_{c,0}(\tau)+\kappa f_{c,0}(\tau_0,\tau)\end{bmatrix}\\
&-\begin{bmatrix} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) \\ b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau)\end{bmatrix}'\begin{bmatrix} a_{\bar{c}_\tau} & m_{\bar{c}_\tau}(\tau) \\ m_{\bar{c}_\tau}(\tau) & d_{\bar{c}_\tau}(\tau)\end{bmatrix}^{-1}\begin{bmatrix} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) \\ b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau)\end{bmatrix}\\
&+\bar{c}_\tau^2\textstyle\int_0^1 W_c(r)^2dr+\bar{c}_\tau W_c(1)^2+\kappa j_{c,\bar{c}_\tau}(\tau_0)+\kappa^2 k_{\bar{c}_\tau}(\tau_0) =: \mathcal{D}_{c,\bar{c}_\tau}^{P_{tb}}(\tau_0,\tau,\kappa)
\end{aligned}$$

Remark 2: It is straightforward to show that setting $\tau=\tau_0$ in the limiting representations given in Theorem 2 yields the known break fraction limiting distributions of $DF_{tb}^{GLS}(\tau)$ and $P_{tb}^{GLS}(\tau)$, as given in HHLT and CKP, respectively. In this case, the limit distributions are invariant to the break magnitude κ. However, in the more general situation where the unit root tests are computed for a break fraction τ that differs from the true break fraction τ_0 (as is required in a local break context where, as we shall see directly below, the break fraction cannot be consistently estimated), the local asymptotic distributions will, as with the without-break unit root tests DF_t^{GLS} and P_t^{GLS}, depend on the timing and magnitude of the trend break, but also on the value of τ used in computing the statistic.

5.3 Limits of $\tilde{\tau}$, $\hat{\tau}$, $\mathring{\tau}$, $\breve{\tau}$

The break fraction estimators employed in the unit root test procedures have the following limits.

Theorem 3 *Let the conditions of Theorem 1 hold. Then,*

(i)

$$\begin{aligned}
\tilde{\tau},\mathring{\tau} \xrightarrow{d} \arg\sup_{\tau\in\Lambda} &\begin{bmatrix} W_c(1)+\kappa(1-\tau_0) \\ W_c(1)-W_c(\tau)+\kappa(1-\tau_0)-\kappa(\tau-\tau_0)\mathbb{I}_{\tau_0}^{\tau}\end{bmatrix}'\begin{bmatrix} 1 & (1-\tau) \\ (1-\tau) & (1-\tau)\end{bmatrix}^{-1}\\
&\times\begin{bmatrix} W_c(1)+\kappa(1-\tau_0) \\ W_c(1)-W_c(\tau)+\kappa(1-\tau_0)-\kappa(\tau-\tau_0)\mathbb{I}_{\tau_0}^{\tau}\end{bmatrix} =: \mathcal{D}_c^{\tilde{\tau}}(\tau_0,\kappa)
\end{aligned}$$

(ii)

$$\hat{\tau} \xrightarrow{d} \arg\sup_{\tau\in\Lambda} |\mathcal{D}_c^{t_0}(\tau_0,\tau,\kappa)| =: \mathcal{D}_c^{\hat{\tau}}(\tau_0,\kappa)$$

with

$$\mathcal{D}_c^{t_0}(\tau_0,\tau,\kappa) := \frac{\int_0^1 V_c(r,\tau_0,\kappa)F(r,\tau)dr}{\sqrt{\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2dr\int_0^1 F(r,\tau)^2dr}}$$

where $V_c(r,\tau_0,\kappa)$ and $N_c(r,\tau_0,\tau,\kappa)$ denote the continuous time residual processes from the projections of $W_c(r)+\kappa\mathbb{I}_{\tau_0}^r(r-\tau_0)$ onto the space spanned by $\{1,r\}$, and $\{1,r,\mathbb{I}_\tau^r(r-\tau)\}$, respectively, and where $F(r,\tau)$ denotes the continuous time residual process from the projection of $\mathbb{I}_\tau^r(r-\tau)$ onto the space spanned by $\{1,r\}$.

(iii)

$$\begin{aligned}\breve{\tau} \xrightarrow{d} \arg\sup_{\tau\in\Lambda} &\begin{bmatrix} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) \\ b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau)\end{bmatrix}' \begin{bmatrix} a_{\bar{c}_\tau} & m_{\bar{c}_\tau}(\tau) \\ m_{\bar{c}_\tau}(\tau) & d_{\bar{c}_\tau}(\tau)\end{bmatrix}^{-1} \\ &\times \begin{bmatrix} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) \\ b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau)\end{bmatrix} - \bar{c}_\tau^2\int_0^1 W_c(r)^2dr - \bar{c}_\tau\{W_c(1)^2-1\} \\ &-\kappa j_{c,\bar{c}_\tau}(\tau_0) - \kappa^2 k_{\bar{c}_\tau}(\tau_0) - \bar{c}_\tau(1-\frac{\sigma_\varepsilon^2}{\omega_\varepsilon^2}) =: \mathcal{D}_{c,\bar{c}_\tau}^{\breve{\tau}}(\tau_0,\kappa,\sigma_\varepsilon/\omega_\varepsilon)\end{aligned}$$

Remark 3: Under our local-to-zero specification for the trend break magnitude, the break fraction estimators are not consistent for τ_0. Rather, they are $O_p(1)$ with limit distributions that depend on τ_0 and κ, and also on the choice set Λ and the local-to-unity parameter, c.[13]

Remark 4: From part (iii) of Theorem 3, we see that, and in contrast to the other break fraction estimators considered in parts (i) and (ii), the limiting distribution for the GLS-based estimator, $\breve{\tau}$, used by CKP is non-pivotal as it depends on the ratio of the short- and long-run variances, $\sigma_\varepsilon^2/\omega_\varepsilon^2$. This arises because the GLS detrending parameter $\bar{c}_\tau$ used in constructing $S(\bar{\rho}_\tau,\tau)$ is allowed to vary with τ, and it clearly represents a somewhat undesirable side-effect of using this level of sophistication for break date location. Also, the problem is an endemic one for this estimator, not one which is dependent on whether c and κ are zero or not. However, notice, crucially, that this issue would not arise if $\bar{c}_\tau$ was held constant across τ; as, for example, with $\tilde{\tau}$, where the implicit value of $\bar{c}_\tau$ is zero for all τ.

[13]It is straightforward to simplify the limit distribution for $\tilde{\tau}$ given in Theorem 3(i), giving a result in line with Yang (2011). However, we maintain use of the quadratic form representation due to its parallels with the $\breve{\tau}$ limit in Theorem 3(iii), and the fact that the $\breve{\tau}$ limit expression does not readily simplify.

5.4 Limits of $\bar{\tau}$, t_λ, PY

The final components of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$, CKP, $A(t_\lambda)$, $A(PY)$, $U^{\delta}_{t_\lambda}$ and U^{δ}_{PY} are the statistics used in the respective break detection procedures. The limit distributions of the (modified) $\bar{\tau}$, t_λ and PY statistics are given in the following theorem.

Theorem 4 *Let the conditions of Theorem 1 hold. Then,*

(i)

$$\begin{aligned}\bar{\tau} &= [1-\exp\{-g'_W W_T(\tilde{\tau})\}]\tilde{\tau} \\ &\xrightarrow{d} [1-\exp\{-g'_W J_c(\tau_0,\mathcal{D}^{\tilde{\tau}}_c(\tau_0,\kappa),\kappa)\}]\mathcal{D}^{\tilde{\tau}}_c(\tau_0,\kappa) \\ &=: \mathcal{D}^{\bar{\tau}}_c(\tau_0,\kappa,g'_W)\end{aligned}$$

where

$$J_c(\tau_0,\tau,\kappa) := \frac{\int_0^1 s_R(r,\tau_0,\kappa)^2 dr}{\int_0^1 s_U(r,\tau_0,\tau,\kappa)^2 dr} - 1$$

and $s_R(r,\tau_0,\kappa)$ and $s_U(r,\tau_0,\tau,\kappa)$ denote the continuous time residuals from the projection of $\int_0^r W_c(s)ds + \kappa \mathbb{I}^r_{\tau_0}\int_{\tau_0}^r (s-\tau_0)ds$ onto the space spanned by $\{r, \int_0^r sds\}$, and $\{r, \int_0^r sds, \mathbb{I}^r_\tau \int_\tau^r (s-\tau)ds\}$, respectively.

(ii)

$$\begin{aligned}t_\lambda &= \lambda\left|g'_t(T/l)^{-1/2}t_0(\hat{\tau})\right| + m_\xi(1-\lambda)\left|t_1(\mathring{\tau})\right|, \qquad \lambda = \exp\{-(g'_S(T/l)^{-1}S_0S_1)^2\} \\ &\xrightarrow{d} \mathcal{D}^{\lambda}_c(\tau_0,\kappa,g'_S)\left|g'_t\mathcal{D}^{t_0}_c(\tau_0,\mathcal{D}^{\hat{\tau}}_c(\tau_0,\kappa),\kappa)\right| + m_\xi(1-\mathcal{D}^{\lambda}_c(\tau_0,\kappa,g'_S))\left|\mathcal{D}^{t_1}_c(\tau_0,\mathcal{D}^{\tilde{\tau}}_c(\tau_0,\kappa),\kappa)\right| \\ &=: \mathcal{D}^{t_\lambda}_c(\tau_0,\kappa,g'_S,g'_t)\end{aligned}$$

where

$$\mathcal{D}^{\lambda}_c(\tau_0,\kappa,g'_S) := \exp[-\{g'_S\mathcal{D}^{S_0}_c(\tau_0,\mathcal{D}^{\hat{\tau}}_c(\tau_0,\kappa),\kappa)\mathcal{D}^{S_1}_c(\tau_0,\mathcal{D}^{\tilde{\tau}}_c(\tau_0,\kappa),\kappa)\}^2]$$

with $\mathcal{D}^{t_0}_c(\tau_0,\tau,\kappa)$ is as defined in Theorem 3 and

$$\begin{aligned}\mathcal{D}^{t_1}_c(\tau_0,\tau,\kappa) &:= \sqrt{\tau(1-\tau)}G_{2,c}(\tau_0,\tau,\kappa), \\ \mathcal{D}^{S_0}_c(\tau_0,\tau,\kappa) &:= \frac{\int_0^1\left(\int_0^r N_c(s,\tau_0,\tau,\kappa)ds\right)^2 dr}{\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2 dr}, \\ \mathcal{D}^{S_1}_c(\tau_0,\tau,\kappa) &:= \int_0^1 Q_c(r,\tau_0,\tau,\kappa)^2 dr\end{aligned}$$

where

$$Q_c(r,\tau_0,\tau,\kappa) := W_c(r) + \kappa(r-\tau_0)\mathbb{I}^r_{\tau_0} - G_{1,c}(\tau_0,\tau,\kappa)r - G_{2,c}(\tau_0,\tau,\kappa)(r-\tau)\mathbb{I}^r_\tau$$

with

$$\begin{aligned} G_{1,c}(\tau_0,\tau,\kappa) &:= \tau^{-1}\{W_c(\tau)+\kappa(\tau-\tau_0)\mathbb{I}^{\tau}_{\tau_0}\}, \\ G_{2,c}(\tau_0,\tau,\kappa) &:= \tau^{-1}(1-\tau)^{-1}\{\tau W_c(1)-W_c(\tau)-\kappa(\tau-\tau_0)\mathbb{I}^{\tau}_{\tau_0}+\kappa\tau(1-\tau_0)\} \end{aligned}$$

(iii)

$$PY \xrightarrow{d} \log\left[\int_{\tau\in\Lambda}\exp\{\tfrac{1}{2}M_c(\tau_0,\tau,\kappa,g'_M,\sigma_\eta/\omega_\varepsilon)\}d\tau\right] =: \mathcal{D}^{PY}_c(\tau_0,\kappa,g'_M,\sigma_\eta/\omega_\varepsilon)$$

where

$$M_c(\tau_0,\tau,\kappa,g'_M,\sigma_\eta/\omega_\varepsilon) := \begin{cases} \mathcal{D}^{W}_{c,-\frac{\sigma_\eta}{\omega_\varepsilon}\mathcal{D}^{\tilde{\rho}_M}_c(\tau_0,\tau,\kappa)}(\tau_0,\tau,\kappa) & \text{if } |\frac{\sigma_\eta}{\omega_\varepsilon}\mathcal{D}^{\tilde{\rho}_M}_c(\tau_0,\tau,\kappa)| > g'_M \\ \mathcal{D}^{W}_{c,0}(\tau_0,\tau,\kappa) & \text{if } |\frac{\sigma_\eta}{\omega_\varepsilon}\mathcal{D}^{\tilde{\rho}_M}_c(\tau_0,\tau,\kappa)| \le g'_M \end{cases},$$

$$\begin{aligned} \mathcal{D}^{W}_{c,c'}(\tau_0,\tau,\kappa) &:= \frac{H_{c,c'}(\tau_0,\tau,\kappa)^2}{R_{c'}(\tau)}, \\ \mathcal{D}^{\tilde{\rho}_M}_c(\tau_0,\tau,\kappa) &:= \mathcal{D}^{\hat{\phi}}_c(\tau_0,\tau,\kappa)+\mathcal{D}^{BC}_c(\tau_0,\tau,\kappa) \end{aligned}$$

with

$$\begin{aligned} H_{c,c'}(\tau_0,\tau,\kappa) &:= c'\textstyle\int_\tau^1(r-\tau)dW_c(r)+c'^2\int_\tau^1(r-\tau)W_c(r)dr+W_c(1)-W_c(\tau)+c'\int_\tau^1 W_c(r)dr \\ &\quad +\kappa\{1-\tau_0-\mathbb{I}^{\tau}_{\tau_0}(\tau-\tau_0)\}+\kappa c'(1-\tau)(1-\tau_0) \\ &\quad +\kappa c'^2[\{1-\tau_0-\mathbb{I}^{\tau}_{\tau_0}(\tau-\tau_0)\}^2\{2+\tau_0-3\tau+4\mathbb{I}^{\tau}_{\tau_0}(\tau-\tau_0)\}]/6 \\ &\quad -(1-\tau)\{a_{c'}-c'^2\tau(1+\tau)/6\}\{c'\textstyle\int_0^1 rdW_c(r)+W_c(1)+c'^2\int_0^1 rW_c(r)dr \\ &\quad +c'\int_0^1 W_c(r)dr+\kappa(1-\tau_0)(a_{c'}-c'^2\tau_0(1+\tau_0)/6)\}/a_{c'}, \\ R_{c'}(\tau) &:= \tau(1-\tau)\left\{1+c'^2\tau(1-\tau)\left(\frac{1}{3}-\frac{c'^2(1+\tau)^2}{36a_{c'}}\right)\right\} \end{aligned}$$

and

$$\mathcal{D}^{\hat{\phi}}_c(\tau_0,\tau,\kappa) := \frac{N_c(1,\tau_0,\tau,\kappa)^2-N_c(0,\tau_0,\tau,\kappa)^2-1}{2\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2dr},$$

$$\mathcal{D}^{BC}_c(\tau_0,\tau,\kappa) := \begin{cases} -\mathcal{D}^{\hat{\phi}}_c(\tau_0,\tau,\kappa) & \text{if } \mathcal{D}^{t_{\hat{\phi}}}_c(\tau_0,\tau,\kappa) > cv_P \\ -\frac{4\mathcal{D}^{\hat{\sigma}_{\hat{\phi}}}_c(\tau_0,\tau,\kappa)}{\left(1+\frac{4-cv_P^2}{cv_P(10+cv_P)}\right)\mathcal{D}^{t_{\hat{\phi}}}_c(\tau_0,\tau,\kappa)+\frac{10(4-cv_P^2)}{cv_P(10+cv_P)}} & \text{if } -10 < \mathcal{D}^{t_{\hat{\phi}}}_c(\tau_0,\tau,\kappa) \le cv_P \\ -\frac{4\mathcal{D}^{\hat{\sigma}_{\hat{\phi}}}_c(\tau_0,\tau,\kappa)}{\mathcal{D}^{t_{\hat{\phi}}}_c(\tau_0,\tau,\kappa)} & \text{if } \mathcal{D}^{t_{\hat{\phi}}}_c(\tau_0,\tau,\kappa) \le -10 \end{cases}$$

with

$$\begin{aligned}\mathcal{D}_c^{\hat{\sigma}_{\hat{\phi}}}(\tau_0,\tau,\kappa) &:= \sqrt{\frac{1}{\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2 dr}},\\ \mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) &:= \frac{N_c(1,\tau_0,\tau,\kappa)^2 - N_c(0,\tau_0,\tau,\kappa)^2 - 1}{2\sqrt{\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2 dr}}\end{aligned}$$

where $a_{c'} := 1 + c' + c'^2/3$ *and* $N_c(r,\tau_0,\tau,\kappa)$ *is as defined in Theorem 3. Here* cv_P *represents a critical value from* $\mathcal{D}_0^{t_{\hat{\phi}}}(\tau_0,\tau_0,\kappa)$.

Remark 5: From (iii), we see that the limiting distribution of PY of the CKP test depends on the ratio $\sigma_\eta/\omega_\varepsilon$. This stems from the dependence of $W_{RQF}(\tau)$ on $\tilde{\rho}_{MS}$ and then of that estimate on $T(\tilde{\rho}_M - 1)$, whose limit is $(\sigma_\eta/\omega_\varepsilon)\mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa)$. Here then, $T(\tilde{\rho}_M - 1)$ is essentially a normalized bias Dickey-Fuller statistic, but one which lacks the necessary variance standardisation that is required to yield a pivotal limit distribution when the ε_t are non-IID.

Remark 6: Also notice that in (iii), $\mathcal{D}_c^{BC}(\tau_0,\tau,\kappa)$, the limit of the step function associated with the bias correction term $C(t_{\hat{\phi}})$, has three regimes rather than four. This arises because the local to unit root asymptotics render the fourth regime in $C(t_{\hat{\phi}})$ asymptotically redundant. Note also that when $\mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) > cv_P$, it follows that $\mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa) = 0$.

5.5 Limits of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$, CKP, $DF_{tb}^{GLS}(\tilde{\tau})$, $A(B)$, U_B^δ

Having established the limit distributions of the constituent components of the test procedures, we are now in a position to determine the asymptotic behaviour of the three original tests $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP under our doubly-local model. The results follow immediately from the previous results and applications of the continuous mapping theorem [CMT].

Theorem 5 *Let the conditions of Theorem 1 hold. Then,*

(i)

$$HHLT_{\bar{\tau}} \xrightarrow{d} \begin{cases} \mathcal{D}_{c,\bar{c}}^{DF_t}(\tau_0,\kappa) & \text{if } \mathcal{D}_c^{\bar{\tau}}(\tau_0,\kappa,g'_W) < \tau_L \\ \mathcal{D}_{c,\bar{c}_{\bar{\tau}}}^{DF_{tb}}(\tau_0,\mathcal{D}_c^{\bar{\tau}}(\tau_0,\kappa,g'_W),\kappa) & \text{if } \mathcal{D}_c^{\bar{\tau}}(\tau_0,\kappa,g'_W) \geq \tau_L \end{cases}$$

where $\bar{c}_{\bar{\tau}}$ *denotes the value of* $\bar{c}_\tau$ *corresponding to* $\mathcal{D}_c^{\bar{\tau}}(\tau_0,\kappa,g'_W)$.

(ii)

$$HHLT_{t_\lambda} \xrightarrow{d} \begin{cases} \mathcal{D}_{c,\bar{c}}^{DF_t}(\tau_0,\kappa) & \text{if } \mathcal{D}_c^{t_\lambda}(\tau_0,\kappa,g'_S,g'_t) < cv_{t_\lambda} \\ \mathcal{D}_{c,\bar{c}_{\tilde{\tau}}}^{DF_{tb}}(\tau_0,\mathcal{D}_c^{\tilde{\tau}}(\tau_0,\kappa),\kappa) & \text{if } \mathcal{D}_c^{t_\lambda}(\tau_0,\kappa,g'_S,g'_t) \geq cv_{t_\lambda} \end{cases}$$

where $\bar{c}_{\tilde{\tau}}$ *denotes the value of* $\bar{c}_\tau$ *corresponding to* $\mathcal{D}_c^{\tilde{\tau}}(\tau_0, \kappa)$.

(iii)

$$CKP \xrightarrow{d} \begin{cases} \mathcal{D}_{c,\bar{c}}^{P_t}(\tau_0, \kappa) & \text{if } \mathcal{D}_c^{PY}(\tau_0, \kappa, g'_M, \sigma_\eta/\omega_\varepsilon) < cv_{PY} \\ \mathcal{D}_{c,\bar{c}_{\check{\tau}}}^{P_{tb}}(\tau_0, \mathcal{D}_{c,\bar{c}_{\check{\tau}}}^{\check{\tau}}(\tau_0, \kappa, \sigma_\varepsilon/\omega_\varepsilon), \kappa) & \text{if } \mathcal{D}_c^{PY}(\tau_0, \kappa, g'_M, \sigma_\eta/\omega_\varepsilon) \geq cv_{PY} \end{cases}$$

where $\bar{c}_{\check{\tau}}$ *denotes the value of* $\bar{c}_\tau$ *corresponding to* $\mathcal{D}_{c,\bar{c}_{\check{\tau}}}^{\check{\tau}}(\tau_0, \kappa, \sigma_\varepsilon/\omega_\varepsilon)$.

Remark 7: It is important to note that the limiting representations given in Theorem 5 do not reduce to those given in HHLT and, in the IID case, CKP, even when $\kappa = 0$; this is because of the changes to $\bar{\lambda}$, t_λ and PY, as outlined above.

The limit distributions of $DF_{tb}^{GLS}(\tilde{\tau})$, $A(t_\lambda)$, $A(PY)$, $U_{t_\lambda}^\delta$ and U_{PY}^δ follow in an entirely straightforward way using the results from the earlier Theorems and applications of the CMT; thus it is not informative to state them explicitly here. The dependence of the asymptotic local power functions of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$, CKP, $DF_{tb}^{GLS}(\tilde{\tau})$, $A(t_\lambda)$, $A(PY)$, $U_{t_\lambda}^\delta$ and U_{PY}^δ on the trend break location and magnitude, τ_0 and κ respectively, under local trend breaks is explored numerically in the next sub-section.

5.6 Asymptotic Size and Local Power

We now consider the asymptotic size and local power of the procedures under the local trend break assumption $\gamma_T = \kappa\omega_\varepsilon T^{-1/2}$, by simulating the limit representations given above. Here, decisions must be made regarding the values of the constants g'_W, g'_S, g'_t and g'_M, which enter the limit distributions of the modified versions of the procedures that we examine. The settings chosen for these constants affect the ability of the limit theory to model the power valley phenomenon observed in finite samples; for example, setting g'_W close to zero would result in $\bar{\tau}$ close to zero, and the asymptotic behaviour of $HHLT_{\bar{\tau}}$ would simply resemble that of DF_t^{GLS}, while at the other extreme, setting g'_W very large would result in the limit of $\bar{\tau}$ being close to the limit of $\tilde{\tau}$, and the asymptotic behaviour of $HHLT_{\bar{\tau}}$ would then replicate that of just $DF_{tb}^{GLS}(\tilde{\tau})$. Similar considerations also apply to the settings for g'_S, g'_t and g'_M. Given that our focus is to establish an asymptotic model of test behaviour that predicts actual finite sample behaviour, we wish to select values for these control constants that result in local asymptotic power functions that closely mimic those seen in finite samples. Perhaps unsurprisingly, we found the most suitable settings for these asymptotic control constants were generally those calibrated according to the implied values used in the implementations of the tests in finite samples. For example, for $HHLT_{\bar{\tau}}$ using $g_W = 1.5$ (as in this paper), equating g'_W to $g_W T^{-1/2}$ yields $g'_W = 0.123$ for $T = 150$ and $g'_W = 0.087$ for $T = 300$. We found that use of the intermediate value $g'_W = 0.1$ delivered an local asymptotic power profile for $HHLT_{\bar{\tau}}$ that closely mirrored observed finite sample behaviour. This same calibration approach was also found to be appropriate for g'_t and g'_M, leading us

to set $g'_t = 7$ and $g'_M = 15$, although for g'_S we found that a larger value, $g'_S = 40000$, resulted in a better predictor of finite sample power behaviour than that indicated by the finite sample-based calibration (which implied values of $g'_S = 18750$ and $g'_S = 30000$ for $T = 150$ and $T = 300$, respectively).

We use the same settings for κ, τ_0 and c as in the finite sample analysis of Sections 3.3 and 4.3, and again report maximum asymptotic sizes of nominal 0.05-level tests across κ (Table 3), and the corresponding asymptotic local power functions for $c = 20$ and $c = 30$ (Figures 8 and 9, respectively). Once again, the powers are size-adjusted, using the scaling constant required to ensure correct size at the point where size is at a maximum (κ^*). The need for size-adjustment arises because the variants of the tests used for the local trend break analysis are not asymptotically correctly sized in the same way as their fixed magnitude trend break counterparts; instead, the asymptotic sizes of the procedures depend on the timing and magnitude of the local break, as the limits in Theorem 5 show. For the trend break pre-tests the critical values cv_{t_λ} and cv_{PY} are those for the the nominal asymptotic 0.05-level. We also set cv_P to its 0.05-level value. Given that the limit distributions of CKP, $A(PY)$ and U^{δ}_{PY} depend on nuisance serial correlation parameters due to their dependence on $\breve{\tau}$ and/or PY, the results we report are those which pertain to the IID case (i.e. for $\sigma_\eta = \sigma_\varepsilon = \omega_\varepsilon$).

As with the finite sample size results, the asymptotic sizes are seen to depend on κ, but with the exception of CKP, we find the maximum asymptotic sizes to be closer to nominal size than in finite samples. More importantly, however, the asymptotic size-adjusted power profiles bear a very close resemblance to their finite sample counterparts in Figures 1-4, especially for $T = 300$, demonstrating the value of our local-to-zero model of the trend magnitude. In particular, and in contrast to asymptotic results obtained under a fixed break assumption (as in HHLT and CKP), the local asymptotic behaviour of $HHLT_{\bar{\tau}}$, $HHLT_{t_\lambda}$ and CKP in Figures 8(a)-(c) and 9(a)-(c) now reproduces the power valley phenomenon observed in finite samples. The potential value of the alternative procedures seen in finite samples is also very clearly mirrored here, with the extent of the drop-off in power for intermediate break magnitudes substantially ameliorated.

6 Conclusions

In this paper we have further investigated the "valleys" observed for trend breaks of relatively small but non-zero magnitude in the finite sample power functions of the recently proposed unit root testing procedures of CKP and HHLT which both employ auxiliary detection devices for whether a trend break is present in the data or not. These valleys, not predicted by the asymptotic analyses of the aforementioned authors who treat the trend break magnitude as fixed, appear when the finite sample power functions of the unit root tests are plotted as functions of the break magnitude.

The contribution of this paper has been two-fold. First, on a practical level we have discussed how the power valley problem can be ameliorated. A valley in the

finite sample power functions can be largely eliminated by implementing a unit root test that always allows for a break in trend, such as the test of PR. However, as well as losing power when no break in trend exists, this approach involves power losses when a break is present due to the necessary use of conservative critical values. We have shown that part of this latter power loss can be recouped by the use of adaptive critical values, whereby the conservative critical value is used if a trend break pre-test fails to reject and otherwise the known break date critical value is used. We have also outlined a union of rejections based approach, whereby we reject the unit root null if either the with-break or the without-break version of the unit root test rejects. This allows us to capture some of the additional power available under the no break case. Overall our results suggest that no one approach is superior in all worlds but that the approach based on the union of rejections principle seems to represent a decent approach, considerably ameliorating the large valleys seen with the tests of CKP and HHLT, yet not losing all the power gains available when no break exists, as happens if a test which always allows for a break is used in isolation. If a greater degree of robustness to the power valley problem is required then the adaptive tests would be recommended.

The second contribution of this paper has been a theoretical one. We have shown that by setting the trend break magnitude to be local to zero (in a Pitman drift sense) the resulting asymptotic local power functions of the unit root tests can in fact closely predict the finite sample power valley phenomenon. Crucially, setting the break magnitude as local to zero reflects in the asymptotic analysis the genuine uncertainty that will exist in finite samples as to whether a trend break exists or not, which is not the case when the break magnitude is taken to be fixed (where the trend break detection devices can distinguish perfectly between break and no-break environments).

We conclude with a suggestion for further research. This paper has focused on testing for a unit root in the presence of a single possible break in trend. It would be interesting to extend the work in this paper to the case where we allow for the possibility of multiple trend breaks. It seems likely that the power valley phenomenon we have discussed in this paper could only be expected to worsen in the multiple breaks case, although in principle the new tests proposed in this paper, namely $A(t_\lambda)$, $A(PY)$, $U^{\delta}_{t_\lambda}$ and U^{δ}_{PY}, could also be extended to the context of multiple possible trend breaks.

Acknowledgements

We thank Peter Robinson, an Associate Editor and three anonymous referees for helpful comments on earlier versions of this paper. We also thank Josep Lluís Carrion-i-Silvestre and Tomoyoshi Yabu for providing GAUSS code for their procedures. Financial support provided by the Economic and Social Research Council of the United Kingdom under research grant RES-000-22-3882 is gratefully acknowledged by the authors.

Appendix

The results stated in Theorems 1 and 2 are derived pointwise in $\tau \in \Lambda$, the generic break fraction argument. The subsequent results in Theorem 3 onwards often require a continuous mapping applied to a sequence of statistics taken across τ. In each of these cases the stated limiting representations follow from the fixed τ representation, using the arguments proved in Zivot and Andrews (1992). It is understood that we appeal to those arguments on each occasion a function is taken across a sequence of statistics indexed by the argument τ.

In what follows, we can set $\mu = \beta = 0$ in (1) without loss of generality. Moreover, the dependence of certain quantities on parameters such as $\bar{c}$, τ etc. is suppressed when not essential to the developments.

Proof of Theorem 1

(i) First consider $\tilde{\mu}$ and $\tilde{\beta}$

$$\begin{bmatrix} \tilde{\mu} \\ \tilde{\beta} \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} \\ g_{12} & g_{22} \end{bmatrix}^{-1} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

where

$$\begin{aligned} g_{11} &:= 1 + (1-\bar{\rho})^2(T-1), \;\; g_{12} := 1 + (1-\bar{\rho})\textstyle\sum_{t=2}^{T}\{t - \bar{\rho}(t-1)\} \\ g_{22} &:= 1 + \textstyle\sum_{t=2}^{T}\{t - \bar{\rho}(t-1)\}^2, \;\; h_1 := y_1 + (1-\bar{\rho})\textstyle\sum_{t=2}^{T}(y_t - \bar{\rho}y_{t-1}) \\ h_2 &:= y_1 + \textstyle\sum_{t=2}^{T}(y_t - \bar{\rho}y_{t-1})\{t - \bar{\rho}(t-1)\}. \end{aligned}$$

The limits involved in the 2×2 matrix are standard: $g_{11} \to 1$, $g_{12} \to 1 + \bar{c} + \bar{c}^2/2$ and $T^{-1}g_{22} \to 1 + \bar{c} + \bar{c}^2/3 = a_{\bar{c}}$. Those for the 2×1 vector are as follows

$$h_1 = y_1 + \bar{c}T^{-1}(y_T - y_1) + \bar{c}^2 T^{-2}\textstyle\sum_{t=2}^{T} y_{t-1} = u_1 + o_p(1).$$

$$\begin{aligned} T^{-1/2}h_2 &= \bar{c}T^{-3/2}\textstyle\sum_{t=2}^{T} t\Delta u_t + T^{-1/2}u_T + \bar{c}^2 T^{-5/2}\sum_{t=2}^{T} t u_{t-1} + \bar{c}T^{-3/2}\sum_{t=2}^{T} u_{t-1} \\ &\quad + \bar{c}\kappa T^{-2}\textstyle\sum_{t=2}^{T} tDU_t(\tau_0) + \kappa T^{-1}(T - \lfloor \tau_0 T \rfloor) + \bar{c}^2\kappa T^{-3}\sum_{t=2}^{T} tDT_{t-1}(\tau_0) \\ &\quad + \bar{c}\kappa T^{-2}\textstyle\sum_{t=2}^{T} DT_{t-1}(\tau_0) + o_p(1) \\ &\xrightarrow{d} \bar{c}\omega_\varepsilon\{W_c(1) - \textstyle\int_0^1 W_c(s)ds\} + \omega_\varepsilon W_c(1) + \bar{c}^2\omega_\varepsilon\textstyle\int_0^1 sW_c(s)ds + \bar{c}\omega_\varepsilon\int_0^1 W_c(s)ds \\ &\quad + \bar{c}\omega_\varepsilon\kappa(1-\tau_0^2)/2 + \omega_\varepsilon\kappa(1-\tau_0) + \bar{c}^2\omega_\varepsilon\kappa\{(1-\tau_0^3)/3 - \tau_0(1-\tau_0^2)/2\} \\ &\quad + \bar{c}\omega_\varepsilon\kappa(1-\tau_0)^2/2 \\ &= (1+\bar{c})\omega_\varepsilon W_c(1) + \bar{c}^2\textstyle\int_0^1 sW_c(s)dr + \omega_\varepsilon\kappa\,(1-\tau_0)\,\{a_{\bar{c}} - \bar{c}^2\tau_0(1+\tau_0)/6\} \\ &= \omega_\varepsilon\{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\} \end{aligned}$$

where, for a generic argument τ, $DU_t(\tau) := 1(t > \lfloor \tau T \rfloor)$, and where $b_{c,\bar{c}}$ and $\kappa f_{c,\bar{c}}(\tau_0)$ are implicitly defined. Consequently,

$$\begin{bmatrix} \tilde{\mu} \\ T^{1/2}\tilde{\beta} \end{bmatrix} = \begin{bmatrix} g_{11} & T^{-1/2}g_{12} \\ T^{-1/2}g_{12} & T^{-1}g_{22} \end{bmatrix}^{-1} \begin{bmatrix} h_1 \\ T^{-1/2}h_2 \end{bmatrix} \xrightarrow{d} \begin{bmatrix} u_1 \\ \omega_\varepsilon\{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\}/a_{\bar{c}} \end{bmatrix}.$$

The limit of $T^{-1/2}\tilde{u}_{\lfloor rT \rfloor}$ can now be obtained

$$\begin{aligned} T^{-1/2}\tilde{u}_{\lfloor rT \rfloor} &= T^{-1/2}y_{\lfloor rT \rfloor} - T^{-1/2}\tilde{\mu} - T^{-1/2}\tilde{\beta}\lfloor rT \rfloor \\ &= T^{-1/2}u_{\lfloor rT \rfloor} + \omega_\varepsilon \kappa (r-\tau_0)\mathbb{I}^r_{\tau_0} - T^{1/2}\tilde{\beta}r + o_p(1) \\ &\xrightarrow{d} \omega_\varepsilon W_c(r) + \omega_\varepsilon \kappa (r-\tau_0)\mathbb{I}^r_{\tau_0} - \omega_\varepsilon \{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\} r/a_{\bar{c}} \\ &= \omega_\varepsilon K_{c,\bar{c}}(r, \tau_0, \kappa). \end{aligned}$$

Next, and in order to simplify the presentation of the proofs, we will assume that $\{\varepsilon_t\}$ follows a stationary AR(p) process, i.e.

$$\delta(L)\varepsilon_t = \eta_t, \qquad \delta(L) = 1 - \delta_1 L - \delta_2 L^2 - ... - \delta_p L^p. \tag{A.1}$$

Setting p appropriately in the regression (5), and defining $v_t := \Delta u_t - \sum_{j=1}^p \delta_j \varepsilon_{t-j}$, we can write

$$\begin{bmatrix} T\hat{\phi} \\ T^{1/2}(\hat{\delta}_1 - \delta_1) \\ \vdots \\ T^{1/2}(\hat{\delta}_p - \delta_p) \end{bmatrix} = \begin{bmatrix} T^{-2}\sum \tilde{u}^2_{t-1} & T^{-3/2}\sum \tilde{u}_{t-1}\Delta\tilde{u}_{t-1} & \dots & T^{-3/2}\sum \tilde{u}_{t-1}\Delta\tilde{u}_{t-p} \\ T^{-3/2}\sum \tilde{u}_{t-1}\Delta\tilde{u}_{t-1} & T^{-1}\sum \Delta\tilde{u}^2_{t-1} & \dots & T^{-1}\sum \Delta\tilde{u}_{t-1}\Delta\tilde{u}_{t-p} \\ \vdots & \vdots & \dots & \vdots \\ T^{-3/2}\sum \tilde{u}_{t-1}\Delta\tilde{u}_{t-p} & T^{-1}\sum \Delta\tilde{u}_{t-1}\Delta\tilde{u}_{t-p} & \dots & T^{-1}\sum \Delta\tilde{u}^2_{t-p} \end{bmatrix}^{-1}$$
$$\left(\begin{bmatrix} T^{-1}\sum \tilde{u}_{t-1}v_t \\ T^{-1/2}\sum \Delta\tilde{u}_{t-1}v_t \\ \vdots \\ T^{-1/2}\sum \Delta\tilde{u}_{t-p}v_t \end{bmatrix} - \begin{bmatrix} T^{-1}\sum \tilde{u}_{t-1}(\Delta u_t - \Delta\tilde{u}_t) \\ T^{-1/2}\sum \Delta\tilde{u}_{t-1}(\Delta u_t - \Delta\tilde{u}_t) \\ \vdots \\ T^{-1/2}\sum \Delta\tilde{u}_{t-p}(\Delta u_t - \Delta\tilde{u}_t) \end{bmatrix} \right.$$
$$\left. + \begin{bmatrix} \sum_{j=1}^p \delta_j T^{-1}\sum \tilde{u}_{t-1}(\varepsilon_{t-j} - \Delta\tilde{u}_{t-j}) \\ \sum_{j=1}^p \delta_j T^{-1/2}\sum \Delta\tilde{u}_{t-1}(\varepsilon_{t-j} - \Delta\tilde{u}_{t-j}) \\ \vdots \\ \sum_{j=1}^p \delta_j T^{-1/2}\sum \Delta\tilde{u}_{t-p}(\varepsilon_{t-j} - \Delta\tilde{u}_{t-j}) \end{bmatrix} \right)$$

where the $\sum$ summation denotes $\sum_{t=p+2}^T$. We find, on simplifying the $K_{c,\bar{c}}(r, \tau_0, \kappa)$ notation to K, and $\{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\}/a_{\bar{c}}$ to b,

$$\begin{bmatrix} T\hat{\phi} \\ T^{1/2}(\hat{\delta}_1 - \delta_1) \\ \vdots \\ T^{1/2}(\hat{\delta}_p - \delta_p) \end{bmatrix} \xrightarrow{d} \omega_\varepsilon^{-2} \begin{bmatrix} \int_0^1 K^2 dr & 0 & \dots & 0 \\ 0 & E(\varepsilon_t^2) & \dots & E(\varepsilon_t \varepsilon_{t-p+1}) \\ \vdots & \vdots & \dots & \vdots \\ 0 & E(\varepsilon_t \varepsilon_{t-p+1}) & \dots & E(\varepsilon_t^2) \end{bmatrix}^{-1}$$
$$\left(\begin{bmatrix} -\omega_\varepsilon^2 c \int_0^1 K W_c(r) dr + \omega_\varepsilon \sigma_\eta \int_0^1 K dW(r) \\ O_p(1) \\ \vdots \\ O_p(1) \end{bmatrix} - \begin{bmatrix} \omega_\varepsilon^2 (b \int_0^1 K dr - \kappa \int_{\tau_0}^1 K dr) \\ O_p(1) \\ \vdots \\ O_p(1) \end{bmatrix} \right.$$
$$\left. + \begin{bmatrix} \omega_\varepsilon^2 \{c \int_0^1 K W_c(r) dr + b \int_0^1 K dr - \kappa \int_{\tau_0}^1 K dr\} \sum_{j=1}^p \delta_j \\ O_p(1) \\ \vdots \\ O_p(1) \end{bmatrix} \right)$$

since

$$\begin{aligned} T^{-1} \textstyle\sum_{t=j+2}^T \Delta\tilde{u}_t \Delta\tilde{u}_{t-j} &= T^{-1} \textstyle\sum_{t=j+2}^T (\Delta u_t + \omega_\varepsilon \kappa T^{-1/2} \mathbb{I}_{\tau_0}^t - \tilde{\beta})(\Delta u_{t-j} + \omega_\varepsilon \kappa T^{-1/2} \mathbb{I}_{\tau_0}^{t-j} - \tilde{\beta}) \\ &= T^{-1} \textstyle\sum_{t=j+2}^T \Delta u_t \Delta u_{t-j} + o_p(1) \\ &= T^{-1} \textstyle\sum_{t=j+2}^T \varepsilon_t \varepsilon_{t-j} + o_p(1) \xrightarrow{p} E(\varepsilon_t \varepsilon_{t-j}) \end{aligned}$$

$$\begin{aligned} T^{-1} \textstyle\sum \tilde{u}_{t-1} v_t &= T\phi_T T^{-2} \textstyle\sum \tilde{u}_{t-1} u_{t-1} + T^{-1} \textstyle\sum \tilde{u}_{t-1} \eta_t \\ &\xrightarrow{d} -\omega_\varepsilon^2 c \int_0^1 K W_c(r) dr + \omega_\varepsilon \sigma_\eta \int_0^1 K dW(r) \end{aligned}$$

and, using $\Delta u_t - \Delta\tilde{u}_t = \tilde{\beta} - \omega_\varepsilon \kappa T^{-1/2} DU_t(\tau_0)$,

$$\begin{aligned} T^{-1} \textstyle\sum \tilde{u}_{t-1}(\Delta u_t - \Delta\tilde{u}_t) &= T^{1/2} \tilde{\beta} T^{-3/2} \textstyle\sum \tilde{u}_{t-1} - \omega_\varepsilon \kappa T^{-3/2} \textstyle\sum DU_t(\tau_0) \tilde{u}_{t-1} \\ &\xrightarrow{d} \omega_\varepsilon^2 [\{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\}/a_{\bar{c}}] \int_0^1 K dr - \omega_\varepsilon^2 \kappa \int_{\tau_0}^1 K dr \end{aligned}$$

and also, using $\varepsilon_t - \Delta\tilde{u}_t = cT^{-1} u_{t-1} + \Delta u_t - \Delta\tilde{u}_t$, for all j,

$$\begin{aligned} T^{-1} \textstyle\sum \tilde{u}_{t-1}(\varepsilon_{t-j} - \Delta\tilde{u}_{t-j}) &= cT^{-2} \textstyle\sum \tilde{u}_{t-1} u_{t-j-1} + T^{-1} \textstyle\sum \tilde{u}_{t-1}(\Delta u_{t-j} - \Delta\tilde{u}_{t-j}) \\ &\xrightarrow{d} \omega_\varepsilon^2 c \int_0^1 K W_c(r) dr + \omega_\varepsilon^2 b \int_0^1 K dr - \omega_\varepsilon^2 \kappa \int_{\tau_0}^1 K dr. \end{aligned}$$

It then follows that, using $\omega_\varepsilon = \sigma_\eta (1 - \sum_{j=1}^p \delta_j)^{-1}$,

$$\begin{aligned} T\hat{\phi} &\xrightarrow{d} \frac{-\omega_\varepsilon c \sigma_\eta \int_0^1 K W_c(r) dr + \omega_\varepsilon \sigma_\eta \int_0^1 K dW(r) - \omega_\varepsilon \sigma_\eta (b \int_0^1 K dr - \kappa \int_{\tau_0}^1 K dr)}{\omega_\varepsilon^2 \int_0^1 K^2 dr} \\ &= \frac{\sigma_\eta}{\omega_\varepsilon} \frac{\left[\int_0^1 K dW_c(r) - b \int_0^1 K dr + \kappa \int_{\tau_0}^1 K dr \right]}{\int_0^1 K^2 dr}. \end{aligned}$$

Next, observe that this limit can alternatively be expressed as

$$\begin{aligned} T\hat{\phi} &\xrightarrow{d} \frac{\sigma_\eta}{\omega_\varepsilon}\frac{\int_0^1 K_{c,\bar{c}}(r,\tau_0,\kappa)dK_{c,\bar{c}}(r,\tau_0,\kappa)}{\int_0^1 K_{c,\bar{c}}(r,\tau_0,\kappa)^2dr} \\ &= \frac{\sigma_\eta}{\omega_\varepsilon}\frac{K_{c,\bar{c}}(1,\tau_0,\kappa)^2-1}{2\int_0^1 K_{c,\bar{c}}(r,\tau_0,\kappa)^2dr} \end{aligned}$$

since

$$dK_{c,\bar{c}}(r,\tau_0,\kappa) = dW_c(r) - \{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\}/a_{\bar{c}} + \kappa \mathbb{I}^r_{\tau_0}.$$

Then, writing the standard error associated with $\hat{\phi}$ as $s.e.(\hat{\phi}) := \sqrt{\hat{\sigma}_\eta^2 V_{11}}$, where $\hat{\sigma}_\eta^2 := (T-p-1)^{-1}\sum_{t=p+2}^T \hat{\eta}_t^2$ and V_{11} denotes the (1,1) element of $(X'X)^{-1}$, with X the regressor matrix associated with (5), we have

$$\hat{\sigma}_\eta^2 \xrightarrow{p} \sigma_\eta^2, \qquad T^2 V_{11} \xrightarrow{d} \{\omega_\varepsilon^2 \int_0^1 K_{c,\bar{c}}(r,\tau_0,\kappa)^2 dr\}^{-1}$$

and

$$DF_t^{GLS} = \frac{T\hat{\phi}}{\sqrt{\hat{\sigma}_\eta^2 T^2 V_{11}}} \xrightarrow{d} \frac{K_{c,\bar{c}}(1,\tau_0,\kappa)^2-1}{2\sqrt{\int_0^1 K_{c,\bar{c}}(r,\tau_0,\kappa)^2 dr}}$$

which is the limit given in (i). It can be shown that this same limit for DF_t^{GLS} is also obtained under the more general conditions for $\{\varepsilon_t\}$ given in Assumption 1, provided the number of lagged differences p satisfies the condition that $1/p + p^3/T \to 0$ as $T \to \infty$; cf. Chang and Park (2002).

(ii) A little manipulation shows that we can write

$$\begin{aligned} S(\bar{\rho}) - y_1^2 - \sum_{t=2}^T (y_t - \bar{\rho}y_{t-1})^2 &= -\begin{bmatrix} h_1 & T^{-1/2}h_2 \end{bmatrix} \begin{bmatrix} g_{11} & T^{-1/2}g_{12} \\ T^{-1/2}g_{12} & T^{-1}g_{22} \end{bmatrix}^{-1} \begin{bmatrix} h_1 \\ T^{-1/2}h_2 \end{bmatrix} \\ &\xrightarrow{d} -\begin{bmatrix} u_1 & \omega_\varepsilon\{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & a_{\bar{c}} \end{bmatrix}^{-1} \begin{bmatrix} u_1 \\ \omega_\varepsilon\{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\} \end{bmatrix} \\ &= -u_1^2 - \omega_\varepsilon^2\{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\}^2/a_{\bar{c}} \end{aligned}$$

from which it follows that

$$S(1) - y_1^2 - \textstyle\sum_{t=2}^T (y_t - y_{t-1})^2 \xrightarrow{d} -u_1^2 - \omega_\varepsilon^2\{b_{c,0} + \kappa f_{c,0}(\tau_0)\}^2$$

and since

$$\begin{aligned} y_1^2 + \textstyle\sum_{t=2}^T (y_t &- \bar{\rho}y_{t-1})^2 - \bar{\rho}y_1^2 - \bar{\rho}\sum_{t=2}^T (y_t - y_{t-1})^2 \\ &\xrightarrow{d} \omega_\varepsilon^2 \bar{c}^2 \textstyle\int_0^1 W_c(r)^2 dr + \omega_\varepsilon^2 \bar{c} W_c(1)^2 \\ &\quad + \omega_\varepsilon^2 \kappa\{2\bar{c}^2 \textstyle\int_{\tau_0}^1 (r-\tau_0) W_c(r) dr + 2\bar{c}\int_{\tau_0}^1 (r-\tau_0) dW_c(r) + 2\bar{c}\int_{\tau_0}^1 W_c(r) dr\} \\ &\quad + \omega_\varepsilon^2 \kappa^2\{\bar{c}^2(1 - 3\tau_0 + 3\tau_0^2 - \tau_0^3)/3 + \bar{c}(1-\tau_0)^2\} \\ &= \omega_\varepsilon^2[\bar{c}^2 \textstyle\int_0^1 W_c(r)^2 dr + \bar{c} W_c(1)^2 + \kappa j_{c,\bar{c}}(\tau_0) + \kappa^2 k_{\bar{c}}(\tau_0)] \end{aligned} \tag{A.2}$$

we obtain, using an application of the CMT, that

$$S(\bar{\rho}) - \bar{\rho}S(1) \xrightarrow{d} \omega_\varepsilon^2[\{b_{c,0} + \kappa f_{c,0}(\tau_0)\}^2 - \{b_{c,\bar{c}} + \kappa f_{c,\bar{c}}(\tau_0)\}^2/a_{\bar{c}} + \bar{c}^2\textstyle\int_0^1 W_c(r)^2 dr + \bar{c}W_c(1)^2 + \kappa j_{c,\bar{c}}(\tau_0) + \kappa^2 k_{\bar{c}}(\tau_0)].$$

Also, assuming $\{\varepsilon_t\}$ is generated as in (A.1), from (i) above, we find

$$\tilde{\omega}^2 = \frac{(T-p-1)^{-1}\sum_{t=p+2}^T \hat{\eta}_t^2}{(1-\sum_{j=1}^p \hat{\delta}_j)^2} \xrightarrow{p} \frac{\sigma_\eta^2}{(1-\sum_{j=1}^p \delta_j)^2} = \omega_\varepsilon^2$$

and the result of (ii) follows, which again holds under the more general conditions for ε_t of Assumption 1, again provided $1/p + p^3/T \to 0$ as $T \to \infty$.

Proof of Theorem 2

(i) First consider $\tilde{\mu}_\tau$, $\tilde{\beta}_\tau$ and $\tilde{\gamma}_\tau$:

$$\begin{bmatrix} \tilde{\mu}_\tau \\ \tilde{\beta}_\tau \\ \tilde{\gamma}_\tau \end{bmatrix} = \begin{bmatrix} g_{11} & g_{12} & g_{13} \\ g_{12} & g_{22} & g_{23} \\ g_{13} & g_{23} & g_{33} \end{bmatrix}^{-1} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$$

where

$$\begin{aligned} g_{13} &:= (1-\bar{\rho})\textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T \{t - \lfloor \tau T\rfloor - \bar{\rho}(t - \lfloor \tau T\rfloor - 1)\} \\ g_{23} &:= \textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T \{t - \bar{\rho}(t-1)\}\{t - \lfloor \tau T\rfloor - \bar{\rho}(t - \lfloor \tau T\rfloor - 1)\} \\ g_{33} &:= \textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T \{t - \lfloor \tau T\rfloor - \bar{\rho}(t - \lfloor \tau T\rfloor - 1)\}^2 \\ \\ h_3 &:= \textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T (y_t - \bar{\rho}y_{t-1})\{t - \lfloor \tau T\rfloor - \bar{\rho}(t - \lfloor \tau T\rfloor - 1)\}. \end{aligned}$$

For these new terms

$$\begin{aligned} T^{-1/2}h_3 &= T^{-1/2}y_T - T^{-1/2}y_{\lfloor \tau T\rfloor} + \bar{c}T^{-3/2}\textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T t\Delta y_t - \bar{c}\tau T^{-1/2}\sum_{t=\lfloor \tau T\rfloor+1}^T \Delta y_t \\ &\quad + \bar{c}T^{-3/2}\textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T y_{t-1} + \bar{c}^2T^{-5/2}\sum_{t=\lfloor \tau T\rfloor+1}^T t y_{t-1} - \bar{c}^2\tau T^{-3/2}\sum_{t=\lfloor \tau T\rfloor+1}^T y_{t-1} \\ &\quad + o_p(1) \\ &= T^{-1/2}u_T - T^{-1/2}u_{\lfloor \tau T\rfloor} + \bar{c}T^{-3/2}\textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T t\Delta u_t - \bar{c}\tau T^{-1/2}\sum_{t=\lfloor \tau T\rfloor+1}^T \Delta u_t \\ &\quad + \bar{c}T^{-3/2}\textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T u_{t-1} + \bar{c}^2T^{-5/2}\sum_{t=\lfloor \tau T\rfloor+1}^T t u_{t-1} - \bar{c}^2\tau T^{-3/2}\sum_{t=\lfloor \tau T\rfloor+1}^T u_{t-1} \\ &\quad + \kappa\omega_\varepsilon(1-\tau_0) - \kappa(\tau-\tau_0)\mathbb{I}_{\tau_0}^\tau + \bar{c}\kappa\omega_\varepsilon T^{-2}\textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T tDU_t(\tau_0) \\ &\quad - \bar{c}\kappa\omega_\varepsilon\tau T^{-1}\textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T DU_{t-1}(\tau_0) + \bar{c}\kappa\omega_\varepsilon T^{-2}\sum_{t=\lfloor \tau T\rfloor+1}^T DT_{t-1}(\tau_0) \\ &\quad + \bar{c}^2\kappa\omega_\varepsilon T^{-3}\textstyle\sum_{t=\lfloor \tau T\rfloor+1}^T tDT_{t-1}(\tau_0) - \bar{c}^2\kappa\omega_\varepsilon\tau T^{-2}\sum_{t=\lfloor \tau T\rfloor+1}^T DT_{t-1}(\tau_0) + o_p(1) \end{aligned}$$

$$
\begin{aligned}
&\xrightarrow{d} \omega_\varepsilon\{W_c(1)-W_c(\tau)\}+\omega_\varepsilon\bar{c}\{W_c(1)-\tau W_c(\tau)-\textstyle\int_\tau^1 W_c(s)ds\}-\omega_\varepsilon\bar{c}\tau\{W_c(1)-W_c(\tau)\}\\
&\quad+\omega_\varepsilon\{\bar{c}\textstyle\int_\tau^1 W_c(s)ds+\bar{c}^2\int_\tau^1 sW_c(s)ds-\bar{c}^2\tau\int_\tau^1 W_c(s)ds\}+\omega_\varepsilon\{\kappa(1-\tau_0)-\kappa(\tau-\tau_0)\mathbb{I}_{\tau_0}^\tau\}\\
&\quad+\omega_\varepsilon\bar{c}\kappa\{(1-\tau_0^2)/2-(\tau^2-\tau_0^2)\mathbb{I}_{\tau_0}^\tau/2\}-\omega_\varepsilon\bar{c}\kappa\tau\{1-\tau_0-(\tau-\tau_0)\mathbb{I}_{\tau_0}^\tau\}\\
&\quad+\omega_\varepsilon\bar{c}\kappa(1-\bar{c}\tau)\{(1-\tau_0)^2/2-(\tau-\tau_0)^2\mathbb{I}_{\tau_0}^\tau/2\}+\omega_\varepsilon\bar{c}^2\kappa[(1-\tau_0^3)/3-\tau_0(1-\tau_0^2)/2\\
&\quad-\{(\tau^3-\tau_0^3)/3-\tau_0(\tau^2-\tau_0^2)/2\}\mathbb{I}_{\tau_0}^\tau]\\
=\;&\omega_\varepsilon(1+\bar{c}-\bar{c}\tau)W_c(1)-\omega_\varepsilon W_c(\tau)+\omega_\varepsilon\bar{c}^2\textstyle\int_\tau^1(s-\tau)W_c(s)ds+\omega_\varepsilon\kappa[(1-\tau_0)\{a_{\bar{c}}-\bar{c}\tau\\
&\quad-\bar{c}^2\tau(1-\tau_0)/2-\bar{c}^2\tau_0(1+\tau_0)/6\}-(\tau-\tau_0)\{1-\bar{c}^2(\tau-\tau_0)^2/6\}\mathbb{I}_{\tau_0}^\tau]\\
=\;&\omega_\varepsilon\{b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau)\}.
\end{aligned}
$$

Collecting results, we therefore have that

$$
\begin{aligned}
\begin{bmatrix}\tilde{\mu}_\tau\\ T^{1/2}\tilde{\beta}_\tau\\ T^{1/2}\tilde{\gamma}_\tau\end{bmatrix} &= \begin{bmatrix} g_{11} & T^{-1/2}g_{12} & T^{-1/2}g_{13}\\ T^{-1/2}g_{12} & T^{-1}g_{22} & T^{-1}g_{23}\\ T^{-1/2}g_{13} & T^{-1}g_{23} & T^{-1}g_{33}\end{bmatrix}^{-1}\begin{bmatrix} h_1\\ T^{-1/2}h_2\\ T^{-1/2}h_3\end{bmatrix}\\
&\xrightarrow{d} \begin{bmatrix}1 & 0 & 0\\ 0 & a_{\bar{c}} & m_{\bar{c}}(\tau)\\ 0 & m_{\bar{c}}(\tau) & d_{\bar{c}}(\tau)\end{bmatrix}^{-1}\begin{bmatrix} u_1\\ \omega_\varepsilon\{b_{c,\bar{c}}+\kappa f_{c,\bar{c}}(\tau_0)\}\\ \omega_\varepsilon\{b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau)\}\end{bmatrix}
\end{aligned}
$$

giving of the limit of $T^{-1/2}\tilde{u}_{\tau,\lfloor rT\rfloor}$ as

$$
\begin{aligned}
T^{-1/2}\tilde{u}_{\tau,\lfloor rT\rfloor} &= T^{-1/2}y_{\lfloor rT\rfloor}-T^{-1/2}\tilde{\mu}_\tau-T^{-1/2}\tilde{\beta}_\tau\lfloor rT\rfloor-T^{-1/2}\tilde{\gamma}_\tau\{\lfloor rT\rfloor-\lfloor\tau T\rfloor\}\mathbb{I}_\tau^r\\
&= T^{-1/2}u_{\lfloor rT\rfloor}+\omega_\varepsilon\kappa(r-\tau_0)\mathbb{I}_{\tau_0}^r-\begin{bmatrix} r & (r-\tau)\mathbb{I}_\tau^r\end{bmatrix}\begin{bmatrix} T^{1/2}\tilde{\beta}_\tau\\ T^{1/2}\tilde{\gamma}_\tau\end{bmatrix}+o_p(1)\\
&\xrightarrow{d} \omega_\varepsilon W_c(r)+\omega_\varepsilon\kappa(r-\tau_0)\mathbb{I}_{\tau_0}^r\\
&\quad-\begin{bmatrix} r & (r-\tau)\mathbb{I}_\tau^r\end{bmatrix}\begin{bmatrix} a_{\bar{c}} & m_{\bar{c}}(\tau)\\ m_{\bar{c}}(\tau) & d_{\bar{c}}(\tau)\end{bmatrix}^{-1}\begin{bmatrix}\omega_\varepsilon\{b_{c,\bar{c}}+\kappa f_{c,\bar{c}}(\tau_0)\}\\ \omega_\varepsilon\{b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau)\}\end{bmatrix}\\
&= \omega_\varepsilon L_{c,\bar{c}}(r,\tau_0,\tau,\kappa).
\end{aligned}
$$

The remainder of the proof to establish the limit behaviour of $DF_{tb}^{GLS}(\tau)$ follows a straightforward parallel of the approach applied for DF_t^{GLS} in Theorem 1 (i).

(ii) Write

$$
\begin{aligned}
&S(\bar{\rho},\tau)-y_1^2-\textstyle\sum_{t=2}^T(y_t-\bar{\rho}y_{t-1})^2\\
&=-\begin{bmatrix} h_1 & T^{-1/2}h_2 & T^{-1/2}h_3\end{bmatrix}\begin{bmatrix} g_{11} & T^{-1/2}g_{12} & T^{-1/2}g_{13}\\ T^{-1/2}g_{12} & T^{-1}g_{22} & T^{-1}g_{23}\\ T^{-1/2}g_{13} & T^{-1}g_{23} & T^{-1}g_{33}\end{bmatrix}^{-1}\begin{bmatrix} h_1\\ T^{-1/2}h_2\\ T^{-1/2}h_3\end{bmatrix}
\end{aligned}
$$

$$\begin{aligned}
&\xrightarrow{d} \; -\left[\begin{array}{ccc} u_1 & \omega_\varepsilon\{b_{c,\bar{c}}+\kappa f_{c,\bar{c}}(\tau_0)\} & \omega_\varepsilon\{b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau)\} \end{array}\right]
\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & a_{\bar{c}} & m_{\bar{c}}(\tau) \\ 0 & m_{\bar{c}}(\tau) & d_{\bar{c}}(\tau) \end{array}\right]^{-1} \\
&\quad \times \left[\begin{array}{c} u_1 \\ \omega_\varepsilon\{b_{c,\bar{c}}+\kappa f_{c,\bar{c}}(\tau_0)\} \\ \omega_\varepsilon\{b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau)\} \end{array}\right] \\
&= \; -u_1^2 - \omega_\varepsilon^2 \left[\begin{array}{cc} b_{c,\bar{c}}+\kappa f_{c,\bar{c}}(\tau_0) & b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau) \end{array}\right]
\left[\begin{array}{cc} a_{\bar{c}} & m_{\bar{c}}(\tau) \\ m_{\bar{c}}(\tau) & d_{\bar{c}}(\tau) \end{array}\right]^{-1} \\
&\quad \times \left[\begin{array}{c} b_{c,\bar{c}}+\kappa f_{c,\bar{c}}(\tau_0) \\ b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau) \end{array}\right] \qquad \text{(A.3)}
\end{aligned}$$

from which it follows that

$$\begin{aligned}
S(1,\tau) - y_1^2 - \textstyle\sum_{t=2}^T (y_t - y_{t-1})^2 \; \xrightarrow{d} \; & -u_1^2 - \omega_\varepsilon^2 \left[\begin{array}{cc} b_{c,0}+\kappa f_{c,0}(\tau_0) & b_{c,0}(\tau)+\kappa f_{c,0}(\tau_0,\tau) \end{array}\right] \\
& \times \left[\begin{array}{cc} 1 & m_0(\tau) \\ m_0(\tau) & d_0(\tau) \end{array}\right]^{-1} \left[\begin{array}{c} b_{c,0}+\kappa f_{c,0}(\tau_0) \\ b_{c,0}(\tau)+\kappa f_{c,0}(\tau_0,\tau) \end{array}\right]
\end{aligned}$$

and since $\{y_1^2 + \sum_{t=2}^T (y_t - \bar{\rho} y_{t-1})^2 - \bar{\rho} y_1^2 - \bar{\rho} \sum_{t=2}^T (y_t - y_{t-1})^2\}$ has the same limit as in (A.2), we find using the CMT that

$$\begin{aligned}
& S(\bar{\rho},\tau) - \bar{\rho} S(1,\tau) \xrightarrow{d} \\
& \omega_\varepsilon^2 \left[\begin{array}{cc} b_{c,0}+\kappa f_{c,0}(\tau_0) & b_{c,0}(\tau)+\kappa f_{c,0}(\tau_0,\tau) \end{array}\right]
\left[\begin{array}{cc} 1 & m_0(\tau) \\ m_0(\tau) & d_0(\tau) \end{array}\right]^{-1}
\left[\begin{array}{c} b_{c,0}+\kappa f_{c,0}(\tau_0) \\ b_{c,0}(\tau)+\kappa f_{c,0}(\tau_0,\tau) \end{array}\right] \\
& -\omega_\varepsilon^2 \left[\begin{array}{cc} b_{c,\bar{c}}+\kappa f_{c,\bar{c}}(\tau_0) & b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau) \end{array}\right]
\left[\begin{array}{cc} a_{\bar{c}} & m_{\bar{c}}(\tau) \\ m_{\bar{c}}(\tau) & d_{\bar{c}}(\tau) \end{array}\right]^{-1}
\left[\begin{array}{c} b_{c,\bar{c}}+\kappa f_{c,\bar{c}}(\tau_0) \\ b_{c,\bar{c}}(\tau)+\kappa f_{c,\bar{c}}(\tau_0,\tau) \end{array}\right] \\
& +\omega_\varepsilon^2 [\bar{c}^2 \textstyle\int_0^1 W_c(r)^2 dr + \bar{c} W_c(1)^2 + \kappa j_{c,\bar{c}}(\tau_0) + \kappa^2 k_{\bar{c}}(\tau_0)].
\end{aligned}$$

The result in (ii) then follows since $\tilde{\omega}^2(\tau) \xrightarrow{p} \omega_\varepsilon^2$.

Proof of Theorem 3

(i) It is straightforward to show that

$$\begin{aligned}
S(1,\tau) \; = \; & \textstyle\sum_{t=2}^T \Delta y_t^2 - \left\{ \left[\begin{array}{cc} T^{-1/2} \sum_{t=2}^T \Delta y_t & T^{-1/2} \sum_{t=\lfloor \tau T \rfloor + 1}^T \Delta y_t \end{array}\right]
\left[\begin{array}{cc} 1 & (1-\tau) \\ (1-\tau) & (1-\tau) \end{array}\right]^{-1} \right. \\
& \left. \times \left[\begin{array}{c} T^{-1/2} \sum_{t=2}^T \Delta y_t \\ T^{-1/2} \sum_{t=\lfloor \tau T \rfloor + 1}^T \Delta y_t \end{array}\right] \right\} + o_p(1).
\end{aligned}$$

Denoting the expression within curly brackets as A, it is then seen that $\arg\min_{\tau \in \Lambda} S(1,\tau)$ is asymptotically equivalent to $\arg\max_{\tau \in \Lambda} A$. Next observe that

$$T^{-1/2} \textstyle\sum_{t=2}^T \Delta y_t = T^{-1/2} \sum_{t=2}^T \Delta u_t + \omega_\varepsilon \kappa T^{-1} \sum_{t=2}^T DU_t(\tau_0) \xrightarrow{d} \omega_\varepsilon \{W_c(1) + \kappa(1-\tau_0)\}$$

and that

$$\begin{aligned} T^{-1/2}\sum_{t=\lfloor \tau T\rfloor+1}^{T}\Delta y_t &= T^{-1/2}\sum_{t=\lfloor \tau T\rfloor+1}^{T}\Delta u_t + \omega_\varepsilon \kappa T^{-1}\sum_{t=\lfloor \tau T\rfloor+1}^{T} DU_t(\tau_0) \\ &\xrightarrow{d} \omega_\varepsilon[W_c(1) - W_c(\tau) + \kappa(1-\tau_0) - \kappa(\tau-\tau_0)\mathbb{I}_{\tau_0}^{\tau}]. \end{aligned}$$

On substituting we therefore obtain using the CMT that

$$\begin{aligned} \tilde{\tau} &= \arg\min_{\tau\in\Lambda} S(1,\tau) \\ &= \arg\min_{\tau\in\Lambda} \omega_\varepsilon^{-2} S(1,\tau) \\ &\xrightarrow{d} \arg\sup_{\tau\in\Lambda} \left[\begin{array}{cc} W_c(1)+\kappa(1-\tau_0) & W_c(1)-W_c(\tau)+\kappa(1-\tau_0)-\kappa(\tau-\tau_0)\mathbb{I}_{\tau_0}^{\tau} \end{array}\right] \\ &\quad \times \left[\begin{array}{cc} 1 & (1-\tau) \\ (1-\tau) & (1-\tau) \end{array}\right]^{-1} \left[\begin{array}{c} W_c(1)+\kappa(1-\tau_0) \\ W_c(1)-W_c(\tau)+\kappa(1-\tau_0)-\kappa(\tau-\tau_0)\mathbb{I}_{\tau_0}^{\tau} \end{array}\right]. \end{aligned}$$

To show that $\mathring{\tau}$ has the same limit distribution as $\tilde{\tau}$, on defining $\hat{\sigma}^2(\check{v}_{\tau,t}) := T^{-1}S(1,\tau)$, we note that $t_1(\tau)^2$ can be written as

$$\begin{aligned} t_1(\tau)^2 &= \frac{\hat{\sigma}^2(\check{v}_{\tau,t})}{\hat{\omega}^2(\check{v}_{\tau,t})} \frac{\check{\gamma}_\tau^2}{\hat{\sigma}^2(\check{v}_{\tau,t})(\sum_{t=2}^{T}\mathbf{x}_{1,\tau,t}\mathbf{x}_{1,\tau,t}')_{22}^{-1}} \\ &= \frac{\hat{\sigma}^2(\check{v}_{\tau,t})}{\hat{\omega}^2(\check{v}_{\tau,t})} T\left\{\frac{S(1)}{S(1,\tau)} - 1\right\} \\ &= \frac{S(1,\tau)}{\hat{\omega}^2(\check{v}_{\tau,t})}\left\{\frac{S(1)}{S(1,\tau)} - 1\right\} = \frac{1}{\hat{\omega}^2(\check{v}_{\tau,t})}\{S(1) - S(1,\tau)\} \end{aligned}$$

Since, as is shown below, $\hat{\omega}^2(\check{v}_{\tau,t}) \xrightarrow{p} \omega_\varepsilon^2$, then, in the limit $\arg\max_{\tau\in\Lambda}|t_1(\tau)|$ is identical to $\arg\min_{\tau\in\Lambda} S(1,\tau)$.

(ii) As is shown below, $(T/l)^{-1/2}t_0(\tau) \xrightarrow{d} \mathcal{D}_c^{t_0}(\tau_0,\tau,\kappa)$ and so

$$\begin{aligned} \hat{\tau} &= \arg\max_{\tau\in\Lambda}|t_0(\tau)| \\ &= \arg\max_{\tau\in\Lambda}\left|(T/l)^{-1/2}t_0(\tau)\right| \xrightarrow{d} \arg\sup_{\tau\in\Lambda}\left|\mathcal{D}_c^{t_0}(\tau_0,\tau,\kappa)\right| \end{aligned}$$

as follows via the CMT.

(iii) In view of (A.3), and using the fact that $y_1 = u_1$, as $\mu = \beta = 0$, we can write

$$\begin{aligned} S(\bar{\rho}_\tau,\tau) - \sum_{t=2}^{T}\Delta y_t^2 &= -\omega_\varepsilon^2 \left\{ \left[\begin{array}{cc} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) & b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau) \end{array}\right] \left[\begin{array}{cc} a_{\bar{c}_\tau} & m_{\bar{c}_\tau}(\tau) \\ m_{\bar{c}_\tau}(\tau) & d_{\bar{c}_\tau}(\tau) \end{array}\right]^{-1} \right. \\ &\quad \left. \times \left[\begin{array}{c} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) \\ b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau) \end{array}\right] \right\} + \sum_{t=2}^{T}(y_t - \bar{\rho}_\tau y_{t-1})^2 - \sum_{t=2}^{T}\Delta y_t^2 \\ &\quad + o_p(1). \end{aligned}$$

Now,

$\sum_{t=2}^{T}(y_t-\bar{\rho}_\tau y_{t-1})^2-\sum_{t=2}^{T}\Delta y_t^2 \xrightarrow{d} \omega_\varepsilon^2\bar{c}_\tau^2\int_0^1 W_c(r)^2dr+\bar{c}_\tau\{\omega_\varepsilon^2 W_c(1)^2-\sigma_\varepsilon^2\}+\omega_\varepsilon^2\kappa j_{\bar{c}_\tau}(\tau_0)+\omega_\varepsilon^2\kappa^2 k_{\bar{c}_\tau}(\tau_0)$

so that

$$
\begin{aligned}
\breve{\tau} &= \arg\min_{\tau\in\Lambda} S(\bar{\rho}_\tau,\tau) \\
&= \arg\min_{\tau\in\Lambda}\left\{S(\bar{\rho}_\tau,\tau)-\textstyle\sum_{t=2}^{T}\Delta y_t^2\right\} \\
&\xrightarrow{d} \arg\sup_{\tau\in\Lambda}\omega_\varepsilon^2\left[\begin{array}{cc} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) & b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau)\end{array}\right]\left[\begin{array}{cc} a_{\bar{c}_\tau} & m_{\bar{c}_\tau}(\tau) \\ m_{\bar{c}_\tau}(\tau) & d_{\bar{c}_\tau}(\tau)\end{array}\right]^{-1} \\
&\quad\times\left[\begin{array}{c} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) \\ b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau)\end{array}\right]-\omega_\varepsilon^2\bar{c}_\tau^2\textstyle\int_0^1 W_c(r)^2dr-\bar{c}_\tau\{\omega_\varepsilon^2 W_c(1)^2-\sigma_\varepsilon^2\} \\
&\quad-\omega_\varepsilon^2\kappa j_{c,\bar{c}_\tau}(\tau_0)-\omega_\varepsilon^2\kappa^2 k_{\bar{c}_\tau}(\tau_0) \\
&= \arg\sup_{\tau\in\Lambda}\left[\begin{array}{cc} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) & b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau)\end{array}\right]\left[\begin{array}{cc} a_{\bar{c}_\tau} & m_{\bar{c}_\tau}(\tau) \\ m_{\bar{c}_\tau}(\tau) & d_{\bar{c}_\tau}(\tau)\end{array}\right]^{-1} \\
&\quad\times\left[\begin{array}{c} b_{c,\bar{c}_\tau}+\kappa f_{c,\bar{c}_\tau}(\tau_0) \\ b_{c,\bar{c}_\tau}(\tau)+\kappa f_{c,\bar{c}_\tau}(\tau_0,\tau)\end{array}\right]-\bar{c}_\tau^2\textstyle\int_0^1 W_c(r)^2dr-\bar{c}_\tau\{W_c(1)^2-1\}-\kappa j_{c,\bar{c}_\tau}(\tau_0) \\
&\quad-\kappa^2 k_{\bar{c}_\tau}(\tau_0)-\bar{c}_\tau\left(1-\frac{\sigma_\varepsilon^2}{\omega_\varepsilon^2}\right).
\end{aligned}
$$

Proof of Theorem 4

(i) Considering a generic argument τ, we have

$$
\begin{aligned}
W_T(\tau) &= \frac{T^{-1}\sum_{t=1}^{T}(T^{-3/2}\hat{s}_{Rt})^2}{T^{-1}\sum_{t=1}^{T}\{T^{-3/2}\hat{s}_{Ut}\}^2}-1 \\
&\xrightarrow{d} \frac{\omega_\varepsilon^2\int_0^1 s_R(r,\tau_0,\kappa)^2dr}{\omega_\varepsilon^2\int_0^1 s_U(r,\tau_0,\tau,\kappa)^2dr}-1
\end{aligned}
$$

where $s_R(r,\tau_0,\kappa)$ and $s_U(r,\tau_0,\tau,\kappa)$ denote the continuous time residuals from the projection of $\int_0^r W_c(s)ds+\kappa\mathbb{I}_{\tau_0}^r\int_{\tau_0}^r(s-\tau_0)ds$ onto the space spanned by $\{r,\int_0^r sds\}$, and $\{r,\int_0^r sds,\mathbb{I}_\tau^r\int_\tau^r(s-\tau)ds\}$, respectively. The result then follows using Theorem 3 (i) and the CMT.

(ii) Examining $\hat{u}_{\tau,t}$, we find that

$$T^{-1/2}\hat{u}_{\tau,[rT]}\xrightarrow{d}\omega_\varepsilon N_c(r,\tau_0,\tau,\kappa).$$

Now, for generic τ, we find that

$$
\begin{aligned}
(T/l)^{-1}S_0 &= \frac{T^{-1}\sum_{t=1}^{T}\left(T^{-3/2}\sum_{i=1}^{t}\hat{u}_{\tau,i}\right)^2}{(lT)^{-1}\hat{\omega}^2(\hat{u}_{\tau,t})} \\
&\xrightarrow{d} \frac{\omega_\varepsilon^2\int_0^1\left(\int_0^r N_c(s,\tau_0,\tau,\kappa)ds\right)^2dr}{\omega_\varepsilon^2\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2dr}=\mathcal{D}_c^{S_0}(\tau_0,\tau,\kappa)
\end{aligned}
$$

the limit result for $(lT)^{-1}\hat{\omega}^2(\hat{u}_{\tau,t})$ following from a generalization of the approach of Kwiatkowski *et al.* (1992), and using equation (24) on p. 168 of that paper in the case of the Bartlett kernel. Also, as is straightforward to show,

$$\begin{aligned}(T/l)^{-1/2}t_0(\tau) &= \frac{T^{1/2}\hat{\gamma}_\tau}{\sqrt{(lT)^{-1}\hat{\omega}^2(\hat{u}_{\tau,t})(T^{-3}\sum_{t=1}^{T}\mathbf{x}_{0,\tau,t}\mathbf{x}'_{0,\tau,t})^{-1}_{33}}} \\ &\xrightarrow{d} \frac{\omega_\varepsilon \int_0^1 V_c(r,\tau_0,\kappa)F(r,\tau)dr}{\sqrt{\omega_\varepsilon^2\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2dr\int_0^1 F(r,\tau)^2dr}} = \mathcal{D}_c^{t_0}(\tau_0,\tau,\kappa).\end{aligned}$$

Next,

$$\begin{aligned}T^{-1/2}\textstyle\sum_{t=2}^{\lfloor rT\rfloor}\tilde{u}_{\tau,t} &= T^{-1/2}\textstyle\sum_{t=2}^{\lfloor rT\rfloor}\Delta y_t - T^{1/2}\check{\beta}_\tau T^{-1}\sum_{t=2}^{\lfloor rT\rfloor}1 - T^{1/2}\check{\gamma}_\tau T^{-1}\sum_{t=2}^{\lfloor rT\rfloor}DU_t(\tau) \\ &\xrightarrow{d} \omega_\varepsilon\{W_c(r)+\kappa(r-\tau_0)\mathbb{I}_{\tau_0}^r - G_{1c}(\tau_0,\tau,\kappa)r - G_{2c}(\tau_0,\tau,\kappa)(r-\tau)\mathbb{I}_\tau^r\} \\ &= \omega_\varepsilon Q_c(r,\tau_0,\tau,\kappa)\end{aligned}$$

using the fact that, as is easily shown, $T^{1/2}\check{\beta}_\tau \xrightarrow{d} \omega_\varepsilon G_{1,c}(\tau_0,\tau,\kappa)$ and $T^{1/2}\check{\gamma}_\tau \xrightarrow{d} \omega_\varepsilon G_{2,c}(\tau_0,\tau,\kappa)$. Then,

$$\begin{aligned}S_1 &= \frac{T^{-1}\sum_{t=2}^{T}\left(T^{-1/2}\sum_{i=2}^{t}\check{v}_{\tau,i}\right)^2}{T^{-2}(T-1)^2\hat{\omega}^2(\check{v}_{\tau,t})} \\ &\xrightarrow{d} \frac{\omega_\varepsilon^2\int_0^1 Q_c(r,\tau_0,\tau,\kappa)^2dr}{\omega_\varepsilon^2} = \mathcal{D}_c^{S_1}(\tau_0,\tau,\kappa)\end{aligned}$$

which uses the fact that $\hat{\omega}^2(\check{v}_{\tau,t}) \xrightarrow{p} \omega_\varepsilon^2$. This arises since $\hat{\omega}^2(\check{v}_{\tau,t})$ employs the sample autocovariances of $\check{v}_{\tau,t}$ and

$$\begin{aligned}T^{-1}\textstyle\sum_{t=j+2}^{T}\check{v}_{\tau,t}\check{v}_{\tau,t-j} &= T^{-1}\sum_{t=j+2}^{T}\Delta u_t\Delta u_{t-j} + o_p(1) \\ &= T^{-1}\sum_{t=j+2}^{T}\varepsilon_t\varepsilon_{t-j} + o_p(1) \xrightarrow{p} E(\varepsilon_t\varepsilon_{t-j}).\end{aligned}$$

Also,

$$\begin{aligned}t_1(\tau) &= \frac{T^{1/2}\check{\gamma}_\tau}{\sqrt{\hat{\omega}^2(\check{v}_{\tau,t})(T^{-1}\sum_{t=2}^{T}\mathbf{x}_{1,\tau,t}\mathbf{x}'_{1,\tau,t})^{-1}_{22}}} \\ &\xrightarrow{d} \frac{\omega_\varepsilon G_{2c}(\tau_0,\tau,\kappa)}{\sqrt{\omega_\varepsilon^2/\tau(1-\tau)}} = \mathcal{D}_c^{t_1}(\tau_0,\tau,\kappa)\end{aligned}$$

Combining the above results, together with those for $\hat{\tau}$ and $\mathring{\tau}$, yields the required limit via the CMT.

(iii) For the OLS estimator $\hat{\phi}$ in (12) we can show, along the same lines as the GLS counterpart in Theorem 1 (i), that

$$T\hat{\phi} \xrightarrow{d} \frac{\sigma_\eta}{\omega_\varepsilon}\frac{N_c(1,\tau_0,\tau,\kappa)^2 - N_c(0,\tau_0,\tau,\kappa)^2 - 1}{2\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2dr} = \frac{\sigma_\eta}{\omega_\varepsilon}\mathcal{D}_c^{\hat{\phi}}(\tau_0,\tau,\kappa).$$

Also, we find

$$\begin{aligned} T\hat{\sigma}_{\hat{\phi}} &\xrightarrow{d} \sqrt{\frac{\sigma_\eta^2}{\omega_\varepsilon^2 \int_0^1 N_c(r,\tau_0,\tau,\kappa)^2 dr}} = \frac{\sigma_\eta}{\omega_\varepsilon} \mathcal{D}_c^{\hat{\sigma}_{\hat{\phi}}}(\tau_0,\tau,\kappa) \\ t_{\hat{\phi}} &\xrightarrow{d} \frac{N_c(1,\tau_0,\tau,\kappa)^2 - N_c(0,\tau_0,\tau,\kappa)^2 - 1}{2\sqrt{\int_0^1 N_c(r,\tau_0,\tau,\kappa)^2 dr}} = \mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa). \end{aligned}$$

Next write, $T(\tilde{\rho}_M - 1) = T\hat{\phi} + C(t_{\hat{\phi}}) T\hat{\sigma}_{\hat{\phi}}$. The limit of the step function $C(t_{\hat{\phi}})$, as given in PY section 2.5, under our specification simplifies to

$$C(t_{\hat{\phi}}) \xrightarrow{d} \begin{cases} -\mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) & \text{if } \mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) > cv_P \\ -\frac{4}{\left(1+\frac{4-cv_P^2}{cv_P(10+cv_P)}\right)\mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa)+\frac{10(4-cv_P^2)}{cv_P(10+cv_P)}} & \text{if } -10 < \mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) \le cv_P \\ -\frac{4}{\mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa)} & \text{if } \mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) \le -10. \end{cases}$$

So,

$$\begin{aligned} C(t_{\hat{\phi}}) T\hat{\sigma}_{\hat{\phi}} &\xrightarrow{d} \frac{\sigma_\eta}{\omega_\varepsilon} \begin{cases} -\mathcal{D}_c^{\hat{\phi}}(\tau_0,\tau,\kappa) & \text{if } \mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) > cv_P \\ -\frac{4\mathcal{D}_c^{\hat{\sigma}_{\hat{\phi}}}(\tau_0,\tau,\kappa)}{\left(1+\frac{4-cv_P^2}{cv_P(10+cv_P)}\right)\mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa)+\frac{10(4-cv_P^2)}{cv_P(10+cv_P)}} & \text{if } -10 < \mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) \le cv_P \\ -\frac{4\mathcal{D}_c^{\hat{\sigma}_{\hat{\phi}}}(\tau_0,\tau,\kappa)}{\mathcal{D}_c^{t_{\hat{\rho}}}(\tau_0,\tau,\kappa)} & \text{if } \mathcal{D}_c^{t_{\hat{\phi}}}(\tau_0,\tau,\kappa) \le -10 \end{cases} \\ &= \frac{\sigma_\eta}{\omega_\varepsilon} \mathcal{D}_c^{BC}(\tau_0,\tau,\kappa) \end{aligned}$$

and therefore

$$\begin{aligned} T(\tilde{\rho}_M - 1) &\xrightarrow{d} \frac{\sigma_\eta}{\omega_\varepsilon} \{\mathcal{D}_c^{\hat{\sigma}_{\hat{\phi}}}(\tau_0,\tau,\kappa) + \mathcal{D}_c^{BC}(\tau_0,\tau,\kappa)\} \\ &= \frac{\sigma_\eta}{\omega_\varepsilon} \mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa). \end{aligned}$$

Finally, then, we obtain

$$T(\tilde{\rho}_{MS} - 1) \xrightarrow{d} \begin{cases} \frac{\sigma_\eta}{\omega_\varepsilon} \mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa) & \text{if } |\frac{\sigma_\eta}{\omega_\varepsilon} \mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa)| > g'_M \\ 0 & \text{if } |\frac{\sigma_\eta}{\omega_\varepsilon} \mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa)| \le g'_M. \end{cases}$$

Now we consider the behaviour of $W_{RQF}(\tau)$. We will evaluate this at the value $\rho' = 1 - c'/T$ where c' is arbitrary. First note that we may write

$$S(\rho') - S(\rho',\tau) = \frac{(T^{-1/2}\mathbf{r}'_{\rho',\tau,3}\mathbf{r}_{\rho'})^2}{T^{-1}\mathbf{r}'_{\rho',\tau,3}\mathbf{r}_{\rho',\tau,3}}$$

where $\mathbf{r}_{\rho',\tau,3}$ denotes the residuals from a regression of the final column of $\mathbf{Z}_{\rho',\tau}$, which we denote $\mathbf{Z}_{\rho',\tau,3}$, on $\mathbf{Z}_{\rho'}$, and $\mathbf{r}_{\rho'}$ denotes the residuals from a regression of $\mathbf{y}_{\rho'}$ on $\mathbf{Z}_{\rho'}$. Then

$$T^{-1}\mathbf{r}'_{\rho',\tau,3}\mathbf{r}_{\rho',\tau,3} = T^{-1}\mathbf{Z}'_{\rho',\tau,3}\mathbf{Z}_{\rho',\tau,3} - T^{-1/2}\mathbf{Z}'_{\rho',\tau,3}\mathbf{Z}_{\rho'}\boldsymbol{\Delta}_3^{-1}(\boldsymbol{\Delta}_3^{-1}\mathbf{Z}'_{\rho'}\mathbf{Z}_{\rho'}\boldsymbol{\Delta}_3^{-1})^{-1}T^{-1/2}\boldsymbol{\Delta}_3^{-1}\mathbf{Z}'_{\rho'}\mathbf{Z}_{\rho',\tau,3}$$

where

$$\boldsymbol{\Delta}_3 = \begin{bmatrix} 1 & 0 \\ 0 & T^{1/2} \end{bmatrix}.$$

This term contains no stochastic components and the following limits are easily verified

$$\begin{aligned}
T^{-1}\mathbf{Z}'_{\rho',\tau,3}\mathbf{Z}_{\rho',\tau,3} &\rightarrow (1-\tau) + c'^2(1-\tau)^3/3 + 2c'(1-\tau)^2/2 \\
\boldsymbol{\Delta}_3^{-1}\mathbf{Z}'_{\rho'}\mathbf{Z}_{\rho'}\boldsymbol{\Delta}_3^{-1} &\rightarrow \begin{bmatrix} 1+c' & 0 \\ 0 & 1+c'+c'^2/3 \end{bmatrix} \\
T^{-1/2}\boldsymbol{\Delta}_3^{-1}\mathbf{Z}'_{\rho'}\mathbf{Z}_{\rho',\tau,3} &\rightarrow \begin{bmatrix} 0 \\ (1-\tau) + c'(1-\tau) + c'^2(1-\tau)^2(2+\tau)/6 \end{bmatrix}
\end{aligned}$$

which, upon simplification, yields

$$\begin{aligned}
T^{-1}\mathbf{r}'_{\rho',\tau,3}\mathbf{r}_{\rho',\tau,3} &\rightarrow \tau(1-\tau)\left\{1 + c'^2\tau(1-\tau)\left(\frac{1}{3} - \frac{c'^2(1+\tau)^2}{36(1+c'+c'^2/3)}\right)\right\} \\
&= R_{c'}(\tau).
\end{aligned}$$

Now consider the term

$$T^{-1/2}\mathbf{r}'_{\rho',\tau,3}\mathbf{r}_{\rho'} = T^{-1/2}\mathbf{Z}'_{\rho',\tau,3}\mathbf{y}_{\rho'} - T^{-1/2}\mathbf{Z}'_{\rho',\tau,3}\mathbf{Z}_{\rho'}\boldsymbol{\Delta}_3^{-1}(\boldsymbol{\Delta}_3^{-1}\mathbf{Z}'_{\rho'}\mathbf{Z}_{\rho'}\boldsymbol{\Delta}_3^{-1})^{-1}\boldsymbol{\Delta}_3^{-1}\mathbf{Z}'_{\rho'}\mathbf{y}_{\rho'}.$$

We require

$$\begin{aligned}
T^{-1/2}\mathbf{Z}'_{\rho',\tau,3}\mathbf{y}_{\rho'} &= T^{-1/2}\textstyle\sum_{t=\tau T+1}^{T}\{DT_t(\tau) - (1-c'/T)DT_{t-1}(\tau)\}\{y_t - (1-c'/T)y_{t-1}\} \\
&= c'T^{-3/2}\textstyle\sum_{t=\tau T+1}^{T}(t-\tau T)\Delta u_t + c'^2T^{-5/2}\sum_{t=\tau T+1}^{T}(t-\tau T)u_{t-1} \\
&\quad + T^{-1/2}\textstyle\sum_{t=\tau T+1}^{T}\Delta u_t + c'T^{-3/2}\sum_{t=\tau T+1}^{T}u_{t-1} \\
&\quad + \omega_\varepsilon\kappa c'T^{-2}\textstyle\sum_{t=\tau T+1}^{T}(t-\tau T)DU_t(\tau_0) + \omega_\varepsilon\kappa c'^2T^{-3}\sum_{t=\tau T+1}^{T}(t-\tau T)DT_{t-1}(\tau_0) \\
&\quad + \omega_\varepsilon\kappa T^{-1}\textstyle\sum_{t=\tau T+1}^{T}DU_t(\tau_0) + \omega_\varepsilon\kappa c'T^{-2}\sum_{t=\tau T+1}^{T}DT_{t-1}(\tau_0) + o_p(1) \\
&\xrightarrow{d} \omega_\varepsilon[c'\textstyle\int_\tau^1(r-\tau)dW_c(r) + c'^2\int_\tau^1(r-\tau)W_c(r)dr + W_c(1) - W_c(\tau) + c'\int_\tau^1 W_c(r)dr \\
&\quad + \kappa c'\{(1-\tau_0)(1+\tau_0-2\tau)/2 + \mathbb{I}(\tau-\tau_0)^2/2\} \\
&\quad + \kappa c'^2[\{1-\tau_0-\mathbb{I}(\tau-\tau_0)\}^2\{2+\tau_0-3\tau+4\mathbb{I}(\tau-\tau_0)\}]/6 \\
&\quad + \kappa\{1-\tau_0-\mathbb{I}(\tau-\tau_0)\} + \kappa c'\{(1-\tau_0)^2/2 - \mathbb{I}(\tau-\tau_0)^2/2\}]
\end{aligned}$$

$$
\begin{aligned}
\boldsymbol{\Delta}_3^{-1}\mathbf{Z}'_{\rho'}\mathbf{y}_{\rho'} &= \begin{bmatrix} y_1 + \sum_{t=2}^{T}\{1-(1-c'/T)\}\{y_t-(1-c'/T)y_{t-1}\} \\ T^{-1/2}[y_1 + \sum_{t=2}^{T}\{t-(1-c'/T)(t-1)\}\{y_t-(1-c'/T)y_{t-1}\}] \end{bmatrix} \\
&= \begin{bmatrix} u_1 + o_p(1) \\ c'T^{-3/2}\sum_{t=2}^{T} t\Delta u_t + T^{-1/2}\sum_{t=2}^{T}\Delta u_t + c'^2T^{-5/2}\sum_{t=2}^{T} tu_{t-1} \\ +c'T^{-3/2}\sum_{t=2}^{T} u_{t-1} + \omega_\varepsilon\kappa c'T^{-2}\sum_{t=2}^{T} tDU_t(\tau_0) + \omega_\varepsilon\kappa T^{-1}\sum_{t=2}^{T} DU_t(\tau_0) \\ +\omega_\varepsilon\kappa c'^2T^{-3}\sum_{t=2}^{T} tDT_{t-1}(\tau_0) + \omega_\varepsilon\kappa c'T^{-2}\sum_{t=2}^{T} DT_{t-1}(\tau_0) + o_p(1) \end{bmatrix} \\
&\xrightarrow{d} \begin{bmatrix} u_1 \\ \omega_\varepsilon\{c'\int_0^1 r dW_c(r) + W_c(1) + c'^2\int_0^1 rW_c(r)dr + c'\int_0^1 W_c(r)dr \\ +\kappa c'(1-\tau_0^2)/2 + \kappa(1-\tau_0) + \kappa c'^2(1-\tau_0)^2(2+\tau_0)/6 + \kappa c'(1-\tau_0)^2/2\} \end{bmatrix}
\end{aligned}
$$

which on substituting and simplifying gives

$$
\begin{aligned}
T^{-1/2}\mathbf{r}'_{\rho',\tau,3}\mathbf{r}_{\rho'} &\xrightarrow{d} \omega_\varepsilon[c'\textstyle\int_\tau^1 (r-\tau)dW_c(r) + c'^2\int_\tau^1 (r-\tau)W_c(r)dr + W_c(1) - W_c(\tau) + c'\int_\tau^1 W_c(r)dr \\
&\quad +\kappa\{1-\tau_0-\mathbb{I}(\tau-\tau_0)\} + \kappa c'(1-\tau)(1-\tau_0) \\
&\quad +\kappa c'^2[\{1-\tau_0-\mathbb{I}(\tau-\tau_0)\}^2\{2+\tau_0-3\tau+4\mathbb{I}(\tau-\tau_0)\}]/6 - (1-\tau) \\
&\quad \times\{1+c'+c'^2/3 - c'^2\tau(1+\tau)/6\}\{c'\textstyle\int_0^1 r dW_c(r) + W_c(1) + c'^2\int_0^1 rW_c(r)dr \\
&\quad +c'\int_0^1 W_c(r)dr + \kappa(1-\tau_0)(1+c'+c'^2/3 - c'^2\tau_0(1+\tau_0)/6)\}/(1+c'+c'^2/3)] \\
&= \omega_\varepsilon H_{c,c'}(\tau_0,\tau,\kappa).
\end{aligned}
$$

Then, since $\hat{h}_\varepsilon \xrightarrow{p} \omega_\varepsilon^2$ we have

$$
\begin{aligned}
W_{RQF}(\tau) &= \frac{S(\rho') - S(\rho',\tau)}{\hat{h}_\varepsilon} \\
&\xrightarrow{d} \frac{H_{c,c'}(\tau_0,\tau,\kappa)^2}{R_{c'}(\tau)} = \mathcal{D}^W_{c,c'}(\tau_0,\tau,\kappa).
\end{aligned}
$$

We obtain the limit of $W_{RQF}(\tau)$ based on $T(\tilde{\rho}_{MS}-1)$ upon replacing c' with the limit of $-T(\tilde{\rho}_{MS}-1)$. That is,

$$
\begin{aligned}
W_{RQF}(\tau) &\xrightarrow{d} \begin{cases} \mathcal{D}^W_{c,-\frac{\sigma_\eta}{\omega_\varepsilon}\mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa)}(\tau_0,\tau,\kappa) & \text{if } |\frac{\sigma_\eta}{\omega_\varepsilon}\mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa)| > g'_M \\ \mathcal{D}^W_{c,0}(\tau_0,\tau,\kappa) & \text{if } |\frac{\sigma_\eta}{\omega_\varepsilon}\mathcal{D}_c^{\tilde{\rho}_M}(\tau_0,\tau,\kappa)| \le g'_M \end{cases} \\
&= M_c(\tau_0,\tau,\kappa,g'_M,\sigma_\eta/\omega_\varepsilon)
\end{aligned}
$$

and so

$$
\begin{aligned}
PY &= \log\left[T^{-1}\textstyle\sum_{\tau\in\Lambda}\exp\left\{\frac{1}{2}W_{RQF}(\tau)\right\}\right] \\
&\xrightarrow{d} \log\left[\int_{\tau\in\Lambda}\exp\{\tfrac{1}{2}M_c(\tau_0,\tau,\kappa,g'_M,\sigma_\eta/\omega_\varepsilon)d\tau\}\right] = \mathcal{D}_c^{PY}(\tau_0,\kappa,g'_M,\sigma_\eta/\omega_\varepsilon).
\end{aligned}
$$

References

Banerjee, A., Lumsdaine, R., Stock, J., 1992. Recursive and sequential tests of the unit root and trend break hypotheses: theory and international evidence. Journal of Business and Economics Statistics 10, 271–288.

Carrion-i-Silvestre, J.L., Kim, D., Perron, P., 2009. GLS-based unit root tests with multiple structural breaks both under the null and the alternative hypotheses. Econometric Theory 25, 1754–1792.

Chang, Y., Park, J.Y., 2002. On the asymptotics of ADF tests for unit roots. Econometric Reviews 21, 431–447.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64, 813–836.

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

Harvey, D.I., Leybourne, S.J., Taylor, A.M.R., 2009a. Unit root testing in practice: dealing with uncertainty over the trend and initial condition (with commentaries and rejoinder). Econometric Theory 25, 587–667.

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

Kejriwal, M., Perron, P., 2010. A sequential procedure to determine the number of breaks in trend with an integrated or stationary noise component. Journal of Time Series Analysis 31, 305–328.

Kim, D., Perron, P., 2009. Unit root tests allowing for a break in the trend function at an unknown time under both the null and alternative hypotheses. Journal of Econometrics 148, 1–13.

Kwiatkowski, D., Phillips, P.C.B., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? Journal of Econometrics 54, 159–178.

Ng, S., Perron, P., 2001. Lag length selection and the construction of unit root tests with good size and power. Econometrica 69, 1519–1554.

Perron, P., 1989. The Great Crash, the oil price shock, and the unit root hypothesis. Econometrica 57, 1361–1401.

Perron, P., 1997. Further evidence of breaking trend functions in macroeconomic variables. Journal of Econometrics 80, 355–385.

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

Perron, P., Rodríguez, G., 2003. GLS detrending, efficient unit root tests and structural change. Journal of Econometrics 115, 1–27.

Perron, P., Yabu, T., 2009a. Estimating deterministic trends with an integrated or stationary noise component. Journal of Econometrics 151, 56–69.

Perron, P., Yabu, T., 2009b. Testing for shifts in trend with an integrated or stationary noise component. Journal of Business and Economic Statistics 27, 369–396.

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

Roy, A., Fuller, W. A., 2001. Estimation for autoregressive processes with a root near one. Journal of Business and Economic Statistics 19, 482–493.

Stock, J.H., Watson, M.W., 1996. Evidence on structural instability in macroeconomic time series relations. Journal of Business and Economic Statistics 14, 11–30.

Stock, J.H., Watson, M.W., 1999. A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In: Engle, R.F., White, H. (Eds.), Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive W.J. Granger. Oxford University Press, Oxford, pp. 1–44.

Stock, J.H., Watson, M.W., 2005. Implications of dynamic factor analysis for VAR models. NBER Working Paper No. 11467.

Vogelsang, T.J., 1997. Wald-type tests for detecting breaks in the trend function of a dynamic time series. Econometric Theory, 13, 818–849.

Vogelsang, T.J., 1998. Trend function hypothesis testing in the presence of serial correlation. Econometrica 66, 123–148.

Yang, J., 2011. Break point estimates for a shift in trend: levels versus first differences. Forthcoming in Econometrics Journal.

Zivot, E., Andrews, D.W.K., 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. Finite sample maximum sizes and corresponding local break magnitudes.

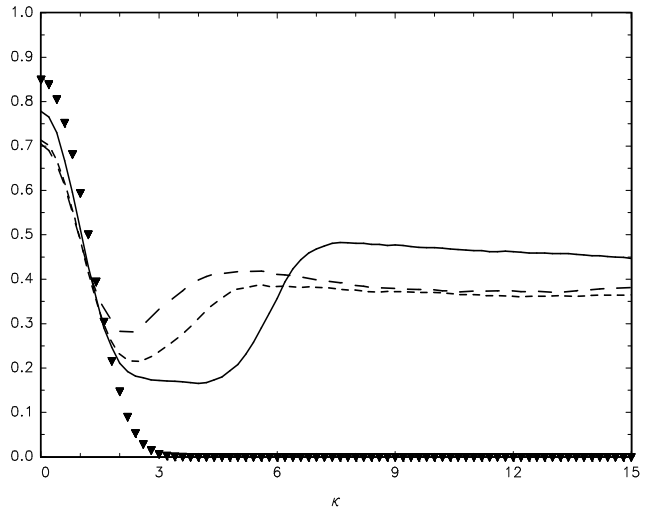
Panel A. $T = 150$						
	$\tau_0 = 0.3$		$\tau_0 = 0.5$		$\tau_0 = 0.7$	
	max. size	κ^*	max. size	κ^*	max. size	κ^*
$HHLT_{\bar{\tau}}$	0.111	1.8	0.110	1.0	0.108	1.4
$HHLT_{t_\lambda}$	0.110	1.8	0.113	1.2	0.111	1.4
CKP	0.052	0.0	0.052	1.2	0.053	8.2
$DF_{tb}^{GLS}(\tilde{\tau})$	0.068	1.8	0.069	1.8	0.069	0.4
$A(t_\lambda)$	0.097	2.2	0.102	3.4	0.105	5.4
$A(PY)$	0.079	2.2	0.087	7.8	0.089	6.6
$U_{t_\lambda}^{\delta}$	0.101	1.8	0.101	1.6	0.102	4.8
U_{PY}^{δ}	0.076	0.0	0.085	7.8	0.085	7.6
Panel B. $T = 300$						
	$\tau_0 = 0.3$		$\tau_0 = 0.5$		$\tau_0 = 0.7$	
	max. size	κ^*	max. size	κ^*	max. size	κ^*
$HHLT_{\bar{\tau}}$	0.086	0.2	0.091	4.2	0.089	0.2
$HHLT_{t_\lambda}$	0.086	0.2	0.087	0.4	0.088	1.0
CKP	0.060	8.6	0.074	7.8	0.065	8.6
$DF_{tb}^{GLS}(\tilde{\tau})$	0.061	0.0	0.062	0.2	0.064	0.4
$A(t_\lambda)$	0.082	5.2	0.092	4.4	0.085	5.6
$A(PY)$	0.070	4.4	0.081	6.2	0.078	6.2
$U_{t_\lambda}^{\delta}$	0.080	5.2	0.088	4.4	0.083	6.2
U_{PY}^{δ}	0.067	10.6	0.076	7.8	0.073	6.8

Table 2. Finite sample maximum sizes and corresponding local break magnitudes: $\tau_0 = 0.5$; tests with lag selection.

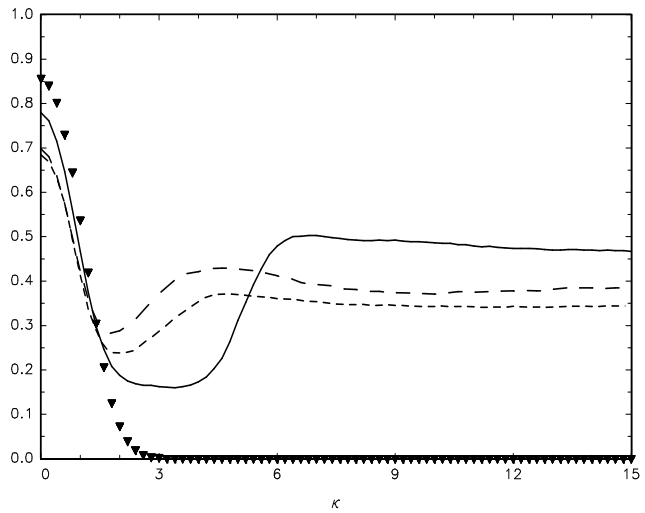
Panel A. $T = 150$						
	IID errors		AR(1) errors		MA(1) errors	
	max. size	κ^*	max. size	κ^*	max. size	κ^*
$HHLT_{\bar{\tau}}$	0.079	0.8	0.077	0.6	0.104	1.6
$HHLT_{t_\lambda}$	0.080	1.2	0.074	0.6	0.099	0.2
CKP	0.040	0.4	0.089	9.8	0.092	1.4
$DF_{tb}^{GLS}(\tilde{\tau})$	0.045	1.6	0.037	0.8	0.073	1.4
$A(t_\lambda)$	0.072	3.4	0.063	3.8	0.093	3.8
$A(PY)$	0.059	7.8	0.054	8.0	0.105	1.4
$U_{t_\lambda}^{\delta}$	0.072	3.2	0.065	1.6	0.095	1.4
U_{PY}^{δ}	0.057	7.8	0.053	8.0	0.122	0.2
Panel B. $T = 300$						
	IID errors		AR(1) errors		MA(1) errors	
	max. size	κ^*	max. size	κ^*	max. size	κ^*
$HHLT_{\bar{\tau}}$	0.068	4.4	0.070	0.0	0.084	0.2
$HHLT_{t_\lambda}$	0.065	0.0	0.068	0.0	0.086	0.2
CKP	0.053	7.8	0.079	8.0	0.085	0.0
$DF_{tb}^{GLS}(\tilde{\tau})$	0.041	0.2	0.041	0.0	0.055	0.2
$A(t_\lambda)$	0.064	4.4	0.065	4.6	0.072	4.2
$A(PY)$	0.057	6.2	0.058	6.0	0.081	0.4
$U_{t_\lambda}^{\delta}$	0.064	4.4	0.064	4.2	0.076	0.0
U_{PY}^{δ}	0.053	9.8	0.055	6.6	0.097	0.0

Table 3. Asymptotic maximum sizes and corresponding local break magnitudes.

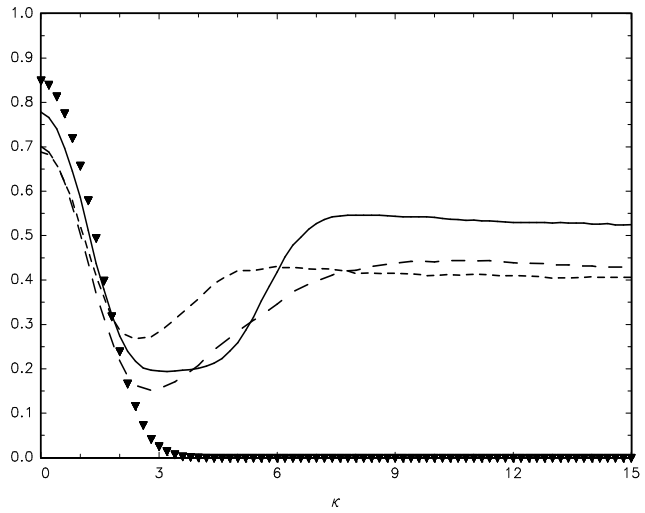
	$\tau_0 = 0.3$		$\tau_0 = 0.5$		$\tau_0 = 0.7$	
	max. size	κ^*	max. size	κ^*	max. size	κ^*
$HHLT_{\bar{\tau}}$	0.081	0.0	0.081	0.0	0.081	0.0
$HHLT_{t_\lambda}$	0.068	0.0	0.068	7.6	0.068	0.0
CKP	0.083	9.4	0.094	8.2	0.081	9.2
$DF_{tb}^{GLS}(\tilde{\tau})$	0.052	1.4	0.051	1.0	0.050	1.0
$A(t_\lambda)$	0.067	7.6	0.072	5.4	0.063	7.4
$A(PY)$	0.066	7.8	0.070	6.6	0.062	7.8
$U_{t_\lambda}^{\delta}$	0.065	8.2	0.069	6.4	0.063	8.2
U_{PY}^{δ}	0.064	9.2	0.068	8.0	0.061	8.8



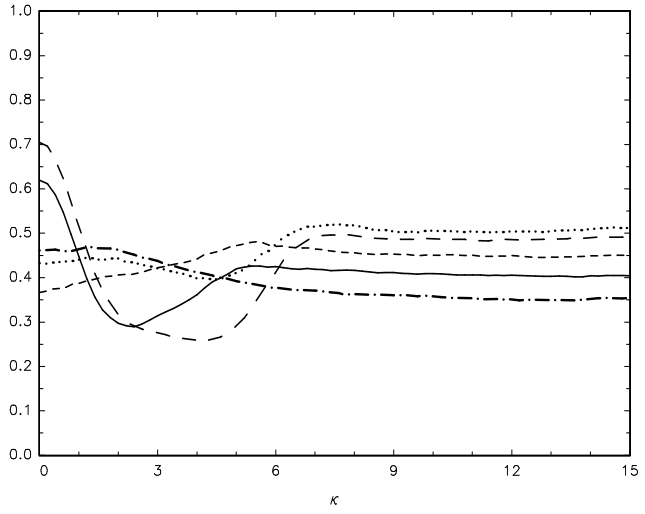
(a) $\tau_0 = 0.3$



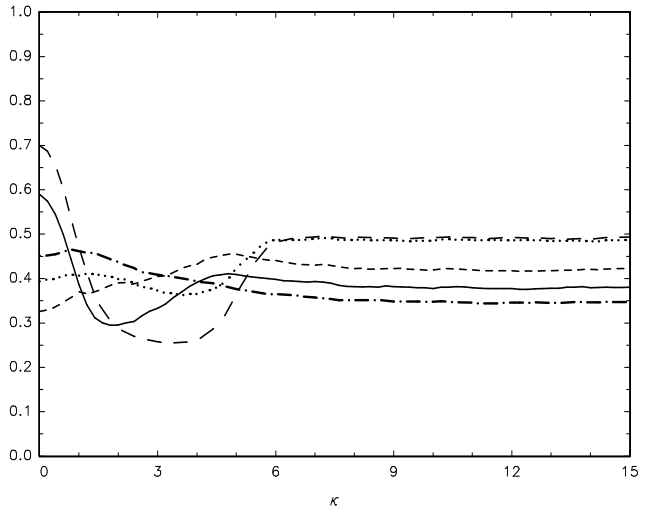
(b) $\tau_0 = 0.5$



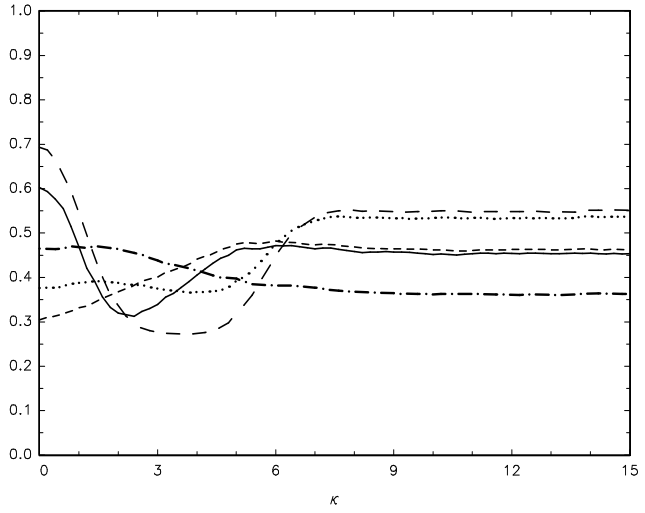
(c) $\tau_0 = 0.7$



(d) $\tau_0 = 0.3$

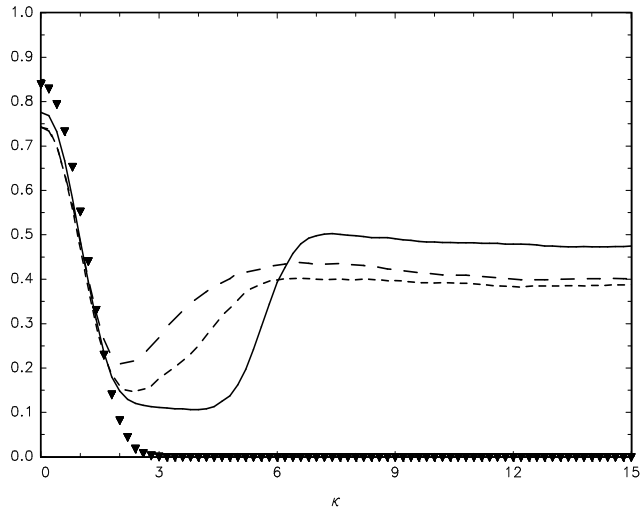


(e) $\tau_0 = 0.5$

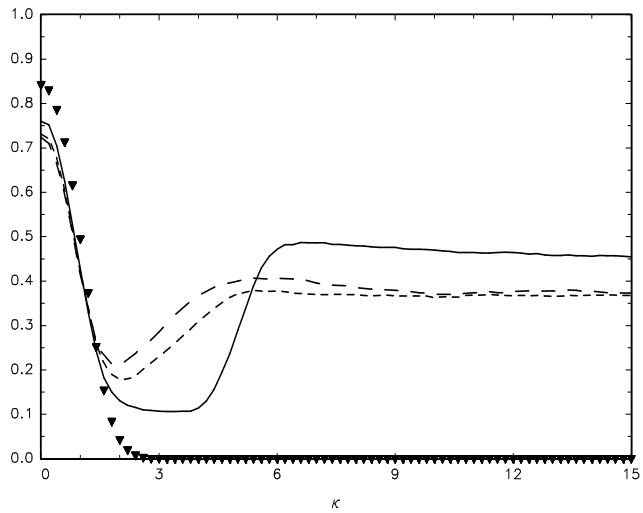


(f) $\tau_0 = 0.7$

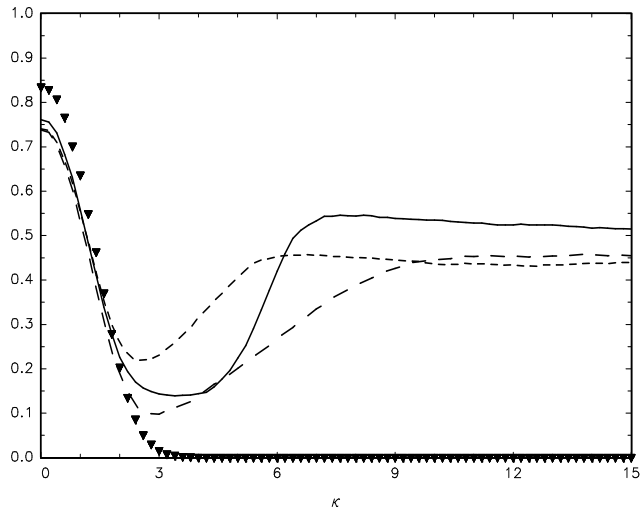
Figure 1. Finite sample size-adjusted power: $T = 150$, $c = 20$
(a)-(c): DF_t^{GLS}: ▼ , $HHLT_{\bar{\tau}}$: – – , $HHLT_{t_\lambda}$: - - - , CKP: ——
(d)-(f): $DF_{tb}^{GLS}(\tilde{\tau})$: – · – , $A(t_\lambda)$: - - - , $A(PY)$: · · · , $U_{t_\lambda}^{\delta}$: ——, U_{PY}^{δ}: – –



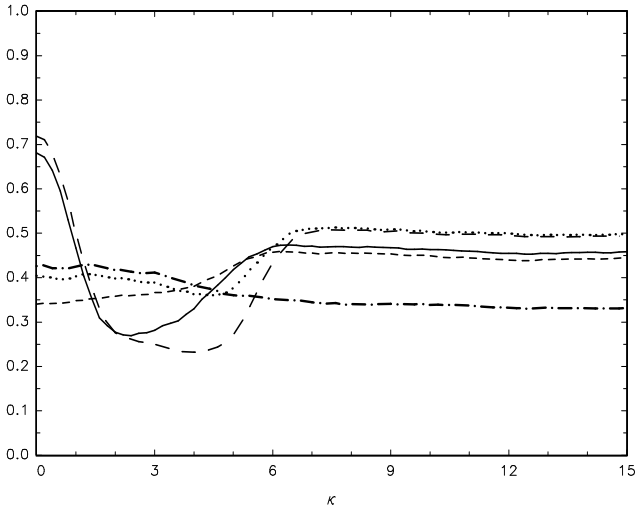
(a) $\tau_0 = 0.3$



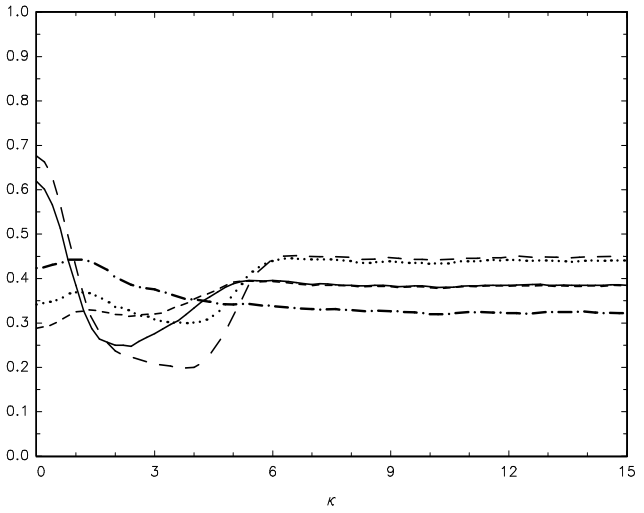
(b) $\tau_0 = 0.5$



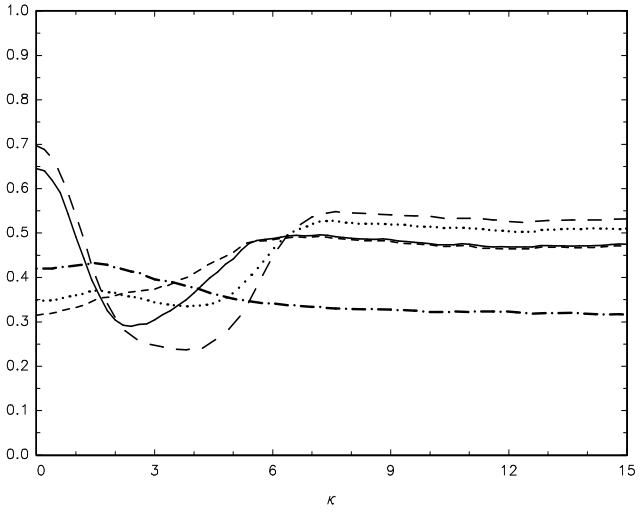
(c) $\tau_0 = 0.7$



(d) $\tau_0 = 0.3$

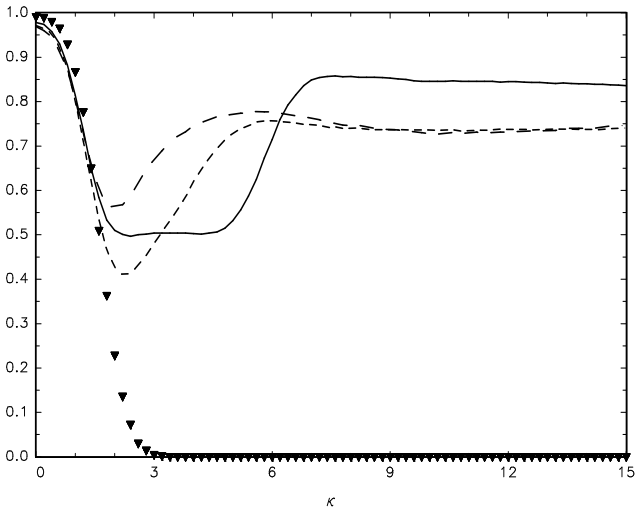


(e) $\tau_0 = 0.5$

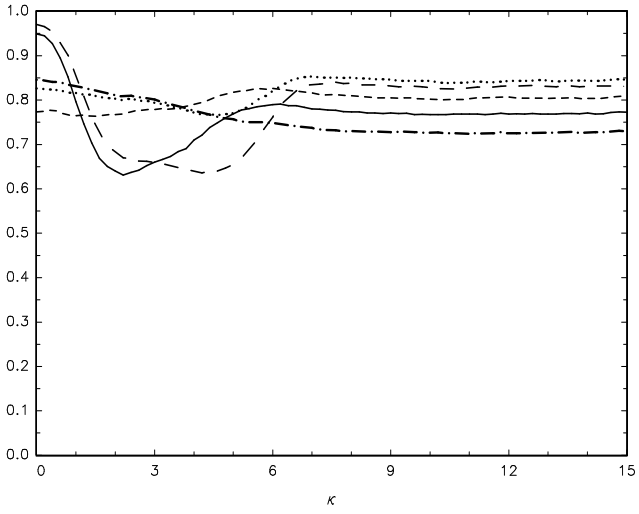


(f) $\tau_0 = 0.7$

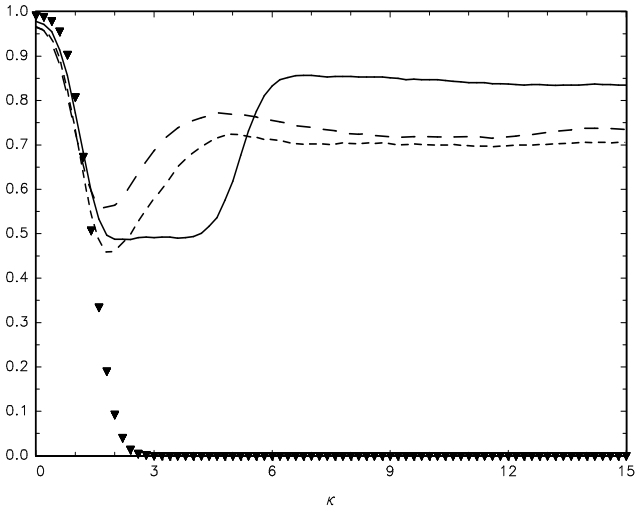
Figure 2. Finite sample size-adjusted power: $T = 300$, $c = 20$
(a)-(c): DF_t^{GLS}: ▼ , $HHLT_{\bar{\tau}}$: – – , $HHLT_{t_\lambda}$: - - - , CKP: ——
(d)-(f): $DF_{tb}^{GLS}(\tilde{\tau})$: – · – , $A(t_\lambda)$: - - - , $A(PY)$: · · · , $U_{t_\lambda}^{\delta}$: ——, U_{PY}^{δ}: – –



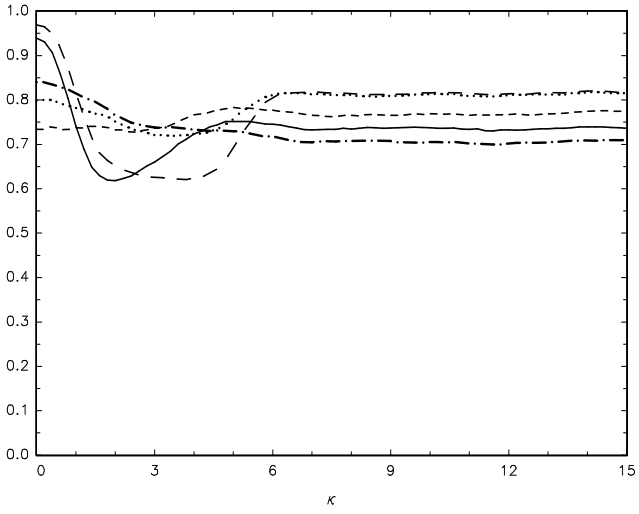
(a) $\tau_0 = 0.3$



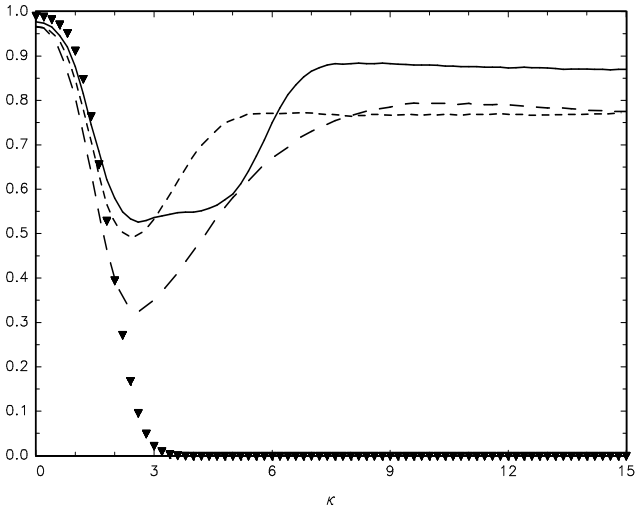
(d) $\tau_0 = 0.3$



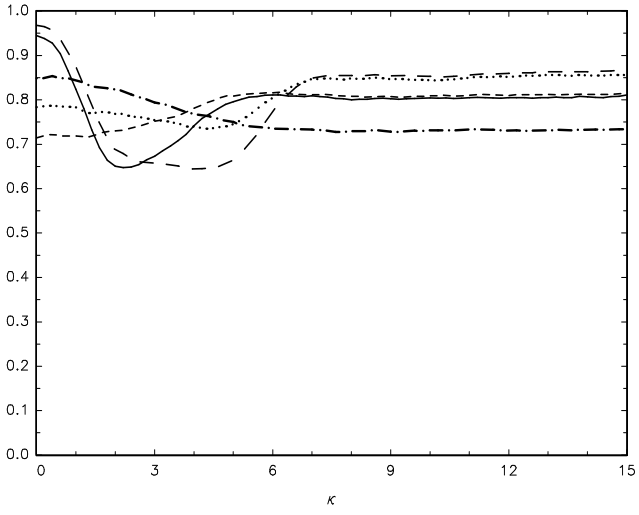
(b) $\tau_0 = 0.5$



(e) $\tau_0 = 0.5$

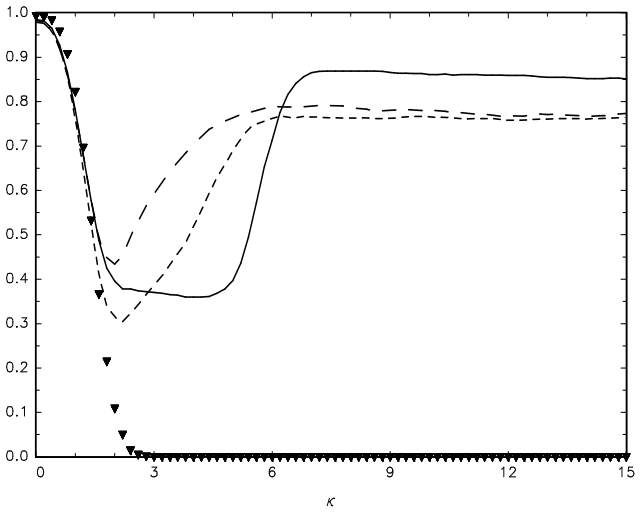


(c) $\tau_0 = 0.7$

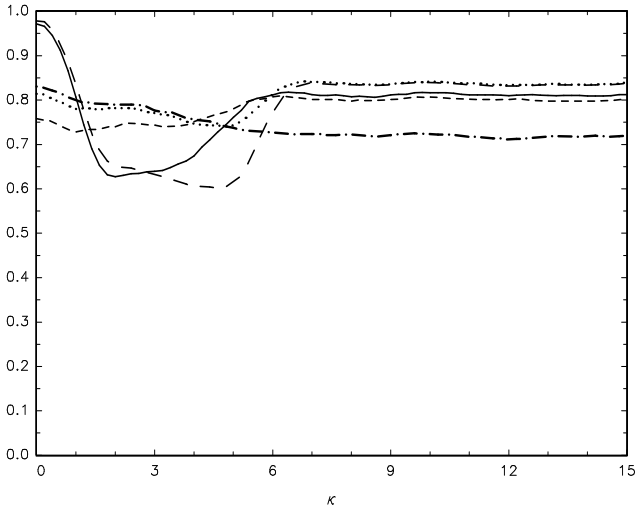


(f) $\tau_0 = 0.7$

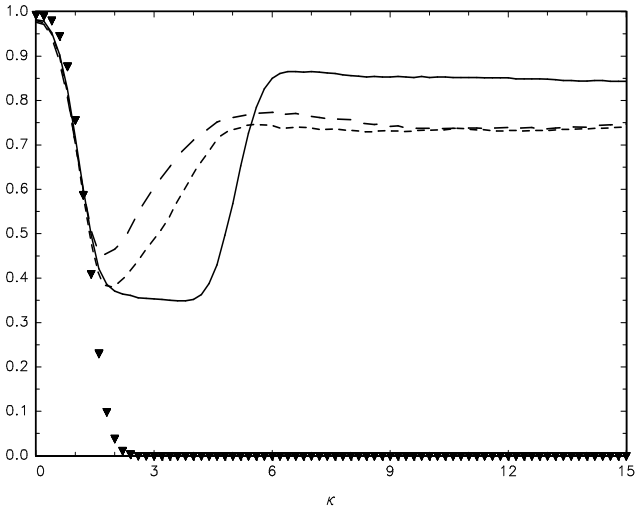
Figure 3. Finite sample size-adjusted power: $T = 150$, $c = 30$
(a)-(c): DF_t^{GLS}: ▼ , $HHLT_{\bar{\tau}}$: – – , $HHLT_{t_\lambda}$: - - - , CKP: ——
(d)-(f): $DF_{tb}^{GLS}(\tilde{\tau})$: – · – , $A(t_\lambda)$: - - - , $A(PY)$: · · · , $U_{t_\lambda}^{\delta}$: ——, U_{PY}^{δ}: – –



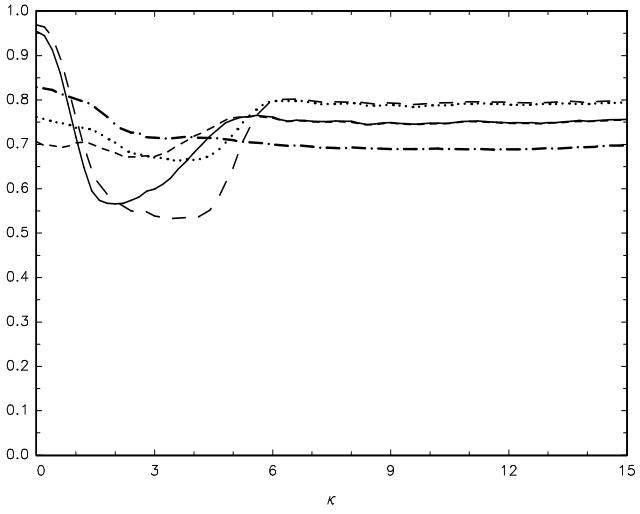
(a) $\tau_0 = 0.3$



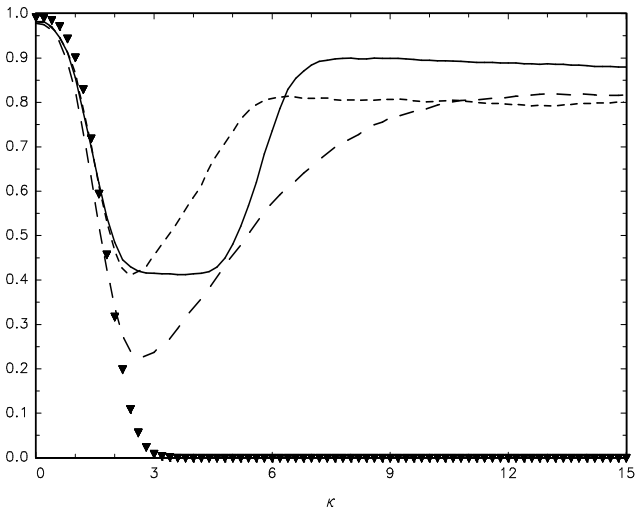
(d) $\tau_0 = 0.3$



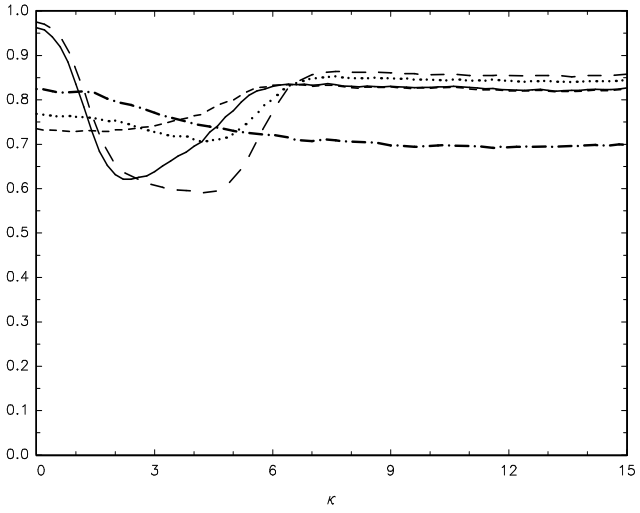
(b) $\tau_0 = 0.5$



(e) $\tau_0 = 0.5$

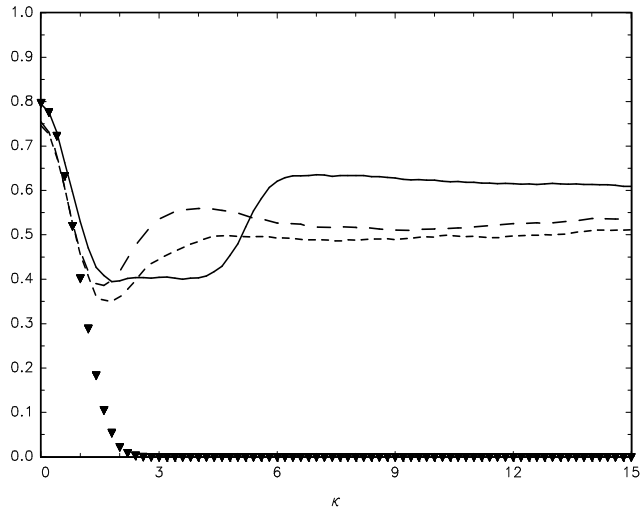


(c) $\tau_0 = 0.7$

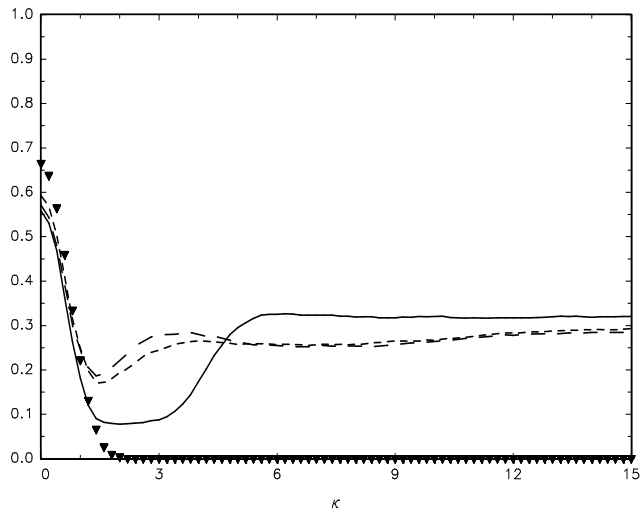


(f) $\tau_0 = 0.7$

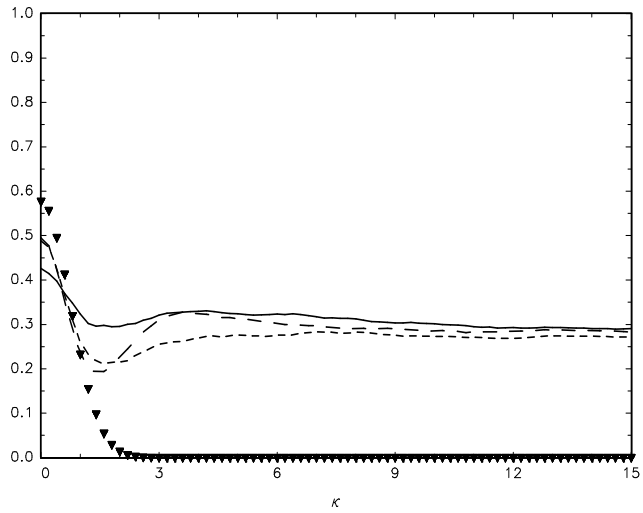
Figure 4. Finite sample size-adjusted power: $T = 300$, $c = 30$
(a)-(c): DF_t^{GLS}: ▼ , $HHLT_{\bar{\tau}}$: – – , $HHLT_{t_\lambda}$: - - - , CKP: ——
(d)-(f): $DF_{tb}^{GLS}(\tilde{\tau})$: – · – , $A(t_\lambda)$: - - - , $A(PY)$: · · · , $U_{t_\lambda}^{\delta}$: ——, U_{PY}^{δ}: – –



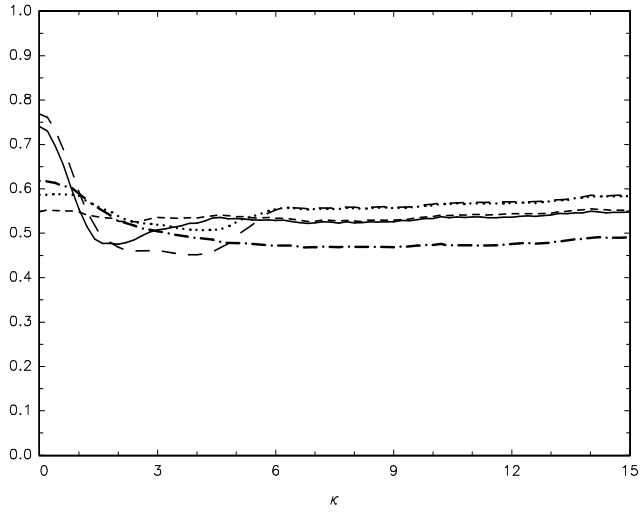
(a) IID errors



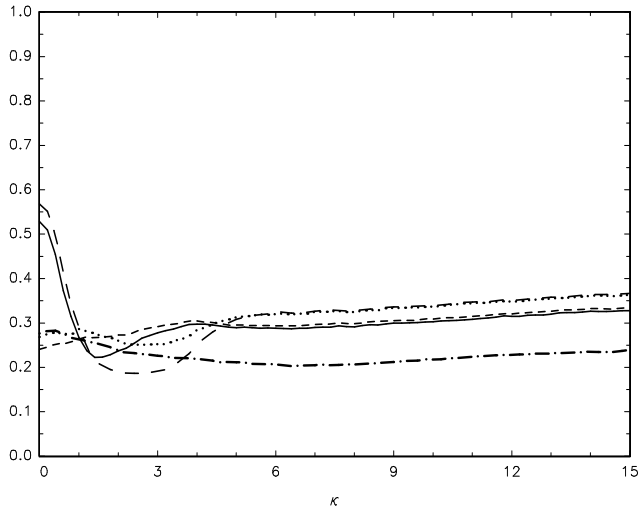
(b) AR(1) errors



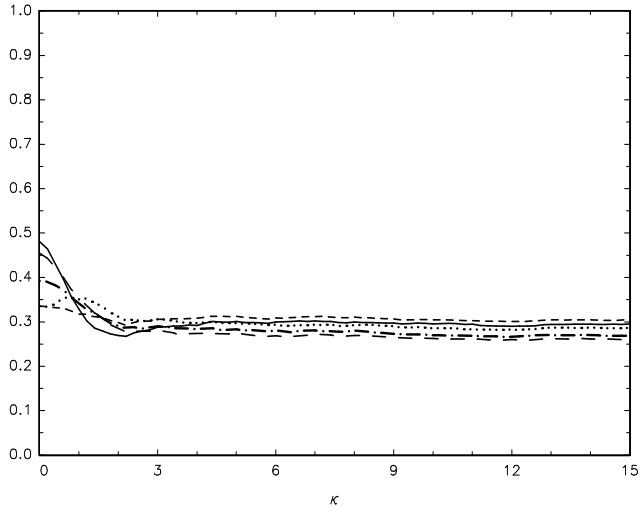
(c) MA(1) errors



(d) IID errors

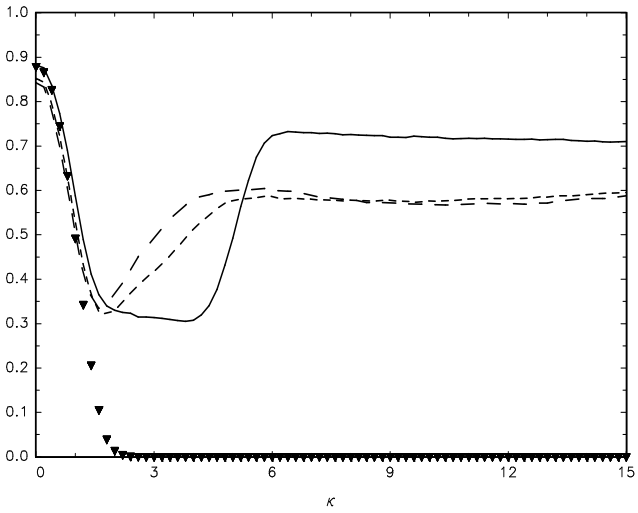


(e) AR(1) errors

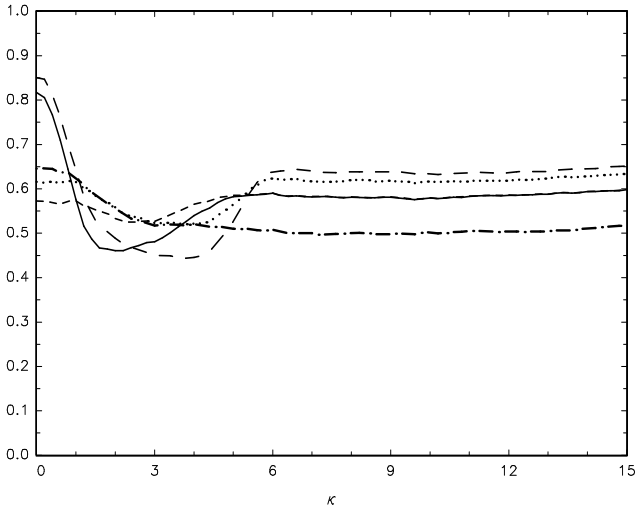


(f) MA(1) errors

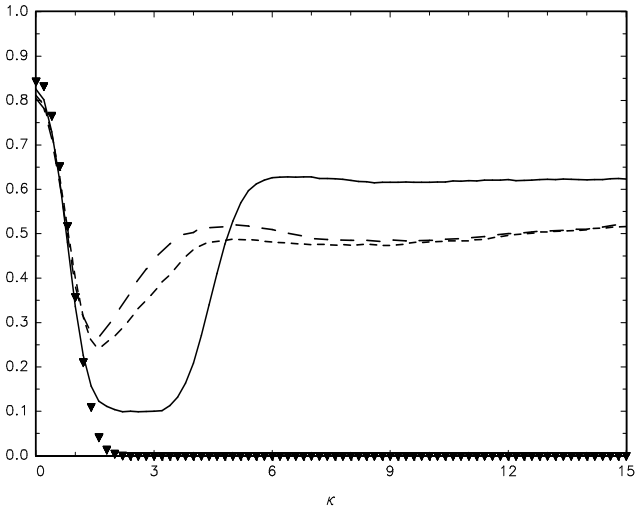
Figure 5. Finite sample size-adjusted power: $T = 150$, $c = 30$, $\tau_0 = 0.5$
(a)-(c): DF_t^{GLS}: ▼ , $HHLT_{\bar{\tau}}$: – – , $HHLT_{t_\lambda}$: - - - , CKP: ——
(d)-(f): $DF_{tb}^{GLS}(\tilde{\tau})$: – · – , $A(t_\lambda)$: - - - , $A(PY)$: · · · , $U_{t_\lambda}^{\delta}$: ——, U_{PY}^{δ}: – –



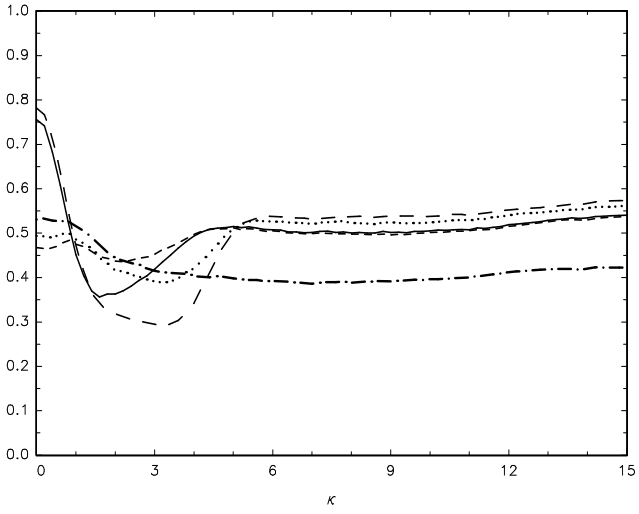
(a) IID errors



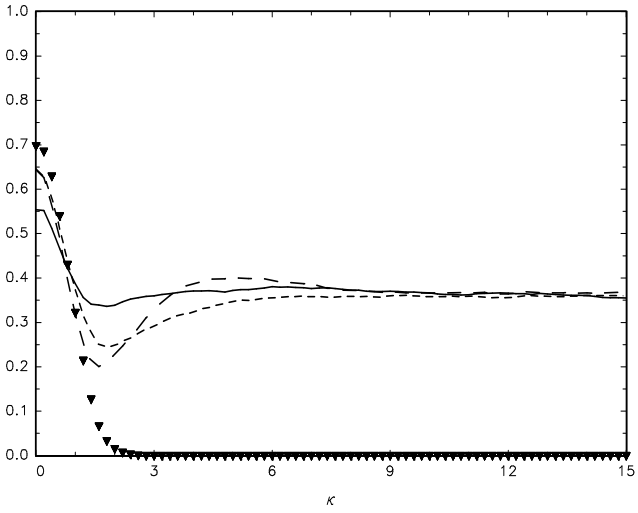
(d) IID errors



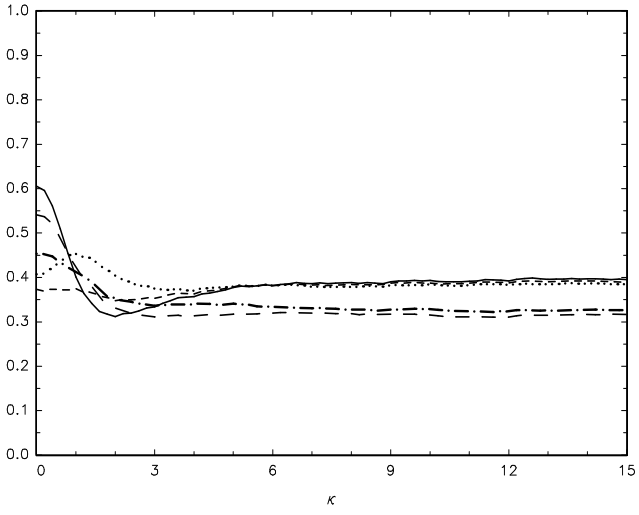
(b) AR(1) errors



(e) AR(1) errors

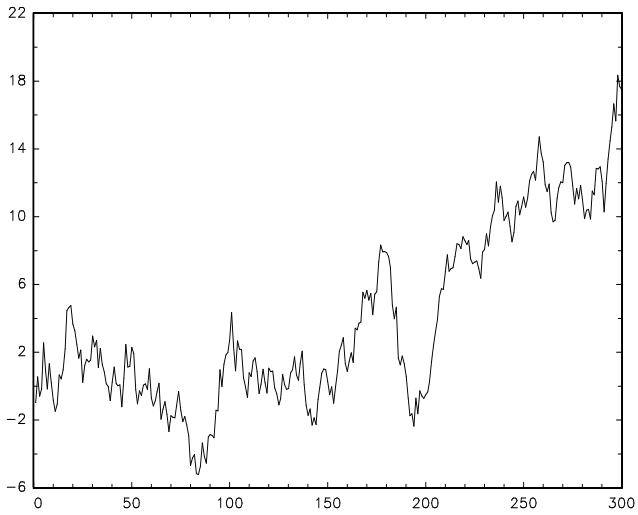


(c) MA(1) errors

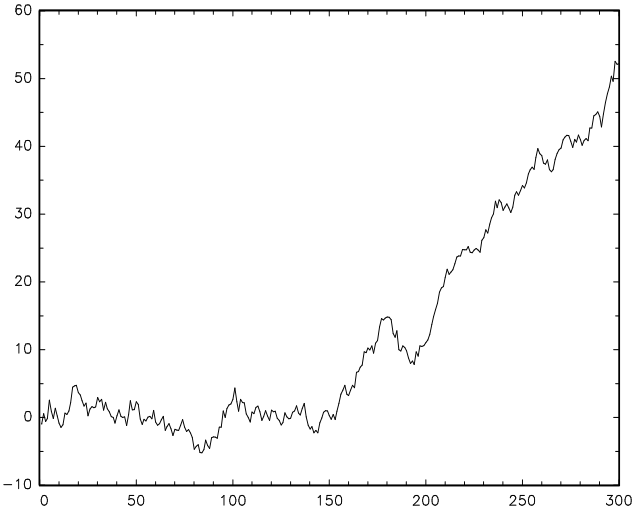


(f) MA(1) errors

Figure 6. Finite sample size-adjusted power: $T = 300$, $c = 30$, $\tau_0 = 0.5$
(a)-(c): DF_t^{GLS}: ▼ , $HHLT_{\bar{\tau}}$: – – , $HHLT_{t_\lambda}$: - - - , CKP: ——
(d)-(f): $DF_{tb}^{GLS}(\tilde{\tau})$: – · – , $A(t_\lambda)$: - - - , $A(PY)$: · · · , $U_{t_\lambda}^{\delta}$: ——, U_{PY}^{δ}: – –

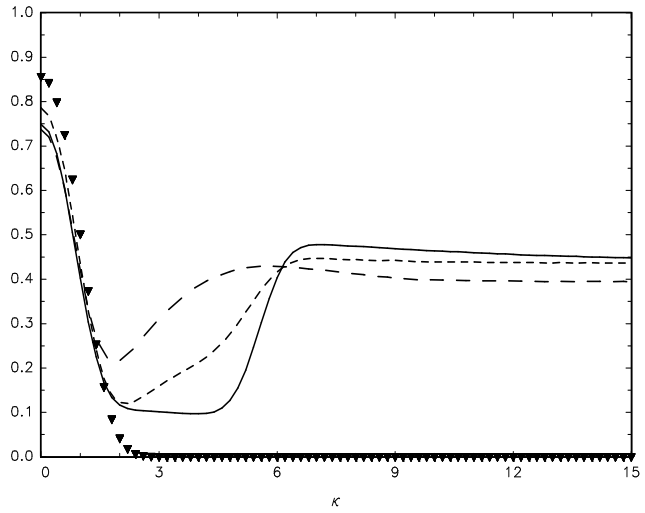


(a) $T = 300$, $c = 20$, $\tau_0 = 0.5$, $\kappa = 2$

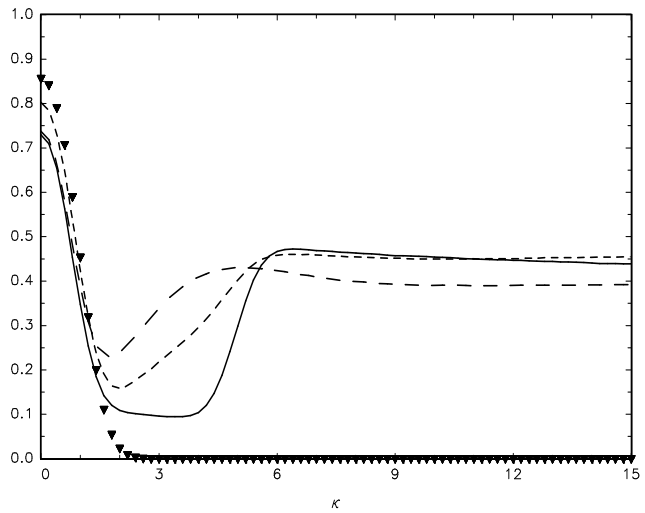


(b) $T = 300$, $c = 20$, $\tau_0 = 0.5$, $\kappa = 6$

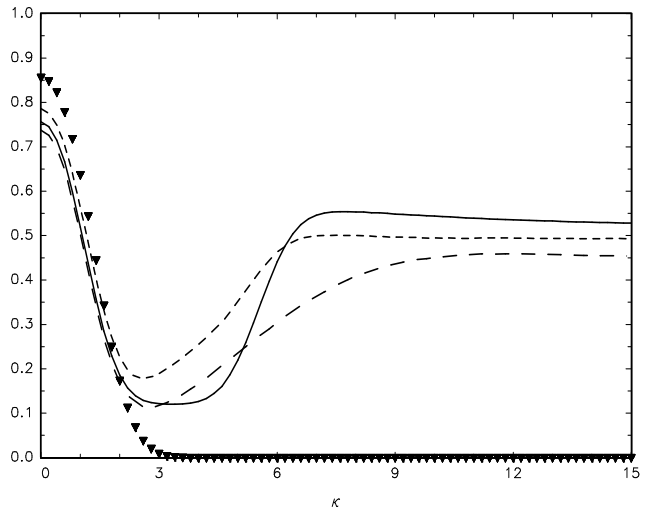
Figure 7. Simulated series with 1 break



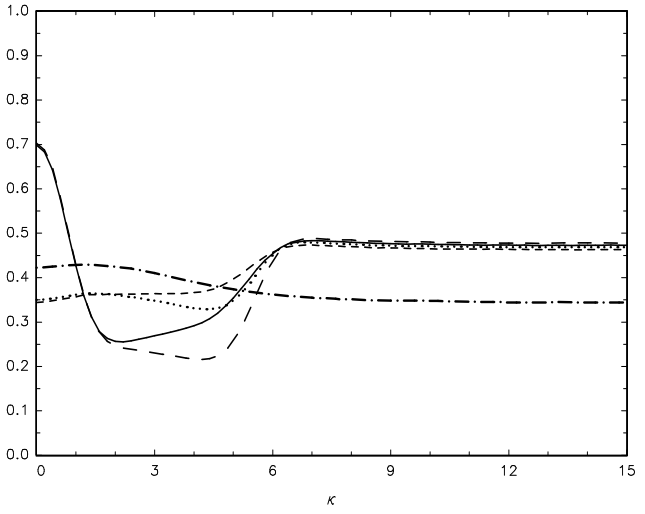
(a) $\tau_0 = 0.3$



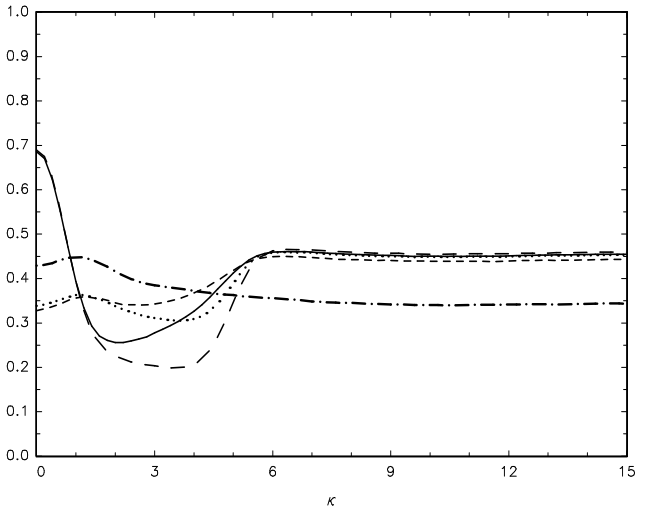
(b) $\tau_0 = 0.5$



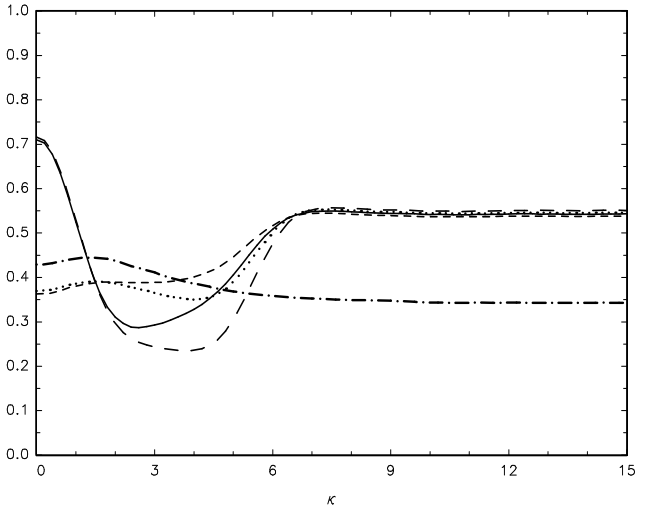
(c) $\tau_0 = 0.7$



(d) $\tau_0 = 0.3$

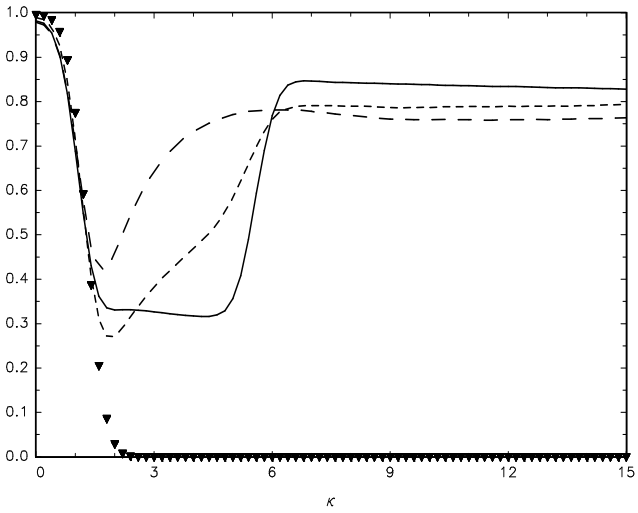


(e) $\tau_0 = 0.5$

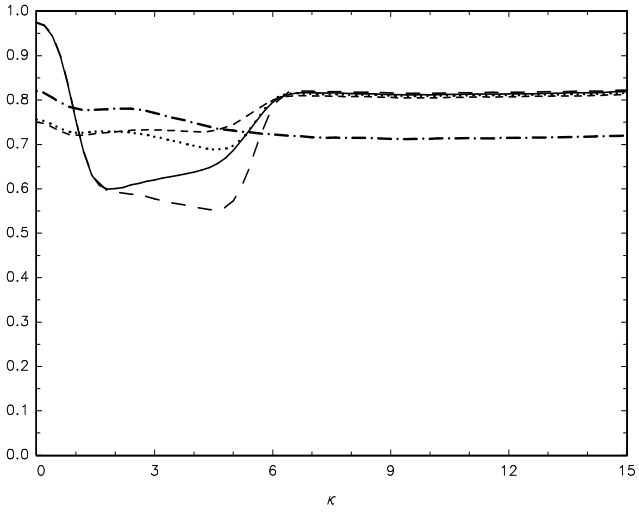


(f) $\tau_0 = 0.7$

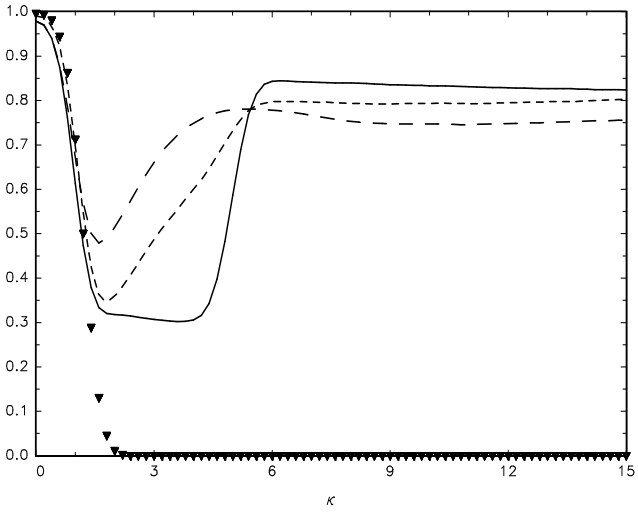
Figure 8. Asymptotic size-adjusted local power: $c = 20$
(a)-(c): DF_t^{GLS}: ▼ , $HHLT_{\bar{\tau}}$: – – , $HHLT_{t_\lambda}$: - - - , CKP: ——
(d)-(f): $DF_{tb}^{GLS}(\tilde{\tau})$: – · – , $A(t_\lambda)$: - - - , $A(PY)$: · · · , $U_{t_\lambda}^{\delta}$: ——, U_{PY}^{δ}: – –



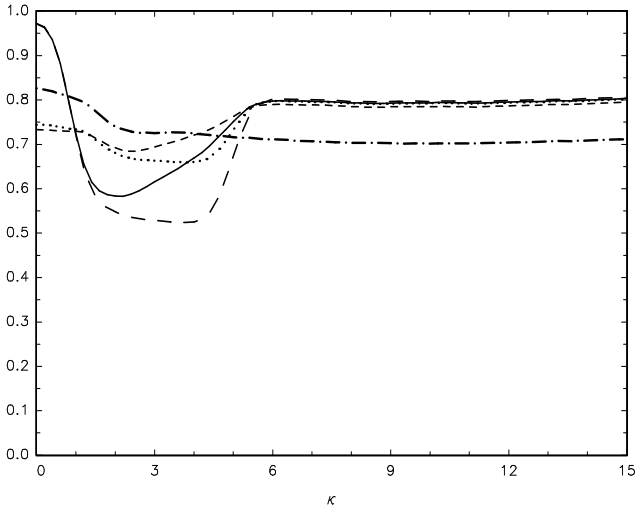
(a) $\tau_0 = 0.3$



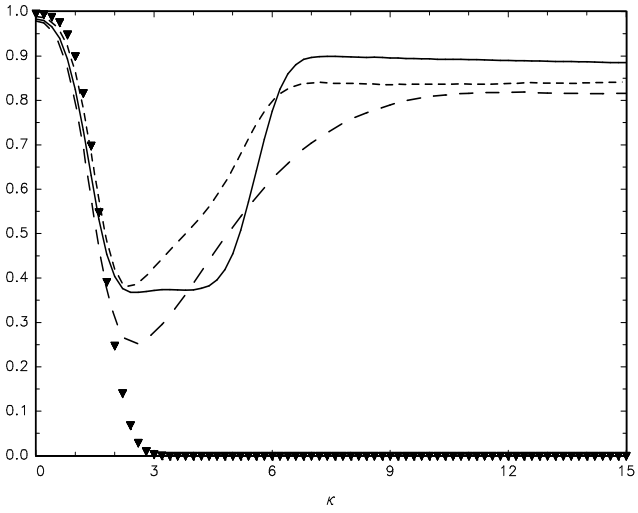
(d) $\tau_0 = 0.3$



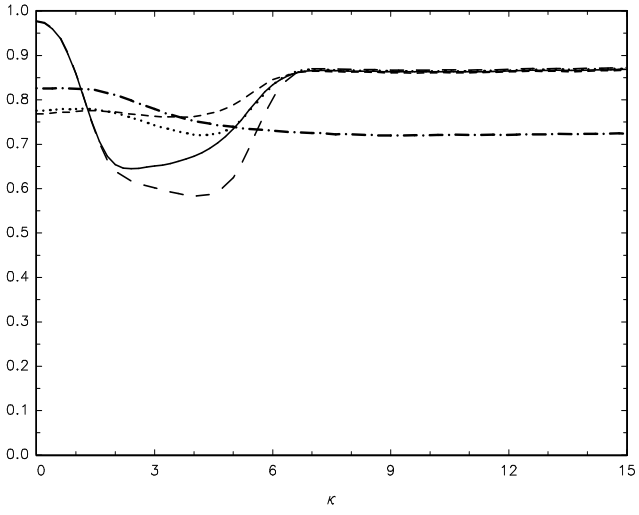
(b) $\tau_0 = 0.5$



(e) $\tau_0 = 0.5$



(c) $\tau_0 = 0.7$



(f) $\tau_0 = 0.7$

Figure 9. Asymptotic size-adjusted local power: $c = 30$
(a)-(c): DF_t^{GLS}: ▼ , $HHLT_{\bar{\tau}}$: – – , $HHLT_{t_\lambda}$: - - - , CKP: ——
(d)-(f): $DF_{tb}^{GLS}(\tilde{\tau})$: – · – , $A(t_\lambda)$: - - - , $A(PY)$: · · · , $U_{t_\lambda}^{\delta}$: ——, U_{PY}^{δ}: – –