Abstract

Trend breaks appear to be prevalent in macroeconomic time series, and unit root tests therefore need to make allowance for these if they are to avoid the catastrophic effects that unmodelled trend breaks have on power. Carrion-i-Silvestre et al. (2009) propose a pre-test-based approach which delivers near asymptotically efficient unit root inference both when breaks do not occur and where multiple breaks occur, provided the break magnitudes are fixed. Unfortunately, however, the fixed magnitude trend break asymptotic theory does not predict well the finite sample power functions of these tests, and power can be very low indeed (even zero) for the magnitudes of trend breaks typically observed in practice. In response to this problem we propose a unit root test that allows for multiple breaks in trend, obtained by taking the infimum of the sequence (across all candidate break points in a trimmed range) of local GLS detrended augmented Dickey-Fuller-type statistics. We show that this procedure has power that is robust to the magnitude of any trend breaks, thereby retaining good finite sample power in the presence of plausibly-sized breaks. We also demonstrate that, unlike the OLS detrended infimum tests of Zivot and Andrews (1992), these tests display no tendency to spuriously reject in the limit when fixed magnitude trend breaks occur under the unit root null.

Keywords: Unit root test; Multiple breaks in trend; Minimum Dickey-Fuller test; Local GLS detrending.

JEL Classification: C22.

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). In a seminal paper, Perron (1989) shows
that failure to account for trend breaks present in the data results in unit root tests with zero power, even asymptotically. Consequently, when testing for a unit root it has become a matter of regular practice to allow for this kind of deterministic structural change. While Perron (1989) initially treated the location of a potential single trend break as known, subsequent 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) [ZA], Banerjee *et al.* (1992), Perron (1997) and Perron and Rodriguez (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 local GLS detrending that is near asymptotically efficient (in the usual sense that the tests lie arbitrarily close to the asymptotic Gaussian local power envelope) when no trend break is present. When a single trend break is known to be present, it is now a test based on PR’s local GLS detrended ADF statistic which allows for a trend break that is asymptotically efficient. The latter holds provided the break point is known, or is unknown but can be dated endogenously with sufficient precision. However, when a trend break does not occur the PR test is not asymptotically efficient, the redundant trend break regressor compromising power. Moreover, the asymptotic critical values for the PR test based on an estimated break point differ markedly according to whether a trend break occurs or not. Precisely, the no break case critical values are substantially left-shifted relative to their break case counterparts. Since the PR test is left-tailed, the no break case critical values need to be employed to avoid over-sizing in the no break case. Consequently, when a break does occur (and can be dated with sufficient precision), the PR test will be under-sized, with an associated loss in power under the alternative, relative to the test based on the non-conservative break critical values. The underlying problem is then essentially one of uncertainty as to whether trend breaks exist in the data or not.

In work which allows for multiple possible breaks in trend, Carrion-i-Silvestre *et al.* (2009) [CKP] propose a solution to the issues raised above.\(^1\) The test procedure outlined in CKP utilizes auxiliary statistics to detect the presence of trend breaks occurring at unknown points in the sample and then uses the outcome of the detection step to indicate whether or not the unit root test employed should include trend breaks in the deterministic specification. CKP use the (multiple) trend break test of Kejriwal and Perron (2010) to select between either one of the standard GLS de-trended M tests of Ng and Perron (2001) or the (feasible) likelihood ratio test of ERS (in each case allowing for a constant plus linear trend), and the corresponding test (in each case the multiple breaks generalisation of the single break test from PR) allowing for the number of trend breaks identified by the Kejriwal and Perron (2010) test, and with their locations estimated as outlined in section 5.1 of CKP.

Assuming the trend break magnitudes to be fixed (independent of sample size) CKP show that their method achieves near asymptotically efficient unit root inference in both the no trend break and trend break environments. In the latter case this occurs because their test employs non-conservative critical values by virtue of the convergence of their break point estimators to the unknown break

---

\(^1\)Alternative procedures along the same lines as the procedure outlined in CKP are also proposed in Harris *et al.* (2009), although these allow only for the possibility of a single break in trend. Consequently, given our focus on multiple trend breaks, we will confine our subsequent discussion to the CKP procedure.
fractions at a sufficiently fast rate. These asymptotic results are, however, somewhat at odds with the finite sample simulations reported in CKP. These show the presence of pronounced “valleys” in the finite sample power functions (when mapped as functions of the break magnitudes) of the tests such that power is initially high for very small breaks, then decreases as the break magnitudes increase, before increasing again. This discrepancy occurs because while the trend break pre-tests used in the CKP procedure are consistent against breaks of fixed magnitude, in finite samples they will not provide perfect discrimination; i.e., some degree of uncertainty will necessarily exist in finite samples as to whether breaks are present or not. The simulation results in CKP suggest, perhaps unsurprisingly, that this problem becomes increasingly pronounced the greater the number of breaks in the series, other things being equal.

For the case of a single break in trend, Harvey et al. (2011a) [HLT] show that treating the trend break magnitude to be local to zero (in a Pitman drift sense), rather than fixed, allows the (local) asymptotic distribution theory to very closely approximate this finite sample effect in the CKP test for a single break. This is because the local-to-zero model for the breaks reflects in the asymptotic theory the uncertainty that necessarily exists in finite samples as to whether trend breaks are present in the data or not. Here we show that the finite sample “valleys” problem worsens as the number of trend breaks present increases, other things being equal. Moreover, we show that in the case of multiple trend breaks the pattern of the breaks is also an important factor, in particular whether the parameters on consecutive trend breaks have equal or opposite sign. Our results suggest that the typical trend break magnitudes seen with real macroeconomic data lie well within the valleys region, suggesting that existing methods may be very poor at discriminating between the unit root null and stochastic stationary alternative in practice.

In response to this problem we propose a practical solution based on a similar approach to that outlined in ZA for the case of a single putative trend break. ZA propose using the infimum of $t$-ratio-type OLS detrended ADF statistics taken across all candidate break points in a trimmed range. However, it is known that the resulting test can have an asymptotic size of one when a trend break of fixed magnitude occurs under the unit root null; see Vogelsang and Perron (1998) and Harvey et al. (2011b). Correspondingly, we show that under a local-to-zero trend break the asymptotic size of the OLS-based infimum test can also exceed the nominal level, approaching unity in some cases. This renders the OLS detrended infimum test too unreliable to be recommended for use in practice. However, for the case of a single possible break in trend, a local GLS detrended implementation of this test is suggested in PR (who also suggest analogous extremum-type tests based on the $M$ and feasible likelihood ratio type tests), although they establish its large sample behaviour only for the case where no trend break occurs. We extend the work of PR by showing that the size problems discussed above for the OLS detrended infimum ADF test do not pertain under local GLS detrending. We further generalise the contribution of PR by developing a local GLS detrended infimum test which allows for multiple possible breaks in trend.

In both the single and multiple trend break cases, we show that these local GLS detrended infimum tests eliminate the aforementioned power valleys, although this necessarily comes at the expense of
some loss of power relative to the CKP test when no breaks are present. In a local-to-zero trend break environment and where the putative break fractions are unknown it is not possible to obtain unit root tests which are invariant (even asymptotically) to the break magnitudes, since the unknown break fractions cannot be consistently estimated. The results presented in this paper show that both the size and power properties of the infimum tests vary very little as a function of the break magnitudes, so that these tests are, in effect, “near”-invariant to the break magnitudes. The infimum test also has the considerable practical advantage that it is easy to compute both for a single and multiple putative trend breaks. In contrast, the alternative procedure outlined in CKP, based around the use of trend break pre-tests, is computationally burdensome even for the case of one or two possible trend breaks.

The plan of the paper is as follows. In section 2 we outline our reference multiple (local-to-zero) trend break model. Here we also detail our proposed infimum test (based based on local GLS detrended ADF tests) which allows for multiple possible breaks in trend. Section 3 details the large sample distributions of the infimum statistic under local-to-zero trend breaks and for a local-to-unity autoregressive root; asymptotic critical values are given for implementing the infimum test when allowing for a maximum of either one, two or three trend breaks. For the case of a single putative trend break, we also compare the asymptotic local power properties of the CKP test and the infimum-based tests based on OLS detrending and local GLS detrending. We also demonstrate that the infimum test, when based on local GLS detrended data, does not suffer the problem seen with its OLS counterpart, whose asymptotic size can tend to unity in the presence of a trend break. Section 4 investigates the finite sample behaviour of the CKP and local GLS detrended infimum tests. For the single possible break case the limit representations are shown to all closely predict the finite sample power functions of the tests. In the case of a DGP with two breaks in trend, the valleys problem with the CKP test is seen to become more severe than in the single break case and also to strongly depend on the pattern of the trend breaks present. An empirical illustration using data on primary commodity prices is also provided to highlight the potential usefulness of the proposed infimum tests. Section 6 concludes. Proofs are collected in an Appendix.

In the following ‘\(\cdot\)’ denotes the integer part of its argument, ‘\(\Rightarrow\)’ and ‘\(\mathbb{P}\)’ denote weak convergence and convergence in probability, respectively, ‘\(x := y\)’ (‘\(x =: y\)’) indicates that \(x\) is defined by \(y\) (\(y\) is defined by \(x\)), \(\circ\) denotes the Hadamard product, and ‘\(1(\cdot)\)’ denotes the indicator function. Finally, \(I_{x} := 1(x \neq 0)\) and \(I_{y}^{x} := 1(y > x)\).

2 The Model and Test Statistic

We consider a time series \(\{y_t\}\) to be generated according to the following DGP,

\[
y_t = \mu + \beta t + \gamma \text{DT}_t(\tau_0) + u_t, \quad t = 1, \ldots, T \\
u_t = \rho u_{t-1} + \varepsilon_t, \quad t = 2, \ldots, T
\] (2.1)

where, \(\text{DT}_t(\tau_0) := [\text{DT}_t(\tau_{0,1}), \ldots, \text{DT}_t(\tau_{0,m})]'\), the elements of which, for a generic fraction \(\tau\), are the indicator variables, \(\text{DT}_t(\tau) := 1(t > \lfloor \tau T \rfloor)(t - \lfloor \tau T \rfloor)\). In this model \(\tau_0 := [\tau_{0,1}, \ldots, \tau_{0,m}]'\) is the vector
of (unknown) putative trend break fractions, with $\gamma := (\gamma_1, \ldots, \gamma_m)'$ the associated break magnitude parameters; a trend break therefore occurs in $\{y_t\}$ at time $\lceil \tau_{0,i} T \rceil$ when $\gamma_i \neq 0$, $i = 1, \ldots, m$. The break fractions are assumed to be such that $\tau_{0,i} \in \Lambda$, for all $i$, where $\Lambda := [\tau_L, \tau_U]$ with $0 < \tau_L < \tau_U < 1$; the fractions $\tau_L$ and $\tau_U$ representing trimming parameters. It is also assumed that $|\tau_{0,i} - \tau_{0,j}| \geq \eta > 0$, for all $i, j$, $i \neq j$, such that the DGP admits (up to) $m$ level breaks occurring at unknown points across the interval $\Lambda$, with a sample fraction of at least $\lceil \eta T \rceil$ observations between breaks. Notice, therefore, that $m$ and $\eta$ must satisfy the relation $m \leq 1 + \lceil (\tau_U - \tau_L)/\eta \rceil$.

In (2.2), $\{u_t\}$ is an unobserved mean zero stochastic process, initialised such that $u_1 = o_p(T^{1/2})$. The disturbance term, $\varepsilon_t$, is taken to satisfy the following conventional stable and invertible linear process-type assumption:

**Assumption 1** Let $\varepsilon_t = C(L) \nu_t$, $C(L) := \sum_{i=0}^{\infty} C_i L^i$, $C_0 := 1$, with $C(z) \neq 0$ for all $|z| \leq 1$ and $\sum_{i=0}^{\infty} i|C_i| < \infty$, and where $\nu_t$ is an independent and identically distributed (IID) sequence with mean zero, variance $\sigma^2$ and finite fourth moment. We also define the short-run and long-run variances of $\varepsilon_t$ as $\sigma_2^2 := E(\varepsilon_t^2)$ and $\omega_2^2 := \lim_{T \to \infty} T^{-1} E(\sum_{t=1}^{T} \varepsilon_t)^2 = \sigma^2 C(1)^2$, respectively.

Our interest in this paper centres on testing the unit root null hypothesis $H_0 : \rho_T = 1$, against the local alternative, $H_e : \rho_T = 1 - c/T$, $c > 0$, without assuming knowledge of whether trend breaks are present in the data or not. The test we propose is a multiple break version of the minimum GLS detrended Dickey-Fuller statistic proposed by PR (which is $MDF_1$ in the notation below), following the approach taken by ZA in a single break OLS detrending environment. Specifically, the test statistic is

$$MDF_m := \inf_{\tau_1, \ldots, \tau_m \in \Lambda, \ |\tau_i - \tau_j| \geq \eta, \ i \neq j} DF^G_{\hat{c}} (\tau)$$

where $DF^G_{\hat{c}} (\tau)$ denotes the standard $t$-ratio associated with $\hat{\pi}$ in the fitted ADF-type regression

$$\Delta \hat{u}_{t} = \hat{\pi} \hat{u}_{t-1} + P \sum_{j=1}^{k} \hat{\psi}_{j} \Delta \hat{u}_{t-j} + \epsilon_{t,k}, \quad t = k + 2, \ldots, T, \quad (2.3)$$

where $\hat{u}_{t} := y_t - \hat{\mu} - \hat{\beta} t - \hat{\gamma}/DT_{t}(\tau)$, with $[\hat{\mu}, \hat{\beta}, \hat{\gamma}]'$ obtained from a local GLS regression of $y_{\hat{\rho}} := [y_1, y_2 - \hat{\rho} y_1, \ldots, y_{T} - \hat{\rho} y_{T-1}]'$ on $Z_{\hat{\rho}, \tau} := [z_1, z_2 - \hat{\rho} z_1, \ldots, z_T - \hat{\rho} z_{T-1}]'$, $z_t := [1, t, DT_{t}(\tau)]'$ with $\hat{\rho} := 1 - \hat{\epsilon}/T$, for some $\hat{\epsilon} > 0$ which is user-supplied. This infimum unit root test rejects for large negative values of the statistic. As discussed in PR, it is assumed that the lag truncation parameter, $k$, is 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), starting from a maximum lag truncation, $k_{\text{max}}$, which satisfies the usual condition that $1/k_{\text{max}} + k_{\text{max}}^3/T \to 0$ as $T \to \infty$.

**Remark 1.** It should be noted that the unit root test based on $MDF_m$ will require the practitioner to specify a value for $m$, the maximum number of breaks allowed to occur in the sample. In practice

\[3\] We suppress the dependence of quantities such as $\hat{u}_{t}$ on $\tau$ for notational economy.
a trade-off will arise between setting \( m \) too low and running the risk of neglecting trend breaks on the one hand or, on the other hand, setting \( m \) too high which will necessarily compromise the finite sample power of the test. A value for \( m \) also needs to be chosen in the context of the (multiple) trend break pre-test of Kejriwal and Perron (2010) used in the CKP procedure; see the discussion on pages 1776 and 1779 of CKP. For the same reason as just given it is also important not to set \( m \) too small here. Although it might be tempting to think one could set \( m \) to be reasonably large here, in practice the Kejriwal and Perron (2010) pre-test does not appear to perform well in finite samples in such cases; indeed, in discussing their finite sample simulation results, Kejriwal and Perron (2010, p.320) argue that “The simulation results point to the importance of the choice of the maximal value of the number of breaks in relation to the size of the sample available. For example, when testing for two breaks in a sample of size 120, one ends up with fewer than 40 observations per segment. It is then not surprising to see low power and/or size distortions. Hence, practitioners must exercise caution to allow a sufficient number of observations in each segment and choose the maximum number of breaks permissible accordingly.”

3 Asymptotic Behaviour of \( MDF_m \)

In order to conduct an asymptotic analysis that appropriately mimics the relevant finite sample power properties of unit root tests when uncertainty exists as to the presence of breaks, we will primarily focus attention on asymptotic behaviour in a doubly-local setting; here, consistent with HLT’s analysis, in addition to allowing local-to-unity behaviour in the autoregressive root, as above, we also assume that the trend break magnitudes are local-to-zero, in contrast to the fixed magnitude assumption in CKP. Specifically, we set the break magnitudes in (2.1) to be \( \gamma_{i,T} = \kappa_i \omega_i T^{-1/2}, i = 1, \ldots, m \), where the \( \kappa_i \) are finite constants, thereby adopting the appropriate Pitman drift for a trend break in a local-to-unit root process.\(^3\) For completeness, however, we will also subsequently discuss the behaviour of our proposed local GLS detrended infimum ADF test under a fixed magnitude break assumption, in order to show that this test does not suffer the drawbacks seen with the corresponding test based on OLS detrending.

3.1 Asymptotic Behaviour Under Local Trend Breaks

In this section we first state the large sample behaviour of the \( MDF_m \) statistic under local-to-unity behaviour in the autoregressive root and when the trend break magnitudes are local-to-zero; i.e., \( \gamma = \gamma T = \kappa \omega T^{-1/2} \), with \( \kappa \) a vector of finite constants. This result is given in Theorem 1. Based on these limiting representations, we then use numerical methods to obtain asymptotic critical values (section 3.2), and explore the asymptotic size (section 3.3) and local power (section 3.4) of the test based on the \( MDF_m \) statistic.

\(^3\)Scaling the trend break by \( \omega_i \) is merely a convenience device allowing it to be factored out of the limit distributions that arise later.
Theorem 1. Let $y_{t}$ be generated according to (2.1) and (2.2) under Assumption 1. Let $H_{c} : \rho_{T} = 1 - c/T$, $c \geq 0$, hold and let $\tau_{T} = \kappa \omega \varepsilon T^{-1/2}$. Then

$$MDF_{m} \Rightarrow \inf_{\tau_{1}, \ldots, \tau_{m} \in \Lambda, \tau_{i} - \tau_{j} \geq \eta, \forall i \neq j} D_{c,\varepsilon}(\tau_{0}, \tau, \kappa)$$

(3.1)

where

$$D_{c,\varepsilon}(\tau_{0}, \tau, \kappa) := \frac{L_{c,\varepsilon}(1, \tau_{0}, \tau, \kappa)^{2} - 1}{2 \sqrt{\int_{0}^{1} L_{c,\varepsilon}(r, \tau_{0}, \tau, \kappa)^{2} dr}}$$

with

$$L_{c,\varepsilon}(r, \tau_{0}, \tau, \kappa) := W_{c}(r) + \kappa' \{(r - \tau_{0}) \circ I_{r}^{\varepsilon}\}$$

$$- \left[ r - (r - \tau) \circ I_{r}^{\varepsilon} \right]^{'} \left[ a_{\varepsilon} m_{\varepsilon}(\tau)^{'} m_{\varepsilon}(\tau) \right]^{-1} \left[ b_{c,\varepsilon} + \kappa f_{c,\varepsilon}(\tau_{0}) \right]^{'}$$

$$b_{c,\varepsilon}(\tau_{0}, \tau)$$

(3.2)

where

$$a_{\varepsilon} := 1 + \bar{c} + \bar{c}^{2}/3, \quad b_{c,\varepsilon} := (1 + \bar{c})W_{c}(1) + \bar{c}^{2} \int_{0}^{1} sW_{c}(s) ds,$$

$I_{r}^{\varepsilon}$ denotes an $m \times 1$ vector with $i$th element $\mathbb{I}_{r_{0}, i}$, $m_{\varepsilon}(\tau)$, $b_{c,\varepsilon}(\tau)$ and $f_{c,\varepsilon}(\tau_{0})$ denote $m \times 1$ vectors with $i$th elements

$$m_{\varepsilon}(\tau_{i}) := a_{\varepsilon} - \tau_{i}(1 + \bar{c} + \bar{c}^{2}/2 - \bar{c}^{2}\tau_{i}^{2}/6),$$

$$b_{c,\varepsilon}(\tau_{i}) := (1 + \bar{c} - \bar{c}^{2})W_{c}(1) - W_{c}(\tau_{i}) + \bar{c}^{2} \int_{\tau_{i}}^{1} (s - \tau_{i}) W_{c}(s) ds,$$

$$f_{c,\varepsilon}(\tau_{0}, i) := (1 - \tau_{0,i})\{a_{\varepsilon} - \bar{c}^{2}\tau_{0,i}(1 + \tau_{0,i})/6\}$$

respectively, $D_{c}(\tau)$ and $F_{c,\varepsilon}(\tau_{0}, \tau)$ denote $m \times m$ matrices with $i$, $j$th elements

$$d_{\varepsilon}(\tau_{i}, \tau_{j}) := (1 - \tau_{\max})\{a_{\varepsilon} + \bar{c}\tau_{\max} - \bar{c}(\tau_{i} + \tau_{j})$$

$$+ \bar{c}^{2}(\tau_{\max}^{2} + \tau_{\max}^{2})/3 - \bar{c}^{2}(\tau_{i} + \tau_{j})(1 + \tau_{\max})/2 + \tau_{i}\tau_{j}\},$$

$$f_{c,\varepsilon}(\tau_{0,j}, \tau_{i}) := (1 - \tau_{0,j})\{a_{\varepsilon} - \bar{c}\tau_{i} - \bar{c}^{2}\tau_{i}(1 - \tau_{0,j})/2 - \bar{c}^{2}\tau_{0,j}(1 + \tau_{0,j})/6\}$$

$$-(\tau_{i} - \tau_{0,j})\{1 - \bar{c}^{2}(\tau_{i} - \tau_{0,j})^{2}/6\}\mathbb{I}_{\tau_{0}, j},$$

respectively, with $\tau_{\max} := \max(\tau_{i}, \tau_{j})$, and where $W_{c}(r) := \int_{0}^{r} e^{-(r-s)c} dW(s)$, with $W(r)$ a standard Brownian motion process.

Remark 2. Notice that when $\kappa = 0$, $D_{c,\varepsilon}(\tau_{0}, \tau, 0)$, and therefore the limit distribution of $MDF_{m}$, does not depend on $\tau_{0}$, as should be expected given that no break occurs.

Remark 3. When $m = 1$ and $\kappa = 0$, we note that the limit $D_{c,\varepsilon}(\tau_{0}, \tau, 0)$ coincides with that contained in Theorem 1 of PR. Consequently, for the case where at most a single break is permitted, but no break is actually present in the DGP, the representation given for $MDF_{1}$ in Theorem 1 simplifies to the expression given in equation (11) of PR.
3.2 Asymptotic Critical Values

In what follows we will set \( L = 0.15 \) and \( U = 0.85 \), and also set \( \eta = 0.15 \), which would seem to be a sensible choice for the break separation. The asymptotic behaviour of \( MDF_m \) under the unit root null hypothesis is obtained on setting \( c = 0 \) in Theorem 1. This limit distribution obviously remains a function of \( \tau_0 \) and \( \kappa \), but also of the user-supplied parameter \( \bar{c} \). First, to make the statistic operational we need to select a value of \( \bar{c} \). Consider first the case where at most a single break in trend is considered. Table 1 of Harris et al. (2009) reports, for each point on the grid of break fractions \( \tau_{0,1} \in \{0.15, 0.20, \ldots, 0.85\} \), the values of \( c = c_{\tau_{0,1}} \) for which the Gaussian power envelope lies at 0.50. We could base \( MDF_1 \) on the minimum of \( DF_{GLS}^{c_{\tau_1}}(\tau_1) \) across \( \tau_1 \), setting \( c_{\tau_1} \) to the corresponding value of \( c_{\tau_{0,1}} \) (and interpolating for intermediate \( \tau_1 \) values). However, since the single break values of \( c_{\tau_{0,1}} \) vary very little across \( \tau_{0,1} \), we adopt the more convenient practice of simply setting \( c \) in \( MDF_1 \) to be the average of the single break \( c_{\tau_{0,1}} \) values in what follows. At the nominal 0.05 level, this average value is \( \bar{c} = 17.6 \).\(^4\) When allowing for multiple trend breaks, we propose adopting the same method for selecting \( \bar{c} \): first we simulate, for each combination of break fractions \( \tau_{0,i} \in \{0.15, 0.20, \ldots, 0.85\} \) (subject to the \( \eta \) separation requirement), the values of \( c = c_{\tau_0} \) for which the nominal 0.05 level Gaussian power envelope lies at 0.50; \( \bar{c} \) is then evaluated as the average of the \( c_{\tau_0} \) values. The resulting \( \bar{c} \) values for \( m = 2 \) and \( m = 3 \) are reported in Table 1. Here and throughout the paper, we approximated the Wiener processes in the limiting functionals using \( N(0,1) \) random variates, and, unless stated otherwise, the integrals are approximated by normalized sums of 2,000 steps and 50,000 replications are used.

Using this designation for \( \bar{c} \), we next obtain asymptotic null critical values for \( MDF_m \). As noted above, its distribution depends on \( \tau_0 \) and \( \kappa \). We simulated (3.1), with \( c = 0 \), setting \( \kappa = 0 \) such that

\[
L_{0,\bar{c}}(r, \tau_0, \tau, 0) := W(r) - \begin{bmatrix} r & \tau \end{bmatrix}' \begin{bmatrix} a_{\bar{c}} & m_{\bar{c}}(\tau)' \\ m_{\bar{c}}(\tau) & D_{\bar{c}}(\tau) \end{bmatrix}^{-1} \begin{bmatrix} b_{0,\bar{c}} \\ b_{0,\bar{c}}(\tau) \end{bmatrix}
\]

These critical values are given in Table 1, again for \( m = 1 \), \( m = 2 \) and \( m = 3 \).\(^5\) We regard allowing for (up to) \( m = 3 \) trend breaks as probably sufficient for the great majority of empirical applications. Moreover, allowing for more than three breaks inevitably compromises the power of the resulting unit root test; see again the discussion in Remark 1. Allowing for many breaks, given a value of the separation parameter \( \eta \), also places constraints on where the breaks may be located, and has the potential to prevent the correct location of a smaller number of true trend breaks. Practical limitations also apply to the maximum number of allowable breaks used in the CKP procedure; see Remark 1.

\(^4\)It is also convenient to use this same value for other significance levels, which ensures that the test statistic does not change its value when implemented at different significance levels; this simplification has little impact on the behaviour of the test since power is relatively insensitive to the choice of \( \bar{c} \).

\(^5\)The critical values for \( m = 2 \) were obtained using 10,000 replications; those for \( m = 3 \) used 5,000 replications, with the limit integrals now approximated by normalized sums of 1,000 steps.
3.3 Asymptotic Size Under a Local Trend Break

We now consider the asymptotic sizes of $MDF_m$ under a local break in trend, focusing on $MDF_1$ and a single break in trend ($m = 1$) in (2.1) due to the computational burden involved in simulating the limit of $MDF_m$ when $m > 1$. In Figure 1 we show the sizes for local trend breaks of magnitude $\kappa_1 = \{0, 0.2, \ldots, 30\}$ and break fractions $\tau_{0,1} = \{0.3, 0.5, 0.7\}$ at the nominal 0.05 level; for purposes of comparison, we also show the asymptotic sizes of the procedure proposed by CKP with up to one break permitted (denoted $CKP_1$), along with those for the corresponding test based on OLS detrended data (i.e. an additive outlier version of ZA’s procedure, denoted $ZA_{AO}$), evaluated with the same allowable break fraction settings as in $MDF_1$.

While we defer from presenting here the details on how $CKP_1$ is constructed (these can be found on page 1776 of CKP), essentially the procedure uses a pre-test (the trend break test of Perron and Yabu, 2009) to detect the presence or otherwise of a trend break occurring at an unknown point, and then uses the outcome of this detection step to indicate whether or not a (likelihood ratio-based) unit root test should include a trend break in the deterministic specification. When no trend break is present, or a fixed magnitude trend break exists, the $CKP_1$ procedure is asymptotically efficient.

For $ZA_{AO}$, it is straightforward to show that its limit distribution under $c = 0$ is given by

$$ZA_{AO} \Rightarrow \inf_{\tau_1 \in A} Z(\tau_{0,1}, \tau_1, \kappa_1)$$

where

$$Z(\tau_{0,1}, \tau_1, \kappa_1) := \frac{N(1, \tau_{0,1}, \tau_1, \kappa_1)^2 - N(0, \tau_{0,1}, \tau_1, \kappa_1)^2}{2\sqrt{\int_0^1 N(r, \tau_{0,1}, \tau_1, \kappa_1)^2dr}}$$

with $N(r, \tau_{0,1}, \tau_1, \kappa_1)$ the continuous time residual process from the projection of $W(r) + \kappa_1 (r - \tau_{0,1})I_{\tau_{0,1}}$ onto the space spanned by $\{1, r, (r - \tau_1)I_r\}$. The limit distribution we use for $CKP_1$ is that given in HLT with the settings employed therein; see HLT for full details.

Broadly speaking, and other things being equal, the size of $MDF_1$ decreases as $\kappa_1$ increases; however, even for the largest value of $\kappa_1$ we consider, the under-sizing is really quite modest. To all intents and purposes then, we can consider that the limit critical values of $MDF_1$ are at their most negative when $\kappa_1 = 0$, and henceforth we will consider $MDF_m$ with use of asymptotic critical values obtained with $\kappa = 0$ as essentially being robust to local trend breaks. The $CKP_1$ test is also seen to have asymptotic size that is close to the nominal 0.05 level; while some over-size is observed for moderate magnitudes of $\kappa_1$, the size never exceeds 0.094, and is correct both in the case of no break and in the presence of large local breaks. In contrast, the size of $ZA_{AO}$ can be very severely distorted, with asymptotic size approaching one as $\kappa_1$ increases when $\tau_{0,1} = 0.3$. This is entirely to be expected, given that for a break of fixed magnitude, Harvey et al. (2011b) show that $ZA_{AO}$ has an asymptotic size of unity whenever $\tau_{0,1} < \frac{2}{3}$. Although this over-size is not apparent in Figure 1 for $\tau_{0,1} = 0.5$, unreported simulations confirm that size does approach one for larger values of $\kappa_1$ (e.g. asymptotic size is 0.210 when $\kappa_1 = 70$, 0.555 when $\kappa_1 = 80$, and 0.877 when $\kappa_1 = 90$). Vogelsang and Perron (1998) note, however, that in finite samples large upward size distortions are only seen for $ZA_{AO}$ for
relatively large breaks, a feature borne out in our asymptotic findings where $ZA_{AO}$ appears to be approximately correctly sized for small and moderate local break magnitudes, even when $\tau_{0,1} = 0.3$. Its behaviour for large local (or fixed) breaks when $\tau_{0,1} < \frac{2}{3}$ does, however, preclude it from being entertained as a reliable testing procedure in general, and we consequently consider it no further in the remainder of this paper.

3.4 Asymptotic Power Under a Local Trend Break

Again in the case of a single break in trend, Figure 2 graphs the asymptotic local power function of the nominal 0.05 level $MDF_1$ test across the range of local break magnitudes $\kappa_1 = \{0, 0.2, \ldots, 15\}$ (as in HLT) for the break fractions $\tau_0 = \{0.3, 0.5, 0.7\}$ and for the two alternatives $c = 20$ and $c = 30$. Also reported for comparison are the asymptotic power functions for $CKP_1$, as obtained by HLT; these powers are size-adjusted by scaling the with-break and without-break unit root test critical values by a common factor, such that the size of the overall procedure is 0.05 when $\kappa_1$ is set to the value where the corresponding uncorrected size was found to be at a maximum, this scaling then being used for all other $\kappa_1$ values.

The standout feature of the power profiles of $MDF_1$ is how little affected they are by the value of $\kappa_1$. The same is patently untrue of $CKP_1$ which, as discussed in HLT, displays very distinct power “valleys” that arise because for a range of local break magnitudes, the break is large enough to seriously compromise the power of the without-break unit root test, while simultaneously being too small to be reliably identified by the Perron and Yabu (2009) trend break pre-test implicit in $CKP_1$. For small $\kappa_1$, $CKP_1$ clearly dominates $MDF_1$ in terms of power, due to the fact that here the former is applying a without-break unit root test which has considerable power advantages over the with-break version employed by $MDF_1$. However, for moderate local break magnitudes, the rankings are reversed and while $MDF_1$ retains an almost constant level of power, the power of $CKP_1$ falls quite dramatically, demonstrating the aforementioned “valley” phenomenon. For larger breaks, the pre-test in $CKP_1$ correctly identifies a break in trend with high probability, and consequently the power of $CKP_1$ recovers, but of course only to those levels associated with a with-break unit root test; in this region, $MDF_1$ and $CKP_1$ have not dissimilar levels of power.

3.5 Asymptotic Behaviour Under a Fixed Trend Break

In this section we investigate the effect of a single ($m = 1$) non-local (fixed magnitude) break on $MDF_1$; that is, for the purposes of this section we set $\gamma_1 = \kappa_1 \omega$, with $\kappa_1 \neq 0$ in (2.1). For purposes of tractability in this section we will consider the case where the disturbance term, $\varepsilon_t$, in (2.2) is an IID process (so that $\omega = \sigma$), with $k$ set to zero in (2.1). Allowance for more than one break and serially correlated innovations is discussed in Remark 6 below.
**Theorem 2** Let \( y_t \) be generated according to (2.1) and (2.2) under the assumptions that \( m = 1 \) and \( \varepsilon_t \sim IID(0, \sigma^2_\varepsilon) \) with finite fourth order moment. Then, for \( \gamma_1 = \kappa_1 \sigma_\varepsilon \), \( \kappa_1 \neq 0 \),

\[
MDF_1 \Rightarrow \inf_{\tau_1 \in \Lambda'} \mathcal{D}'_{c,e}(\tau_{0,1}, \pi'_{\tau_1 - \tau_{0,1}, \kappa_1})
\]

(3.3)

where \( \Lambda' \) is the set of \( \tau_1 \) satisfying \( \pi'_{\tau_1 - \tau_{0,1}} := \lim T^{1/2} (\tau_1 - \tau_{0,1}) \) and \( 0 \leq |\pi'_{\tau_1 - \tau_{0,1}}| < \infty \), and where

\[
\mathcal{D}'_{c,e}(\tau_{0,1}, \pi'_{\tau_1 - \tau_{0,1}, \kappa_1}) := \frac{J_{c,e}(1, \tau_{0,1}, \pi'_{\tau_1 - \tau_{0,1}, \kappa_1})^2 - 1}{2 \int_0^1 J_{c,e}(r, \tau_{0,1}, \pi'_{\tau_1 - \tau_{0,1}, \kappa_1})^2 dr}
\]

with

\[
J_{c,e}(r, \tau_{0,1}, \pi'_{\tau_1 - \tau_{0,1}, \kappa_1}) := L_{c,e}(r, \tau_{0,1}, \tau_{0,1}, 0) + \kappa_1 \pi'_{\tau_1 - \tau_{0,1}} \mathbb{I}_{\tau_{0,1}, 0} \frac{r}{(r - \tau_{0,1}) \mathbb{I}_{\tau_{0,1}, 0}} \left[ \begin{array}{ccc} \alpha_{\varepsilon} & m_{\varepsilon}(\tau_{0,1}) \\ m_{\varepsilon}(\tau_{0,1}) & d_{\varepsilon}(\tau_{0,1}, \tau_{0,1}) \end{array} \right]^{-1} \left[ \begin{array}{c} l_{1,\varepsilon}(\tau_{0,1}) \\ l_{2,\varepsilon}(\tau_{0,1}) \end{array} \right]
\]

where

\[
l_{1,\varepsilon}(\tau_{0,1}) := 1 + \tilde{c} \left( 1 + \tilde{c} \left( 1 - \frac{\tau_{0,1}^2}{2} \right) \right)
\]

\[
l_{2,\varepsilon}(\tau_{0,1}) := \left\{ \begin{array}{ll}
1 + \tilde{c} \left( 1 - \tau_{0,1} + \tilde{c} \left( 1 - \tau_{0,1} \right)^2 / 2 \right) & \pi'_{\tau_1 - \tau_{0,1}} \leq 0 \\
\tilde{c} \left( 1 - \tau_{0,1} + \tilde{c} \left( 1 - \tau_{0,1} \right)^2 / 2 \right) & \pi'_{\tau_1 - \tau_{0,1}} > 0
\end{array} \right.
\]

**Remark 4.** An implication of Theorem 2 is that the break fraction estimator implicit in \( MDF_1 \), i.e. \( \hat{\tau}_1 := \arg \min_{\tau_1 \in \Lambda} D_{c,e}^{GLS}(\tau_1) \), is such that \( \hat{\tau}_1 = \tau_{0,1} + O_p(T^{-1/2}) \), since it is clear that \( \hat{\tau}_1 \) lies in the region \( \Lambda' \) asymptotically, with \( \hat{\tau}_1 \) lying within a \( O_p(T^{-1/2}) \) neighbourhood of \( \tau_{0,1} \). However, we obtained the 0.05 level asymptotic critical values for \( MDF_1 \) under the fixed trend break assumption with \( \kappa_1 = \{1, 2\} \) and \( \tau_{0,1} = \{0.3, 0.5, 0.7\} \), simulating (3.3) using \( \Lambda' = \{\tau_{0,1} \pm 2T^{-1/2}\} \), and found the critical values to lie in the range \(-3.54\) to \(-3.68\); by contrast, the corresponding asymptotic 0.05 critical values for \( D_{c,e}^{GLS}(\tau_{0,1}) \) are in the range \(-3.33\) to \(-3.44\), which makes clear that the limiting null distribution of \( MDF_1 \) under a fixed break does not coincide with that which would be obtained if \( \tau_{0,1} \) were known. Since it is well established that a rate of consistency of a break fraction estimator that is faster than \( O_p(T^{-1/2}) \) results in the corresponding unit root statistic having a limit distribution coincident with that of the same statistic evaluated using a known break fraction, it follows that while \( \hat{\tau}_1 = \tau_{0,1} + O_p(T^{-1/2}) \), it cannot be true that \( \hat{\tau}_1 = \tau_{0,1} + o_p(T^{-1/2}) \).

**Remark 5.** The properties of \( MDF_1 \) shown in Theorem 2 contrast sharply with those of the OLS detrended analogue. As mentioned above, Harvey et al. (2011b) demonstrate that when a fixed break occurs under the null the behaviour of the OLS detrended variant depends on the true break fraction, \( \tau_{0,1} \); specifically, at all but the true break fraction, \( \tau_{0,1} \), the ADF-type statistics in the sequence over which the infimum is taken will diverge in \( T \) to either \(-\infty \) if \( \tau_{0,1} < 2/3 \), or \(+\infty \) if \( \tau_{0,1} \geq 2/3 \).

\(^a\)The range \( \Lambda' \) was selected on the basis of a prior calibrating simulation, which suggested that \( \inf_{\tau_1 \in \Lambda'} \) almost always lay within the range \( \tau_{0,1} \pm 2T^{-1/2} \).
Consequently, the OLS detrended infimum test will have an asymptotic size of one, along with an inconsistent implicit break fraction estimator, if $\tau_{0,1} < 2/3$, although the statistic will be of $O_p(1)$ and the implicit break fraction estimator will be consistent if $\tau_{0,1} \geq 2/3$. Consequently, without certain knowledge that $\tau_{0,1} \geq 2/3$, one could not entertain the use of the OLS detrended variant in practice. PR also consider $MDF_1$, but presume that the problems with the OLS detrended variant would carry over to the GLS detrended implementation; in fact these problems do not persist in the GLS context, as Theorem 2 demonstrates.

**Remark 6.** We conjecture that in the case of multiple breaks in trend, a similar analysis could be used to establish that the asymptotic behaviour of $MDF_m$ is qualitatively similar to that of $MDF_1$. Additionally, we conjecture that the IID assumption for the disturbance process $\{\varepsilon_t\}$ underlying Theorem 2 is not critical for the asymptotic size robustness of $MDF_m$ in the presence of fixed magnitude trend breaks, provided $k$ in (2.1) is chosen appropriately, as in section 2. While establishing these results analytically would be extremely involved, unreported finite sample simulation results provided evidence in their support.

We conclude this section by investigating the asymptotic size properties of $MDF_1$ under fixed trend breaks of magnitude $\kappa_1 = \{1, 2\}$ and break fractions $\tau_{0,1} = \{0.3, 0.5, 0.7\}$ at the nominal 0.05 level, using the critical values from Table 1. These results, reported in Table 2, were again obtained from simulation of (3.3) using $\Lambda' = \{\tau_{0,1} \pm 2T^{-1/2}\}$. A modest amount of under-sizing is present, indicating that the limiting null distribution of $MDF_1$ is right-shifted (at least in the region of the lower 0.05 quantile) under a fixed break in trend, as was also observed under a local break in trend. The degree of under-sizing is also closely akin to that observed for the larger values of the local trend break $\kappa_1$ in Figure 1. It would appear, therefore, that our local trend break size results for the larger $\kappa_1$ also provide a very good predictor of size behaviour under a fixed trend break.

Table 2 also shows the corresponding powers of $MDF_1$ when $c = 20$ and $c = 30$. Again we observe these are very much in line with those from Figure 2, for the larger values of the local trend break $\kappa_1$. The results in Table 2 also highlight the fact that the size and power of $MDF_1$ under a fixed trend break are quite insensitive to the break magnitude.

## 4 Finite Sample Results

In this section we present simulation results on the power of the tests in finite samples, using $T = 150$ and $T = 300$. For these sample sizes, the computational requirement for simulating the power curves of the $MDF_m$ and $CKP$ tests when allowing for multiple breaks is now considerably less burdensome, thus here we provide results both for the single break case ($m = 1$) and for the two break case ($m = 2$). We also illustrate the potential value of the $MDF_m$ approach using a selection of relative primary commodity price series.
4.1 The Single Break Case

Figures 3 and 4 report the finite sample power functions of the \( MDF_1 \) and \( CKP_1 \) tests under the DGP (2.1) and (2.2) with \( m = 1 \). We generated \( \varepsilon_t \sim NIID(0,1) \) (so that \( \omega_e = 1 \)), set \( k = 0 \) in (2.3), and employed parameter settings that correspond to the local asymptotic power simulations of section 3.4, i.e. \( \tau_{0,1} = \{0.3, 0.5, 0.7\} \), \( \gamma_1 = \kappa_1 T^{-1/2} \) with \( \kappa_1 = \{0, 0.2, ..., 15\} \), and \( \rho_T = 1 - c/T \) with \( c = 20 \) and \( c = 30 \). We set \( \mu = \beta = 0 \) in (2.1), without loss of generality. The tests were conducted at the nominal 0.05 level using asymptotic critical values, with the powers being size-adjusted in the same way as described in section 3.4. The simulations here were based on 5,000 Monte Carlo replications.

The relative shapes of the finite sample power curves very closely mirror those for the corresponding limits reported in Figure 2, particularly for \( T = 300 \); indeed, the local asymptotic analysis appears to provide an excellent predictor of the finite sample behaviour of these tests. The same remarks therefore apply as in section 3.4, with \( MDF_1 \) once again having an appealing power function that displays robustness to the magnitude of the break in trend, unlike \( CKP_1 \) which, despite achieving high power for zero/small breaks and reliable power for very large breaks, exhibits considerably lower power for intermediate trend break magnitudes. It is reassuring to see that the robust behaviour observed for \( MDF_1 \) in the limit carries over to small samples, adding to the appeal of adopting such an approach in practical applications.

4.2 The Two Break Case

We next examine finite sample power in the case of two breaks in trend, i.e. \( m = 2 \) in (2.1), with the break fractions \( \tau_{0,1} = 0.33 \) and \( \tau_{0,2} = 0.66 \), and break magnitudes \( \gamma_i = \kappa_i T^{-1/2} \), \( i = 1, 2 \) for the two cases (i) \( \kappa_1 = \kappa_2 = \kappa \), and (ii) \( \kappa_1 = -\kappa, \kappa_2 = \kappa \), with \( \kappa = \{0, 0.2, ..., 15\} \). As before we consider \( T = \{150, 300\} \) with (2.2) generated using \( \rho_T = 1 - c/T \) with \( c = 20 \) and \( c = 30 \), and \( \varepsilon_t \sim NIID(0,1) \).

Figures 5 and 6 report the powers of \( MDF_2 \) tests conducted at the nominal 0.05 level, again with \( k = 0 \) in (2.3) and using asymptotic critical values. Also reported are the powers of the corresponding CKP test. In the multiple break setting, the first stage of the CKP procedure involves pre-testing to identify the number of breaks present, using sequential application of the single break detection test of Perron and Yabu (2009) test, as proposed by Kejriwal and Perron (2010); for further details, see pages 1776 and 1779 of CKP. We permit here a maximum of two breaks, denoting the resulting test by \( CKP_2 \). Where the tests display over-size across \( \kappa \), the powers are size-adjusted in the same way as before. The simulations here were based on 1,000 Monte Carlo replications.

Once more we see that the power profiles of \( MDF_2 \) are remarkably flat with respect to the break magnitudes \( \kappa_1 \) and \( \kappa_2 \). As would be expected, the powers are somewhat lower than in the one break case, but the key feature of insensitivity to the break magnitude(s) is retained. As regards \( CKP_2 \), while power gains are again achieved relative to \( MDF_2 \) for breaks of very small or zero magnitude, the power “valleys” are still very much in evidence. Indeed, they are considerably wider than in the one break case, and, apart from very small magnitudes, it is only for much larger break magnitudes that decent power is recovered. This feature is particularly marked in the case where the break coefficients
are opposite in sign (see Figures 5(c)-(d) and 6(c)-(d)); here the power of \( \text{CKP}_2 \) actually drops to zero over a wide range of values of \( \kappa \). The difference in power performance between the reinforcing breaks case (\( \kappa_1 = \kappa_2 \)) and the opposite signed breaks case (\( \kappa_1 = -\kappa_2 \)) comes about from the behaviour of the implicit auxiliary methodology for detecting the number of trend breaks. When the break coefficients are opposite in sign, the inherent mis-specification in the first stage of testing zero versus one break results in very low detection rates unless the breaks are very large. This still happens, albeit to a lesser extent, when the breaks are reinforcing. Hence while \( \text{CKP}_2 \) can clearly exhibit power advantages over \( \text{MDF}_2 \) for some break magnitudes, its simultaneous potential for zero power for other break magnitudes seriously questions its suitability for application in practice, particularly in comparison to the more robust \( \text{MDF}_2 \) procedure.

### 4.3 Empirical Illustration

By way of a demonstration that the potential power gains offered by \( \text{MDF}_m \) for intermediate break magnitudes are practically relevant and not merely a theoretical curiosity, we now apply the tests to four relative primary commodity price series. The series we consider are (the logarithms of) indices of primary commodity prices relative to the price of manufactures, observed annually over the period 1900–2003 (\( T = 104 \)), and are a subset of the 24 series dataset compiled by Pfaffenzeller et al. (2007), who updated the widely used dataset of Grilli and Yang (1988). The four commodities we consider are Copper, Hides, Lead and Silver, and are selected to illustrate the relevance of the simulation results, being cases where (multiple) breaks of intermediate magnitude appear to be present, and where \( \text{MDF}_m \) offers rejections of the unit root null but \( \text{CKP} \) does not.

Table 3 presents results for the \( \text{MDF}_1 \), \( \text{MDF}_2 \) and \( \text{CKP}_2 \) tests applied at the nominal asymptotic 0.05 significance level, with \( k \) selected using the MAIC procedure of Ng and Perron (2001) and Perron and Qu (2007) with \( k_{\text{max}} = \lceil 12(T/100)^{1/4} \rceil \). We find that the test which permits only a single break in trend (\( \text{MDF}_1 \)) rejects the unit root null only for Lead. However, when allowance is made for two breaks in trend, the \( \text{MDF}_2 \) test rejects in favour of stationarity about a broken trend path for all four series. In contrast, the CKP procedure that allows for up to two breaks (\( \text{CKP}_2 \)) delivers no rejections of the null across the four series; in each case the first stage Perron and Yabu (2009) pre-test fails to reject the null of no trend breaks, thus for each series \( \text{CKP}_2 \) amounts to applying a without-break unit root test.

To provide some insight into the pattern of rejections and non-rejections observed in Table 3, we estimated model (2.1) for each series assuming \( m = 2 \), and computed estimates of the implied local break magnitudes associated with our local-to-unity model. Specifically, we first estimated the timing of the trend breaks by minimizing the residual sum of squares across \( \tau_1, \tau_2 \in [0.15, 0.85], |\tau_1 - \tau_2| \geq 0.15 \), using a first differenced version of (2.1). Conditioned on these break fraction estimates (denoted by \( \hat{\tau}_1, \hat{\tau}_2 \)), we then estimated the local GLS regression of \( y_{\hat{\rho}} := [y_1, y_2 - \hat{\rho}_1, \ldots, y_T - \hat{\rho}_{T-1}]' \) on

---

7. These data were also considered by HLT in the context of allowing up to a single break in trend.

8. The first difference regression-based method delivers break fraction estimators that are \( O_p(T^{-1}) \) consistent under unit root, local-to-unity and stationary errors; see, for example, Harris et al. (2009) in the single break case.
\[ Z_{\hat{p},T} := [z_1, z_2 - \hat{p}z_1, ..., z_T - \hat{p}z_{T-1}], \] where \( z_t := [1, t, DT_t(\hat{r}_1), DT_t(\hat{r}_2)]' \) with \( \hat{p} := 1 - \hat{c}/T \) and \( \hat{c} = 21.5 \) (as in the implementation of \( MDF_2 \)). Denoting the trend break coefficient estimates from this regression by \( \hat{\gamma}_1 \) and \( \hat{\gamma}_2 \), we then computed the implied local break magnitudes \( \hat{\kappa}_i = \hat{\gamma}_i \hat{\omega}_z^{1/2}T^{1/2}, \) \( i = 1, 2 \), with \( \hat{\omega}_z^2 \) obtained from the residuals from the local GLS regression, using a Bartlett kernel with bandwidth parameter \( [12(T/100)^{1/4}] \). Table 3 reports the results, and Figure 7 also plots the fitted broken trend paths using the local GLS regression estimates. For Copper, Hides and Silver, the estimated magnitudes for both breaks are comfortably away from zero, suggesting that two breaks in trend are present in these series. However, the \( \hat{\kappa}_1 \) and \( \hat{\kappa}_2 \) values obtained are intermediate magnitudes, and are also of opposite sign, coinciding with the range of local breaks for which \( CKP_2 \) suffered the power “valley” phenomenon due to the lack of detection of these breaks by its pre-test procedure. That breaks of this order of magnitude are readily found in economic time series highlights the importance of the local break analysis and study of the power “valley” region, and reinforces the simulation findings that the \( MDF_m \) approach can be more powerful than approaches designed to achieve asymptotic efficiency under the extremes of a zero break or fixed break assumption. For Lead, one break magnitude is close to zero while the other is again in the “valley” region; here it seems that one break is present, again compromising the power of \( CKP_2 \), while \( MDF_2 \) rejects. As might be expected in this situation, \( MDF_1 \) also rejects for this series.

## 5 Conclusions

In this paper we have focused on the problem of testing for an autoregressive unit root in the presence of trend breaks. Uncertainty as to whether trend breaks are present in the data or not represents a considerable complication in this exercise. We have shown that existing approaches based on trend break pre-tests, while asymptotically near efficient when breaks have either zero or fixed (independent of the sample size) magnitudes, can display catastrophically low (even zero) finite sample power for even quite substantially sized breaks. This occurs because some degree of uncertainty will always exist in finite samples as to whether breaks occur or not, the degree of uncertainty depending on the precise pattern of breaks and their magnitudes. Low power then arises in regions where the pre-tests have very low power to detect breaks and, hence, incorrectly make no allowance for trend breaks in the resulting unit root tests.

In response to these findings we have advocated alternative unit root tests, both for the single and multiple possible break cases, based on the infimum of the sequence (across all candidate break points in a trimmed range) of local GLS detrended augmented Dickey-Fuller-type statistics. Although these tests have no formal asymptotic optimality properties in the presence of breaks of fixed magnitude, we have shown them to have the highly desirable property that they eliminate the aforementioned power “valleys”, thereby retaining good finite sample power in the presence of plausibly-sized trend breaks. Indeed, the region of local trend break magnitudes where the infimum-based tests display a significant power advantage over the procedures based around trend break pre-testing is far from a theoretical irrelevance; on the contrary, the evidence presented in this paper suggests that it is a region that is
really rather important to those analyzing empirical data. Unlike the corresponding OLS detrended infimum tests, we have also shown that these tests display no tendency to spuriously reject when trend breaks (either of fixed or local-to-zero magnitude) occur under the unit root null hypothesis. We recommend these tests to practitioners as we believe they come as close as is possible to providing invariant inference with respect to the trend break magnitudes in the context of this testing problem.

References


A Appendix

Proof of Theorem 1. For expositional brevity we demonstrate the result of Theorem 1 in the serially uncorrelated case \( \varepsilon_t = v_t \), setting \( k = 0 \) in (2.3). The result continues to hold under the more general conditions for \( \varepsilon_t \) of Assumption 1, provided \( 1/k + k^3/T \to 0 \) as \( T \to \infty \). In what follows, we may set \( \mu = \beta = 0 \) and \( \omega_\varepsilon = 1 \) without loss of generality.

First, for any \( \tau_1, \ldots, \tau_m \in \Lambda \), consider the estimators \( \tilde{\mu}, \tilde{\beta} \) and \( \tilde{\kappa} \)

\[
\begin{bmatrix}
\tilde{\mu} \\
\tilde{\beta} \\
\tilde{\kappa}
\end{bmatrix} = 
\begin{bmatrix}
g_{11} & g_{12} & \mathbf{j}_1 \\
g_{12} & g_{22} & \mathbf{j}_2 \\
\mathbf{j}_1 & \mathbf{j}_2 & \mathbf{Q}
\end{bmatrix}^{-1}
\begin{bmatrix}
h_1^y \\
h_2^y \\
p
\end{bmatrix}
\]

where

\[
g_{11} = 1 + (1 - \bar{\rho})^2(T - 1),
g_{12} = 1 + (1 - \bar{\rho}) \sum_{t=2}^{T} \{ t - \bar{\rho}(t - 1) \},
g_{22} = 1 + \sum_{t=2}^{T} \{ t - \bar{\rho}(t - 1) \}^2
\]

and

\[
h_1^y = y_1 + (1 - \bar{\rho}) \sum_{t=2}^{T} \{ y_t - \bar{\rho}y_{t-1} \},
h_2^y = y_1 + \sum_{t=2}^{T} \{ y_t - \bar{\rho}y_{t-1} \} \{ t - \bar{\rho}(t - 1) \}
\]

and where the vectors \( \mathbf{j}_1, \mathbf{j}_2 \) and \( \mathbf{p} \), and the matrix \( \mathbf{Q} \), have elements

\[
\begin{aligned}
j_{1i} &:= (1 - \bar{\rho}) \sum_{t=\lceil \tau_iT \rceil + 1}^{T} \{ t - \lceil \tau_iT \rceil - \bar{\rho}(t - \lceil \tau_iT \rceil - 1) \}, \\
j_{2i} &:= \sum_{t=\lceil \tau_iT \rceil + 1}^{T} \{ t - \bar{\rho}(t - 1) \} \{ t - \lceil \tau_iT \rceil - \bar{\rho}(t - \lceil \tau_iT \rceil - 1) \}, \\
p_i &:= \sum_{t=\lceil \tau_iT \rceil + 1}^{T} \{ y_t - \bar{\rho}y_{t-1} \} \{ t - \lceil \tau_iT \rceil - \bar{\rho}(t - \lceil \tau_iT \rceil - 1) \}, \\
g_{ij} &:= \sum_{t=\max(\tau_i, \tau_j)T + 1}^{T} \{ t - \lceil \tau_iT \rceil - \bar{\rho}(t - \lceil \tau_iT \rceil - 1) \} \{ t - \lceil \tau_jT \rceil - \bar{\rho}(t - \lceil \tau_jT \rceil - 1) \}
\end{aligned}
\]

respectively. The limits of the \( g_{ij} \) terms 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 elements of \( \mathbf{j}_1, \mathbf{j}_2 \) and \( \mathbf{Q} \) are as follows

\[
\begin{aligned}
j_{1i} &= \bar{c}T^{-1} \sum_{t=\lceil \tau_iT \rceil + 1}^{T} \{ 1 + \bar{c}T^{-1}(t - \lceil \tau_iT \rceil - 1) \} \\
&= \bar{c}(1 - \tau_i + \bar{c}T^{-2} \sum_{t=\lceil \tau_iT \rceil + 1}^{T} t - \bar{c}T^{-2}(\lceil \tau_iT \rceil + 1)(T - \lceil \tau_iT \rceil)) \\
&= \bar{c}(1 - \tau_i)(1 - \bar{c}T) + \bar{c}^2(1 - \tau_i^2)/2,
\end{aligned}
\]

\[
\begin{aligned}
T^{-1}j_{2i} &= 1 - \tau_i + \bar{c}^2T^{-3} \sum_{t=\lceil \tau_iT \rceil + 1}^{T} t(t - \lceil \tau_iT \rceil) \\
&\quad + \bar{c}T^{-2} \sum_{t=\lceil \tau_iT \rceil + 1}^{T} t + \bar{c}T^{-2} \sum_{t=\lceil \tau_iT \rceil + 1}^{T} (t - \lceil \tau_iT \rceil) + o(1) \\
&= (1 - \tau_i)(1 - \bar{c}T) + \bar{c}^2(1 - \tau_i^2)/3 - \bar{c}^2\tau_i(1 - \tau_i^2)/2 + \bar{c}(1 - \tau_i^2) \\
&= a_{\bar{c}} - \tau_i(1 + \bar{c} + \bar{c}^2/2 - \bar{c}^2\tau_i^2/6) =: m_{\bar{c}}(\tau_i),
\end{aligned}
\]
\[ T^{-1}q_{ij} = T^{-1}\sum_{t=[\tau_{\max}]1}^{T}\{(1 + \tilde{c}T^{-1}(t - [\tau_iT]))\{1 + \tilde{c}T^{-1}(t - [\tau_jT])\} + o(1) \]
\[ = 1 - \tau_{\max} + \tilde{c}^2T^{-3}\sum_{t=[\tau_{\max}]1}^{T}(t - [\tau_iT])(t - [\tau_jT]) + o(1) \]
\[ + \tilde{c}T^{-2}\sum_{t=[\tau_{\max}]1}^{T}(2t - [\tau_iT] - [\tau_jT]) + o(1) \]
\[ \Rightarrow 1 - \tau_{\max} + \tilde{c}^2(1 - \tau_{\max}^2)/3 - \tilde{c}^2(\tau_i + \tau_j)(1 - \tau_{\max}^2)/2 \]
\[ + \tilde{c}^2\tau_i\tau_j(1 - \tau_{\max}) + 2\tilde{c}(1 - \tau_{\max}^2)/2 - \tilde{c}(\tau_i + \tau_j)(1 - \tau_{\max}) \]
\[ = (1 - \tau_{\max})\{a_c - \tilde{c}\tau_{\max} - \tilde{c}(\tau_i + \tau_j) + \tilde{c}^2(\tau_{\max} + \tau_{\max}^2)/3 - \tilde{c}^2(\tau_i + \tau_j)(1 + \tau_{\max})/2 + \tilde{c}^2\tau_i\tau_j\} \]
\[ =: d_c(\tau_i, \tau_j) \]

while for \( h^y_1, h^y_2 \) and the elements of \( p \) we obtain

\[ h^y_1 = y_1 + \tilde{c}T^{-1}(y_T - y_1) + \tilde{c}^2T^{-2}\sum_{t=2}^{T}y_{t-1} = u_1 + o_p(1), \]

\[ T^{-1/2}h^y_2 = \tilde{c}T^{-3/2}\sum_{t=1}^{T}t\Delta u_t + T^{-1/2}u_T + \tilde{c}^2T^{-5/2}\sum_{t=1}^{T}tu_{t-1} + \tilde{c}^2T^{-3/2}\sum_{t=1}^{T}u_{t-1} \]
\[ + \tilde{c}\sum_{j=1}^{m}k_jT^{-2}\sum_{t=1}^{T}tDU_{j}(\tau_0,i) + \sum_{j=1}^{m}k_jT^{-1}(T - [\tau_0,i]) \]
\[ + \tilde{c}^2\sum_{j=1}^{m}k_jT^{-2}\sum_{t=1}^{T}tDU_{j-1}(\tau_0,i) + \tilde{c}\sum_{j=1}^{m}k_jT^{-2}\sum_{t=1}^{T}DU_{j-1}(\tau_0,i) + o_p(1) \]
\[ \Rightarrow \tilde{c}\{W_c(1) - \int_{1}^{t}W_c(s)ds\} + W_c(1) + \tilde{c}\int_{0}^{1}sW_c(s)ds + \tilde{c}\int_{0}^{1}W_c(s)ds \]
\[ + \tilde{c}\sum_{j=1}^{m}k_jT^{-1}\{1 - \tau_{\max}\}/2 + \sum_{j=1}^{m}k_j(1 - \tau_{\max}) \]
\[ + \tilde{c}^2\sum_{j=1}^{m}k_j\{(1 - \tau_{\max}^2)/3 - \tau_{0,i}(1 - \tau_{\max}^2)/2\} + \tilde{c}\sum_{j=1}^{m}k_j(1 - \tau_{\max})^2/2 \]
\[ = (1 + \tilde{c})W_c(1) + \tilde{c}\int_{0}^{1}sW_c(s)ds + \sum_{j=1}^{m}k_j(1 - \tau_{\max})\{a_c - \tilde{c}^2\tau_{\max} - \tilde{c}(\tau_i + \tau_j) + \tilde{c}^2(\tau_{\max} + \tau_{\max}^2)/3 - \tilde{c}^2(\tau_i + \tau_j)(1 + \tau_{\max})/2 + \tilde{c}^2\tau_i\tau_j\} \]
\[ =: b_{c,c}(\tau_i, \tau_j), \]

\[ T^{-1/2}p_i = T^{-1/2}u_T - T^{-1/2}u_{[\tau_i,T]} + \tilde{c}T^{-3/2}\sum_{t=[\tau_i,T]}^{T}t\Delta u_t - \tilde{c}\tau_{[\tau_i,T]}^{-1/2}\sum_{t=[\tau_i,T]}^{T}t\Delta u_t \]
\[ + \tilde{c}T^{-5/2}\sum_{t=[\tau_i,T]}^{T}tu_{t-1} + \tilde{c}^2T^{-3/2}\sum_{t=[\tau_i,T]}^{T}u_{t-1} \]
\[ + \sum_{j=1}^{m}k_j(1 - \tau_{0,j}) - \sum_{j=1}^{m}k_j(\tau_i - \tau_{0,j})\tilde{t}_{\tau_i}^{\tau_{0,j}} \]
\[ + \tilde{c}\sum_{j=1}^{m}k_jT^{-1}\sum_{t=[\tau_i,T]}^{T}tDU_{j-1}(\tau_0,i) + \tilde{c}\sum_{j=1}^{m}k_jT^{-2}\sum_{t=[\tau_i,T]}^{T}DU_{j-1}(\tau_0,i) \]
\[ + \tilde{c}^2\sum_{j=1}^{m}k_jT^{-3}\sum_{t=[\tau_i,T]}^{T}DU_{j-1}(\tau_0,i) - \tilde{c}\sum_{j=1}^{m}k_j\tau_iT^{-2}\sum_{t=[\tau_i,T]}^{T}DU_{j-1}(\tau_0,i) \]
\[ =: W_c(1) - W_c(\tau_i) + \tilde{c}\{W_c(1) - \tau_iW_c(\tau_i) - \int_{1}^{\tau_i}W_c(s)ds\} - \tilde{c}\tau_i\{W_c(1) - W_c(\tau_i)\} \]
\[ + \tilde{c}\int_{1}^{\tau_i}W_c(s)ds + \tilde{c}\int_{1}^{\tau_i}sW_c(s)ds - \tilde{c}\tau_i\int_{1}^{\tau_i}W_c(s)ds \]
\[ + \sum_{j=1}^{m}k_j(1 - \tau_{0,j}) - \sum_{j=1}^{m}k_j(\tau_i - \tau_{0,j})\tilde{t}_{\tau_i}^{\tau_{0,j}} \]
\[ + \tilde{c}\sum_{j=1}^{m}k_j\{(1 - \tau_{0,j})/2 - (\tau_i - \tau_{0,j})\tilde{t}_{\tau_i}^{\tau_{0,j}}/2\} - \tilde{c}\sum_{j=1}^{m}k_j\tau_i\{1 - \tau_{0,j} - (\tau_i - \tau_{0,j})\tilde{t}_{\tau_i}^{\tau_{0,j}}\} \]
\[ + \tilde{c}\sum_{j=1}^{m}k_j\{(1 - \tau_{0,j})/2 - (\tau_i - \tau_{0,j})\tilde{t}_{\tau_i}^{\tau_{0,j}}/2\} \]
\[ + \tilde{c}^2\sum_{j=1}^{m}k_j\{(1 - \tau_{0,j})/3 - \tau_{0,j}(1 - \tau_{0,j})/2 - (\tau_i - \tau_{0,j})/3 - \tau_{0,j}(\tau_i - \tau_{0,j})/2\} \]
\[ = (1 + \tilde{c})W_c(1) - W_c(\tau_i) + \tilde{c}\int_{1}^{\tau_i}(s - \tau_i)W_c(s)ds \]
\[ + \sum_{j=1}^{m}k_j\{1 - \tau_{0,j}\}a_c - \tilde{c}\tau_i - \tilde{c}^2\tau_i(1 + \tau_{0,j})/2 - \tilde{c}^2\tau_{0,j}(1 + \tau_{0,j})/6 \]
\[ - (\tau_i - \tau_{0,j})\{1 - \tilde{c}^2(\tau_i - \tau_{0,j})^2/6\} \]
\[ =: b_{c,c}(\tau_i, \tau_j), \]
Collecting results, we therefore have that
\[
\begin{bmatrix}
\hat{\mu} \\
T^{1/2}\hat{\beta}
\end{bmatrix}
= \begin{bmatrix}
g_{11} & T^{-1/2}g_{12} & T^{-1/2}j_1' \\
T^{-1/2}g_{12} & T^{-1}g_{22} & T^{-1}j_2' \\
T^{-1/2}j_1 & T^{-1}j_2 & T^{-1}Q
\end{bmatrix}^{-1}
\begin{bmatrix}
h_1 \\
T^{-1/2}h_2 \\
T^{-1/2}p
\end{bmatrix}
\Rightarrow
\begin{bmatrix}
1 & 0 & 0' \\
0 & a_c & m_c(\tau)' \\
0 & m_c(\tau) & D_c(\tau)
\end{bmatrix}^{-1}
\begin{bmatrix}
u_1 \\
b_{c,e} + \kappa'f_{c,e}(\tau_0) \\
b_{c,e}(\tau) + F_{c,e}(\tau_0, \tau)\kappa
\end{bmatrix}
\]
giving of the limit of \( T^{-1/2}\hat{u}_{[rT]} \) as
\[
T^{-1/2}\hat{u}_{[rT]} = \begin{bmatrix}
T^{-1/2}y_{[rT]} - T^{-1/2}\hat{\mu} - T^{-1/2}\hat{\beta}_{[rT]} - T^{-1/2}\hat{\kappa}'\{(rT) - [rT]\} \circ I_r' \\
T^{-1/2}u_{[rT]} + \kappa'\{(r - \tau_0) \circ I_r'\}
\end{bmatrix}
= \begin{bmatrix}
\begin{bmatrix}
r \\
(r - \tau) \circ I_r'
\end{bmatrix}' \\
\begin{bmatrix}
a_c & m_c(\tau)' \\
m_c(\tau) & D_c(\tau)
\end{bmatrix}^{-1}
\end{bmatrix}
\begin{bmatrix}
\begin{bmatrix}
0 \\
b_{c,e} + \kappa'f_{c,e}(\tau_0) \\
b_{c,e}(\tau) + F_{c,e}(\tau_0, \tau)\kappa
\end{bmatrix}
\end{bmatrix}
\]
\[
= L_{c,e}(r, \tau_0, \tau, \kappa).
\]

It is straightforward to show that the regression error variance estimator from (2.3) converges in probability to one, and application of the continuous mapping theorem (CMT) then gives \( D_{c,e}^{GLS}(\tau) \Rightarrow D_{c,e}(\tau_0, \tau, \kappa) \). The stated result for \( MDF_{m} \) then follows from this fixed \( \tau \) representation, using the arguments proved in Zivot and Andrews (1992) and an application of the CMT.

**Proof of Theorem 2.** In what follows, we may set \( \sigma_e = 1 \) without loss of generality. To ease notation, we write \( [\tau_{0,1} T] \) and \( [\tau_{1} T] \) as simply \( \tau_{0,1}T \) and \( \tau_{1}T \), respectively. We first show that it is only necessary to consider the region \( \tau_1 \in \Lambda' \). Write
\[
y_{\hat{\theta}} = Z_{\hat{\theta},\tau_{0,1}}\begin{bmatrix}
\mu \\
\beta \\
\gamma_1
\end{bmatrix} + \mathbf{u}_\hat{\theta}
\]
where \( \mathbf{u}_\hat{\theta} = [u_1, u_2 - \hat{\rho}u_1, ..., u_T - \hat{\rho}u_{T-1}]' \). Then, for \( \theta = [\mu, \beta, \gamma_1]' \),
\[
\tilde{\theta} = (Z_{\theta,\tau_{1}} Z_{\theta,\tau_{1}})^{-1}Z_{\theta,\tau_{1}} y_{\hat{\theta}} = (Z_{\theta,\tau_{1}} Z_{\theta,\tau_{1}})^{-1}Z_{\theta,\tau_{1}} (Z_{\theta,\tau_{0,1}} \theta + \mathbf{u}_\theta) = (Z_{\theta,\tau_{1}} Z_{\theta,\tau_{1}})^{-1}Z_{\theta,\tau_{1}} (Z_{\theta,\tau_{0,1}} \theta + (Z_{\theta,\tau_{0,1}} - Z_{\theta,\tau_{1}}) \theta + \mathbf{u}_\theta)
\]
\[
\tilde{\theta} - \theta = (Z_{\theta,\tau_{1}} Z_{\theta,\tau_{1}})^{-1}Z_{\theta,\tau_{1}} (Z_{\theta,\tau_{0,1}} - Z_{\theta,\tau_{1}}) \theta + (Z_{\theta,\tau_{1}} Z_{\theta,\tau_{1}})^{-1}Z_{\theta,\tau_{1}} \mathbf{u}_\theta.
\]
Now, using notation from the proof of Theorem 1 above,
\[
Z_{\theta,\tau_{1}} Z_{\theta,\tau_{1}} = \begin{bmatrix}
g_{11} & g_{12} & j_{11} \\
g_{12} & g_{22} & j_{21} \\
j_{11} & j_{21} & g_{11}
\end{bmatrix} =: \mathbf{G}
\]

20
with the limits of the elements of \( G \) as given in the proof of Theorem 1. Next,

\[
Z'_{\bar{p}, \tau, 1} u_{\bar{p}} = \begin{bmatrix} h_1^u \\ h_2^u \\ h_3^u \end{bmatrix} = \mathbf{h}
\]

where

\[
h_1^u := u_1 + (1 - \bar{p}) \sum_{t=2}^{T} (u_t - \bar{p} u_{t-1}) \Rightarrow u_1.
\]

\[
 T^{-1/2} h_2^u := T^{-1/2} \sum_{t=2}^{T} \{u_t - \bar{p} u_{t-1}\} \{t - \bar{p} (t - 1)\} \Rightarrow b_{c,c},
\]

\[
 T^{-1/2} h_3^u := T^{-1/2} \sum_{t=|\tau_1 T| + 1}^{T} \{u_t - \bar{p} u_{t-1}\} \{t - \tau_1 T - \bar{p} (t - \tau_1 T - 1)\} \Rightarrow b_{c,c}(\tau_1).
\]

Also,

\[
Z'_{\bar{p}, \tau, 1} (Z_{\bar{p}, \tau_0, 1} - Z_{\bar{p}, \tau_1}) = \begin{bmatrix} 0 & 0 & k_1 \\ 0 & 0 & k_2 \\ 0 & 0 & k_3 \end{bmatrix} =: \mathbf{K}
\]

where, letting \( d_T := \tau_1 T - \tau_0, T \)

\[
k_1 := \begin{cases} 
(1 - \bar{p}) \sum_{t=|\tau_0, T + 1}^{T} \{t - \tau_0 T - \bar{p} (t - \tau_0 T - 1) - (t - \tau_1 T) + \bar{p} (t - \tau_1 T - 1)\} & \tau_1 T \leq \tau_0 T \\
- \sum_{t=|\tau_1 T + 1}^{T} \{t - \tau_1 T - \bar{p} (t - \tau_1 T - 1)\} & \tau_1 T > \tau_0 T
\end{cases}
\]

\[
k_2 := \begin{cases} 
\sum_{t=|\tau_0, T + 1}^{T} \{t - \bar{p} (t - 1)\} \{t - \tau_0 T - \bar{p} (t - \tau_0 T - 1) - (t - \tau_1 T) + \bar{p} (t - \tau_1 T - 1)\} & \tau_1 T \leq \tau_0 T \\
- \sum_{t=|\tau_1 T + 1}^{T} \{t - \bar{p} (t - 1)\} \{t - \tau_1 T - \bar{p} (t - \tau_1 T - 1)\} & \tau_1 T > \tau_0 T
\end{cases}
\]

and, when \( d_T = O(T^\delta), \delta \leq 1 \), i.e. \( \tau_1 - \tau_0, 1 = O(T^{\delta-1}) \)

\[
T^{-1/2} k_2 := \begin{cases} 
(1 + \bar{c})(1 - \tau_0^2)/2 + (T^{-1/2} d_T) & \tau_1 T \leq \tau_0 T \\
+ \bar{c}(T^{-1/2} d_T) (T^{-1/2} d_T)/2 - c(T^{-1/2} d_T)^2/2 & \tau_1 T > \tau_0 T
\end{cases}
\]

\[
= \begin{cases} 
o(1) & \delta \leq \frac{1}{2} \\
o(1) & \delta = \frac{1}{2} \\
o(T^{\delta-1/2}) & \frac{1}{2} < \delta < 1
\end{cases}
\]

21
Scaling appropriately we can write

$$\Omega_T^{-1}(\tilde{\theta} - \theta) = (\Gamma_T G \Omega_T)^{-1} \Gamma_T K \theta + (\Delta_T G \Omega_T)^{-1} \Delta_T h$$

where

$$\Omega_T := \begin{bmatrix} 1 & 0 & 0 \\ 0 & T^{\delta'} & 0 \\ 0 & 0 & T^{\delta'} \end{bmatrix}, \quad \Gamma_T := \begin{bmatrix} 1 & 0 & 0 \\ 0 & T^{-\delta'} & 0 \\ 0 & 0 & T^{-\delta'} \end{bmatrix}, \quad \Delta_T := \begin{bmatrix} 1 & 0 & 0 \\ 0 & T^{-1/2} & 0 \\ 0 & 0 & T^{-1/2} \end{bmatrix}$$

with $\delta' := \max(\delta, \frac{1}{2})$. Note that when $\delta \leq \frac{1}{2}$, $\Omega_T = \Gamma_T = \Delta_T$. Also,

$$\Delta_T^{-1}(\tilde{\theta} - \theta) = (\Delta_T G \Delta_T)^{-1} \Delta_T h + o(1) \quad \delta < \frac{1}{2}$$

$$\Delta_T^{-1}(\tilde{\theta} - \theta) = (\Delta_T G \Delta_T)^{-1} \Delta_T K \theta + (\Delta_T G \Delta_T)^{-1} \Delta_T h \quad \delta = \frac{1}{2}$$

$$\Omega_T^{-1}(\tilde{\theta} - \theta) = (\Gamma_T G D T)^{-1} \Gamma_T K \theta + o_p(1) \quad \frac{1}{2} < \delta \leq 1.$$

Now $\tilde{u}_t$ can be written as follows, using $z_t := [t, t, DT(t')]'$,

$$\tilde{u}_t = \sum_{t=1}^{T} \Gamma_T (\tilde{\theta} - \theta) z_t + \gamma_1 (t - \tau_0 T) T_{\tau_0 T}^T \hat{u}_t$$

Since $T^{\delta'}(t - \tau_0 T) T^{\delta'}(t - \tau_1 T)$ are bounded by an $O(1)$ variate when $\tau_0 T + \tau_1 T = 1$ and $\tau_0 T + \tau_1 T = 1$, respectively, it follows that $T^{\delta'} \tilde{u}_t = O_p(1)$, and the following orders can be established. First, since $\tau_1 - \tau_0$ is bounded by an $O(1)$ variate and $\tilde{\theta} - \theta$ is bounded by an $O_p(1)$ variate, it is straightforward to show that $T^{-1} \sum_{t=2}^{T} \Delta \tilde{u}_t^2 = O_p(1)$ regardless of the value of $\delta$. Then,

$$\sum_{t=2}^{T} \Delta \tilde{u}_t \tilde{u}_{t-1} = (\tilde{u}_T^2 - \tilde{u}_1^2 - \sum_{t=2}^{T} \Delta \tilde{u}_t^2) / 2 = O_p(T^{2\delta'})$$

$$\sum_{t=2}^{T} \tilde{u}_t^2 = O_p(T^{1+2\delta'})$$
\[
\hat{\sigma}_e^2 = T^{-1} \left\{ \sum_{t=2}^{T} \Delta \hat{u}_t^2 - \left( \frac{\sum_{t=2}^{T} \Delta \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^{T} \hat{u}_t^2} \right)^2 \right\} = O_p(1)
\]

and so

\[
DF^{GLS}_e(\tau_1) = O_p(T^{\delta' - 1/2}).
\]

When \( \delta \leq \frac{1}{2} \), since \( \delta' = \frac{1}{2} \) we find that \( DF^{GLS}_e(\tau_1) = O_p(1) \). Conversely, when \( \frac{1}{2} < \delta \leq 1 \), since \( \delta' = \delta \) we find \( DF^{GLS}_e(\tau_1) = O_p(T^{\delta - 1/2}) \), i.e. it is divergent. Moreover, in this latter case, the numerator is

\[
T^{-2\delta} \sum_{t=2}^{T} \Delta \hat{u}_t \hat{u}_{t-1} = \left\{ (T^{-\delta} \hat{u}_t^2)^2 - (T^{-\delta} \hat{u}_1)^2 \right\}/2 - T^{1-2\delta}(T-1 \sum_{t=2}^{T} \Delta \hat{u}_t^2)/2 + o_p(1)
\]

since \( T^{-1} \sum_{t=2}^{T} \Delta \hat{u}_t^2 = O_p(1) \) and \( T^{-\delta} \hat{u}_1 = T^{-\delta} u_1 - T^{-\delta}(\hat{\mu} - \mu) - T^{-\delta}(\hat{\beta} - \beta) = o_p(1) \). Therefore, when \( \frac{1}{2} < \delta \leq 1 \), we obtain that \( DF^{GLS}_e(\tau_1) \Rightarrow +\infty \).

As a result, the minimum of \( DF^{GLS}_e(\tau_1) \) over \( \tau_1 \in \Lambda \) cannot in the limit be obtained for a value of \( \tau_1 \) satisfying \( d_T = O(T^\delta) \) with \( \frac{1}{2} < \delta \leq 1 \), i.e. any \( \tau_1 \) lying outside of a \( T^{-1/2} \) neighbourhood of \( \tau_{0,1} \). Hence, asymptotically, it follows that

\[
\inf_{\tau_1 \in \Lambda} DF^{GLS}_e(\tau_1) = \inf_{\tau_1 \in \Lambda} DF^{GLS}_e(\tau_1).
\]

As we can restrict attention to \( \tau_1 \in \Lambda' \), we need only consider \( \tau_1 \) satisfying \( \tau_1 - \tau_{0,1} = O(T^{\delta-1}) \) with \( \delta \leq \frac{1}{2} \) (and \( \delta' = \frac{1}{2} \)). If \( \delta = \frac{1}{2} \), let \( \tau_1 - \tau_{0,1} = \pi_{\tau_1-\tau_{0,1}} T^{-1/2} \) for some non-zero finite constant \( \pi_{\tau_1-\tau_{0,1}} \), so that \( T^{-1/2} d_T \rightarrow \lim T^{1/2}(\tau_1 - \tau_{0,1}) = \pi_{\tau_1-\tau_{0,1}} \). On the other hand, if \( \delta < \frac{1}{2} \) so that \( \tau_1 - \tau_{0,1} = o(T^{-1/2}) \), it is clear that \( T^{-1/2} d_T \rightarrow 0 \). Nesting these two possibilities, we can write

\[
\lim T^{1/2}(\tau_1 - \tau_{0,1}) = \pi^{'}_{\tau_1-\tau_{0,1}}, \quad \text{where} \quad 0 \leq |\pi^{'}_{\tau_1-\tau_{0,1}}| < \infty \quad (\text{i.e.} \quad \pi^{'}_{\tau_1-\tau_{0,1}} = \pi_{\tau_1-\tau_{0,1}} \text{ if } \delta = \frac{1}{2} \text{ and } \pi^{'}_{\tau_1-\tau_{0,1}} = 0 \text{ if } \delta < \frac{1}{2}).
\]

Then, using \( \tau_1 - \tau_{0,1} = o(1) \)

\[
\Delta_T^{-1}(\hat{\theta} - \theta) = \left[ \begin{array}{c} \hat{\mu} - \mu \\ T^{1/2} (\hat{\beta} - \beta) \\ T^{1/2} (\hat{\gamma}_1 - \gamma_1) \end{array} \right] = (\Delta_T G \Delta_T)^{-1} \Delta_T K \theta + (\Delta_T G \Delta_T)^{-1} \Delta_T h
\]

Therefore, using \( \Delta_T G \Delta_T \)

\[
\Rightarrow \left[ \begin{array}{ccc} 1 & 0 & 0 \\ 0 & a_e & m_e(\tau_{0,1}) \\ 0 & m_e(\tau_{0,1}) & d_e(\tau_{0,1}, \tau_{0,1}) \end{array} \right]^{-1} \left[ \begin{array}{cccc} 0 & 0 & 0 \\ 0 & \pi^{'}_{\tau_1-\tau_{0,1}} l_1, e(\tau_{0,1}) & \beta \\ 0 & \pi^{'}_{\tau_1-\tau_{0,1}} l_2, e(\tau_{0,1}) & \gamma_1 \end{array} \right] \left[ \begin{array}{c} \mu \\ \beta \\ \gamma_1 \end{array} \right] + \left[ \begin{array}{ccc} 0 & 0 & 0 \\ 0 & a_e & m_e(\tau_{0,1}) \\ 0 & m_e(\tau_{0,1}) & d_e(\tau_{0,1}, \tau_{0,1}) \end{array} \right]^{-1} \left[ \begin{array}{c} u_1 \\ b_{e,\hat{\theta}} \\ b_{e,\hat{\theta}} \end{array} \right]
\]

23
The limits of the terms in the numerator can be established as follows

\[
\begin{align*}
T^{-1/2}\widetilde{u}_T &= T^{-1/2}u_T - T^{-1/2}(\tilde{\mu} - \mu) - T^{-1/2}(\tilde{\beta} - \beta)t/T - T^{-1/2}(\tilde{\gamma}_1 - \gamma_1)\left(t/T - \tau_1\right)\mathbb{I}^t_{r_1T} \\
&+ \gamma_1 T^{-1/2}(\tau_1 - \tau_0,1)\mathbb{I}^t_{\tau_0,1T} + \gamma_1 T^{-1/2}(t - \tau_0,1)\mathbb{I}^t_{\tau_0,1T} + \gamma_1 T^{-1/2}(t - \tau_1)\mathbb{I}^{\tau_0,1T + 1\mathbb{I}^t_{r_1T}}
\end{align*}
\]

The unit root statistic evaluated at any given break fraction \(\tau_1\) is given by

\[
DF^{\text{GLS}}_c(\tau_1) = \frac{\sum_{t=2}^{T} \Delta \tilde{u}_{t-1}}{\sqrt{\hat{\sigma}_c^2 \sum_{t=2}^{T} \tilde{u}^2_{t-1}}}
= \frac{(T^{-1/2}\widetilde{u}_T)^2 - (T^{-1/2}\tilde{u}_1)^2}{2\sqrt{\hat{\sigma}_c^2 T - 2\sum_{t=2}^{T} \tilde{u}^2_{t-1}}}.
\]

The limits of the terms in the numerator can be established as follows

\[
\begin{align*}
T^{-1/2}\tilde{u}_1 &= T^{-1/2}u_1 - T^{-1/2}(\tilde{\mu} - \mu) - T^{-1/2}(\tilde{\beta} - \beta) \Rightarrow 0 \\
T^{-1/2}\tilde{u}_T &= T^{-1/2}u_T - T^{-1/2}(\tilde{\mu} - \mu) - T^{-1/2}(\tilde{\beta} - \beta) - T^{-1/2}(\tilde{\gamma}_1 - \gamma_1)(1 - \tau_1) + \gamma_1 T^{-1/2}(\tau_1 - \tau_0,1) \\
&\Rightarrow W_c(1) - \gamma_1 \pi'_{\tau_1 - \tau_0,1} \left[ \frac{1}{1 - \tau_0,1} \right]' \left[ \begin{array}{cc} a_c & m_c(\tau_0,1) \\ m_c(\tau_0,1) & d_c(\tau_0,1, \tau_0,1) \end{array} \right]^{-1} \left[ \begin{array}{c} l_{1,c}(\tau_0,1) \\ l_{2,c}(\tau_0,1) \end{array} \right] \\
&+ \left[ \frac{1}{1 - \tau_0,1} \right]' \left[ \begin{array}{cc} a_c & m_c(\tau_0,1) \\ m_c(\tau_0,1) & d_c(\tau_0,1, \tau_0,1) \end{array} \right]^{-1} \left[ \begin{array}{c} b_{c,c} \\ b_{c,c} \end{array} \right] + \gamma_1 \pi'_{\tau_1 - \tau_0,1} \\
&= J_c(1, \tau_1, \tau_0,1, \pi'_{\tau_1 - \tau_0,1}, \kappa_1)
\end{align*}
\]

noting that \(\gamma_1 = \kappa_1\), and

\[
\begin{align*}
T^{-1/2}\Delta \tilde{u}_t &= T^{-1/2}u_t - T^{-1/2}(\tilde{\beta} - \beta) - T^{-1/2}(\tilde{\gamma}_1 - \gamma_1)\mathbb{I}^t_{r_1T} \\
&+ \gamma_1 T^{-1/2}\tilde{u}^t_{\tau_0,1T} + \gamma_1 T^{-1/2}(\tau_1 - \tau_0,1)\mathbb{I}^t_{\tau_0,1T} \\
T^{-1}\sum_{t=2}^{T} \Delta \tilde{u}^2_t &= T^{-1} \sum_{t=2}^{T} \Delta u^2_t + (\tilde{\beta} - \beta)^2 + (1 - \tau_1)(\tilde{\gamma}_1 - \gamma_1)^2 + \gamma_1^2(\tau_1 - \tau_0,1)\mathbb{I}^{\tau_0,1}_{r_0,1T} + \gamma_1^2(\tau_0,1 - \tau_1)\mathbb{I}^\tau_{r_1T} \\
&+ \text{cross product terms} + o_p(1) \\
&= T^{-1} \sum_{t=2}^{T} \Delta u^2_t + o_p(1) \Rightarrow 1
\end{align*}
\]

using the Cauchy-Schwarz inequality to show that each cross product term is asymptotically negligible.

For the denominator terms, first we have that \(\hat{\sigma}_c^2 = T^{-1} \sum_{t=2}^{T} \Delta \tilde{u}^2_t + o_p(1) \Rightarrow 1\). Next, let

\[
A_t = T^{-1/2}u_t - T^{-1/2}(\tilde{\beta} - \beta)t/T - T^{-1/2}(\tilde{\gamma}_1 - \gamma_1)(t/T - \tau_1)\mathbb{I}^t_{r_1T} + \gamma_1 T^{-1/2}(\tau_1 - \tau_0,1)\mathbb{I}^t_{\tau_0,1T} + \mathbb{I}^t_{r_1T}
\]
and write
\[ T^{-1/2} \tilde{u}_t = A_t - T^{-1/2}(\tilde{\mu} - \mu) + \gamma_1 T^{-1/2}(t - \tau_{0,1}T)\bar{\pi}_{\tau_{0,1}T} + \gamma_1 T^{-1/2}(t - \tau_1T)\bar{\pi}_{\tau_1T} - \gamma_1 T^{-1/2}(t - \tau_1T)\bar{\pi}_{\tau_1T} \]

Then
\[
T^{-2} \sum_{t=2}^{T} \tilde{u}_{t-1}^2 = T^{-1} \sum_{t=2}^{T} A_t^2 - T^{-1}(\tilde{\mu} - \mu)^2 + \gamma_1^2 T^{-2} \sum_{t=\tau_0,1}^{T+1} (t - \tau_{0,1}T)^2 \\
+ \gamma_1^2 T^{-2} \sum_{t=\tau_1T+1}^{T+1} (t - \tau_1T)^2 + \text{cross product terms} + o_p(1) \\
= T^{-1} \sum_{t=2}^{T} A_t^2 + o_p(1)
\]
since \( T^{-1}(\tilde{\mu} - \mu)^2 = o_p(1), \)

\[
T^{-2} \sum_{t=\tau_0,1}^{T+1} (t - \tau_{0,1}T)^2 = T^{-1} (\tau_1 - \tau_{0,1})/6 + (\tau_1 - \tau_{0,1})^2/2 + T(\tau_1 - \tau_{0,1})^3/3 = O(T^{-1/2})
\]

and, similarly \( T^{-2} \sum_{t=\tau_1T+1}^{T+1} (t - \tau_1T)^2 = O(T^{-1/2}); \) since \( T^{-1} \sum_{t=2}^{T} A_t^2 \) is \( O_p(1) \) (as shown below), the cross product terms are then again asymptotically negligible, via applications of the Cauchy-Schwarz inequality. For the limit of \( A_t \), using \( \tau_1 - \tau_{0,1} = o(1) \) we find

\[
A_{[rT]} = T^{-1/2} u_{[rT]} - T^{-1/2}(\tilde{\beta} - \beta)r - T^{1/2}(\tau_1 - \gamma_1)(r - \tau_1)\bar{\pi}_{[rT]} + \gamma_1 T^{1/2}(\tau_1 - \tau_{0,1})\bar{\pi}_{[rT]1} \\
\Rightarrow W_c(r) - \gamma_1 \pi_{\tau_1-\tau_{0,1}} \left[ \begin{array}{c} r \\ (r - \tau_{0,1})\bar{\pi}_{\tau_{0,1}} \end{array} \right] \left[ \begin{array}{c} a_{\tilde{\varepsilon}} \\ m_{\varepsilon}(\tau_{0,1}) \end{array} \right]^{-1} \left[ \begin{array}{c} l_1,\tilde{\varepsilon}(\tau_{0,1}) \\ l_2,\tilde{\varepsilon}(\tau_{0,1}) \end{array} \right] \\
- \left[ \begin{array}{c} r \\ (r - \tau_{0,1})\bar{\pi}_{\tau_{0,1}} \end{array} \right] \left[ \begin{array}{c} a_{\tilde{\varepsilon}} \\ m_{\varepsilon}(\tau_{0,1}) \end{array} \right]^{-1} \left[ \begin{array}{c} b_{c,\tilde{\varepsilon}} \\ b_{c,\varepsilon}(\tau_{0,1}) \end{array} \right] \\
=: J_{c,\tilde{\varepsilon}}(r, \tau_{0,1}, \pi_{\tau_1-\tau_{0,1}}^{rT}, \kappa_1)
\]

and so
\[
T^{-2} \sum_{t=2}^{T} \tilde{u}_{t-1}^2 = T^{-1} \sum_{t=2}^{T} A_t^2 + o_p(1) \Rightarrow \int_0^1 J_{c,\tilde{\varepsilon}}(r, \tau_{0,1}, \pi_{\tau_1-\tau_{0,1}}^{rT}, \kappa_1)^2 dr.
\]

Application of the CMT then gives \( D\tilde{\varepsilon}_{GLS}^{CMT}(\tau_1) \Rightarrow D_{c,\tilde{\varepsilon}}(\tau_{0,1}, \pi_{\tau_1-\tau_{0,1}}, \kappa_1), \) and the result for \( MDF_1 \) then follows from this fixed \( \tau_1 \) representation, again using the arguments proved in Zivot and Andrews (1992) and an application of the CMT.
Table 1. Asymptotic \( \bar{c} \) values and \( \xi \) level critical values for \( MDF_m \)

<table>
<thead>
<tr>
<th>( m )</th>
<th>( \bar{c} )</th>
<th>( \xi = 0.10 )</th>
<th>( \xi = 0.05 )</th>
<th>( \xi = 0.01 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>17.6</td>
<td>-3.57</td>
<td>-3.85</td>
<td>-4.40</td>
</tr>
<tr>
<td>2</td>
<td>21.5</td>
<td>-4.30</td>
<td>-4.58</td>
<td>-5.10</td>
</tr>
<tr>
<td>3</td>
<td>25.5</td>
<td>-4.81</td>
<td>-5.06</td>
<td>-5.58</td>
</tr>
</tbody>
</table>

Table 2. Asymptotic size and local power of \( MDF_1 \) in the presence of a single fixed break in trend

<table>
<thead>
<tr>
<th>( c )</th>
<th>( \kappa_1 )</th>
<th>( \tau_{0.1} = 0.3 )</th>
<th>( \tau_{0.1} = 0.5 )</th>
<th>( \tau_{0.1} = 0.7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>0.030</td>
<td>0.030</td>
<td>0.021</td>
</tr>
<tr>
<td>2</td>
<td>0.030</td>
<td>0.030</td>
<td>0.021</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>0.403</td>
<td>0.405</td>
<td>0.388</td>
</tr>
<tr>
<td>2</td>
<td>0.402</td>
<td>0.404</td>
<td>0.387</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0.805</td>
<td>0.807</td>
<td>0.800</td>
</tr>
<tr>
<td>2</td>
<td>0.804</td>
<td>0.806</td>
<td>0.800</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Application of \( MDF_1, MDF_2 \) and \( CKP_2 \) to relative primary commodity prices

<table>
<thead>
<tr>
<th></th>
<th>( MDF_1 )</th>
<th>( MDF_2 )</th>
<th>( CKP_2 )</th>
<th>( \hat{\kappa}_1 )</th>
<th>( \hat{\kappa}_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Copper</td>
<td>-3.15</td>
<td>-4.81*</td>
<td>6.61</td>
<td>5.62</td>
<td>-6.63</td>
</tr>
<tr>
<td>Hides</td>
<td>-3.45</td>
<td>-6.34*</td>
<td>25.67</td>
<td>-7.04</td>
<td>3.48</td>
</tr>
<tr>
<td>Lead</td>
<td>-4.62*</td>
<td>-4.86*</td>
<td>18.23</td>
<td>0.21</td>
<td>-4.10</td>
</tr>
<tr>
<td>Silver</td>
<td>-2.04</td>
<td>-4.71*</td>
<td>14.15</td>
<td>6.84</td>
<td>-10.55</td>
</tr>
</tbody>
</table>

Note: * denotes rejection at the 0.05-level.
Figure 1. Asymptotic size: 1 break; $MDF_1$: ---, $CKP_1$: - - -, $Z_{AO}$: --
Figure 2. Asymptotic local power: 1 break; $MDF_1$: --, $CKP_1$: - - -
Figure 3. Finite sample power: 1 break, $T = 150$; $MDF_1$: —, $CKP_1$: --
Figure 4. Finite sample power: 1 break, $T = 300$; $MDF_1$: —, $CKP_1$: -- -
Figure 5. Finite sample power: 2 breaks, $T = 150$, $\tau_{0.1} = 0.33$, $\tau_{0.2} = 0.66$; $MDF_2$: ---, $CKP_2$: - - -
Figure 6. Finite sample power: 2 breaks, $T = 300$, $\tau_{0,1} = 0.33$, $\tau_{0,2} = 0.66$; $MDF_2$: $\ldots$, $CKP_2$: $- - -$
Figure 7. Relative primary commodity price series and fitted broken trends