Granger Centre
for time series econometrics

Granger Centre Discussion Paper Series

Testing for unit roots in the presence of a possible break in trend and non-stationary volatility

by

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

---

Granger Centre Discussion Paper No. 09/05

The University of Nottingham

Testing for Unit Roots in the Presence of a Possible Break in Trend and Non-Stationary Volatility*

Giuseppe Cavaliere[a], David I. Harvey[b], Stephen J. Leybourne[b] and A.M. Robert Taylor[b]

[a]Department of Statistical Sciences, University of Bologna

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

December 2009

Abstract

In this paper we analyse the impact of non-stationary volatility on the recently developed unit root tests which allow for a possible break in trend occurring at an unknown point in the sample, considered in Harris, Harvey, Leybourne and Taylor (2009) [HHLT]. HHLT's analysis hinges on a new break fraction estimator which, when a break in trend occurs, is consistent for the true break fraction at rate $O_p(T^{-1})$. Unlike other available estimators, however, when there is no trend break HHLT's estimator converges to zero at rate $O_p(T^{-1/2})$. In their analysis HHLT assume the shocks to follow a linear process driven by IID innovations. Our first contribution is to show that HHLT's break fraction estimator retains the same consistency properties as demonstrated by HHLT for the IID case when the innovations display non-stationary behaviour of a quite general form, including, for example, the case of a single break in the volatility of the innovations which may or may not occur at the same time as a break in trend. However, as we subsequently demonstrate, the limiting null distribution of unit root statistics based around this estimator are not pivotal in the presence of non-stationary volatility. Associated Monte Carlo evidence is presented to quantify the impact of a one-time change in volatility on both the asymptotic and finite sample behaviour of such tests. A solution to the identified inference problem is then provided by considering wild bootstrap-based implementations of the HHLT tests, using the trend break estimator from the original sample data. The proposed bootstrap method does not require the practitioner to specify a parametric model for volatility, and is shown to perform very well in practice across a range of models.

Keywords: Unit root tests; quasi difference de-trending; trend break; non-stationary volatility; wild bootstrap.

JEL Classification: C22.

*We thank Peter Philips, Tim Vogelsang, two anonymous referees and participants at the Granger Centre Conference on Bootstrap and Numerical Methods in Time Series, held in Nottingham on September 10–11, 2008 for useful comments on earlier versions of this paper. We also gratefully acknowledge the financial support provided by the Economic and Social Research Council of the UK under research grant RES-000-22-3882. Correspondence to: David Harvey, School of Economics, University of Nottingham, Nottingham, NG7 2RD, UK. E-mail: dave.harvey@nottingham.ac.uk

1 Introduction

In a recent paper Harris, Harvey, Leybourne and Taylor (2009) [HHLT] discuss solutions to a long-standing inference problem in time series econometrics. The problem in question dates back to the pioneering work of Perron (1989) who developed augmented Dickey-Fuller [ADF] type unit root tests which allow for a break in trend. Such breaks are considered to be prevalent in macroeconomic series; see, *inter alia*, Stock and Watson (1996,1999,2005) and Perron and Zhu (2005), and it is now regular applied practice to allow for the possibility of deterministic structural change in the trend function when testing for a unit root.

In his original work Perron (1989) treated the location of the potential trend break as known, *a priori*. This assumption has attracted significant criticism (see, for example, Christiano, 1992). In response to this criticism, subsequent approaches have focused on the case where the (possible) break occurs at an unknown point in the sample which needs to be estimated in some way; see, *inter alia*, Zivot and Andrews (1992), Banerjee *et al.* (1992), Perron (1997) and Perron and Rodríguez (2003), the last of which considers tests based on both the ADF principle and the more recent autocorrelation-robust $\mathcal{M}$-type tests of Perron and Ng (1996), Stock (1999) and Ng and Perron (2001). However, while the tests proposed in Perron (1989) are (exact) invariant to the magnitude of the trend break (provided any trend break does indeed occur at the assumed date), this is not true of these later tests. Tests of the type proposed in Zivot and Andrews (1992), which base inference on the minimum of a sequence of ADF statistics calculated for all possible break dates within a given range, are particularly problematic owing to the fact that the location of the minimum of the ADF statistics is not a consistent estimator of the true break fraction when a break occurs. For this reason it is necessary with these tests to make the infeasible assumption that no break in trend occurs under the unit root null hypothesis, such that tests with pivotal limiting null distributions can be obtained. Even with that assumption these tests have the undesirable property that their power functions depend on the magnitude of the trend break.

The tests developed in Perron (1997) and Perron and Rodríguez (2003) are explicitly based around an estimator of the unknown break fraction. Provided this estimator is consistent at a sufficiently fast rate (the rate in question must be faster than $O_p(T^{-1/2})$) then, where a break occurs, the resulting tests are asymptotically equivalent to the corresponding known break fraction test. In the case of the tests of Perron and Rodríguez (2003), which are based on quasi difference (QD) detrending, this entails that the resulting tests are near asymptotically efficient[1] where a break occurs. However, it is the behaviour of the breakpoint estimators used in these approaches when no break occurs which causes the inference problem. Here

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

the tests proposed in Perron and Rodríguez (2003) are not efficient, and the efficiency losses can be quite substantial relative to the efficient tests of Elliott *et al.* (1996) [ERS], as HHLT demonstrate for the ADF-type statistic. This occurs because, in the absence of a trend break, the break point estimator they propose has a non-degenerate limit distribution over the range of possible break dates from which it is calculated and, as such, will spuriously indicate the presence of an unnecessary trend break.

In practice, since it will be *unknown* as to whether a trend break occurs or not, this differing behaviour of the break fraction estimator also renders the *true* asymptotic critical values of the aforementioned tests dependent on whether a break occurs or not. For these tests to be feasible one faces a choice. Either, as in Perron and Rodríguez (2003), to use conservative critical values corresponding to the case where it is assumed that no break is present, with an associated loss of efficiency in cases where a break is present (and, indeed, where it is not, as noted above), or to use critical values which assume that a break is present but run the risk of over-sizing in the unit root tests when a break is in fact not present (coupled with the loss in efficiency which occurs when there is no break).

HHLT rectify these drawbacks with the existing tests by proposing a new break fraction estimator. This estimator is a data-dependent modification of the estimator of the break fraction obtained by using an OLS estimator on the first differences of the data (hereafter, the first difference estimator). This estimator possesses two key properties. First, when a break actually occurs, the estimator collapses to the first difference estimator, which converges to the true break fraction at rate $O_p(T^{-1})$, allowing critical values from the Perron and Rodríguez (2003) fixed break fraction distribution to be applied in the limit. Second, and crucially, when no break occurs, the estimator collapses to zero at rate $O_p(T^{-1/2})$ which is sufficiently fast such that the efficient ERS test with only a linear trend may be applied using the critical values given in ERS. Alternative approaches to this problem, based on pre-testing for a break in trend, have also been proposed in Kim and Perron (2009) and Carrion-i-Silvestre *et al.* (2009).

While the unit root tests discussed above allow for the possibility of breaks in the deterministic trend function of the series under consideration, significantly they do not allow for time-varying behaviour in the unconditional volatility of the driving shocks. This is a considerable drawback with these procedures especially bearing in mind the tendency (see, *inter alia*, the discussions in Busetti and Taylor, 2003, and Sensier and van Dijk, 2004) for series which display breaks in trend to display simultaneous breaks in unconditional volatility. Indeed, there is a large body of recent applied work suggesting non-constant behaviour, in particular a general decline, in the unconditional volatility of the shocks driving macroeconomic time-series over the past twenty years or so is a relatively common phenomenon; see, in particular the literature review in Cavaliere and Taylor (2008). As an example, and controlling for the possibility of level and/or trend breaks, Sensier and van Dijk (2004) report that over 80% of the real and price variables in the Stock and Watson (1999) data-set reject the null of constant innovation variance against the alternative of a one-off change in variance.

These findings have helped stimulate a research agenda analysing the effects of non-constant (or non-stationary) volatility on univariate unit root and stationarity tests; see, *inter alia*, Kim, Leybourne and Newbold (2002), Busetti and Taylor (2003), Cavaliere (2004), and Cavaliere and Taylor (2005,2007,2008). These authors show that standard unit root and stationarity tests based on the assumption of constant unconditional volatility can display significant size distortions in the presence of non-stationary volatility. Cavaliere and Taylor (2008) develop wild bootstrap-based implementations of standard unit root tests which are shown to yield pivotal inference in the presence of non-stationary volatility.

In this paper we analyse the impact of non-stationary volatility in the shocks on the ADF-type trend break unit root test of HHLT together with analogous tests based on the $\mathcal{M}$ testing principle. We allow for innovation processes whose unconditional variances evolve over time according to a quite general mechanism which allows, for example, single and multiple abrupt variance breaks, smooth transition variance breaks, and trending variances. We demonstrate that although the large sample properties of HHLT's modified break fraction estimator are unaltered from those which apply in the homoskedastic case in both the trend break and no trend break environments, the asymptotic distributions of the HHLT statistic and the corresponding $\mathcal{M}$ statistics are not pivotal and in both trend break and no trend break environments depend on the structure of the underlying volatility process. Simulation results for a one-time change in volatility suggest that this can have a large impact on both the size and power properties of the tests both in large and small samples.

In order to solve the identified inference problem, at least within the class of non-stationary volatility processes considered, we propose an implementation of the ADF-type test of HHLT and the corresponding $\mathcal{M}$-type tests using the wild bootstrap principle. The wild bootstrap replicates in the re-sampled data the pattern of heteroskedasticity present in the original shocks, and has been shown to be highly effective in the case of standard unit root tests which allow for either a constant or linear trend; see Cavaliere and Taylor (2008). As we shall demonstrate, the wild bootstrap implementations of the HHLT-type trend break unit root statistics share the same limiting null distribution as the original statistics under the class of non-stationary volatility considered, with this result holding in both the trend break and no trend break environments. Consequently, inference that is asymptotically robust under the unit root null hypothesis to non-stationary volatility in the shocks can be performed without the need to specify any parametric model of volatility and regardless of whether a trend break is present or not.

The paper is organised as follows. We outline our reference heteroskedastic trend break model in section 2. In section 3 we first review the modified break fraction estimator and associated ADF-type test developed for the homoskedastic case in HHLT, and also extend their analysis to the corresponding tests based on the $\mathcal{M}$ testing principle. We then demonstrate that HHLT's modified break fraction estimator retains the same rates of consistency in both the trend break and no trend break cases in the presence of non-stationary volatility as were stated in HHLT for the case of constant unconditional volatility. We also show that in both the trend break and no trend break cases, the unit root test proposed in HHLT, based around this modified estimator, has a non-pivotal limiting distribution with its form depending on the underlying volatility process. The same is shown to be true of the corresponding $\mathcal{M}$-type tests. The impact of a one-time change in volatility - including the case where this occurs simultaneously with a break in trend - on the asymptotic properties of these statistics is explored numerically through Monte Carlo simulations. In section 4 we propose a wild bootstrap-based implementation of the HHLT procedure. We demonstrate that the wild bootstrap analogue of the HHLT statistic replicates the first-order asymptotic null distribution of the standard HHLT statistic, such that the corresponding bootstrap tests are asymptotically valid, in the presence of non-stationary volatility. The same is shown to be true for the corresponding $\mathcal{M}$-type tests. Simulation evidence presented in Section 5 suggests that the proposed bootstrap tests perform well in small samples. Concluding remarks are offered in section 6. Proofs are collected in an Appendix.

In what follows we use the following notation: '$\lfloor \cdot \rfloor$' denotes the integer part of its argument; '$\xrightarrow{w}$' denotes weak convergence, '$\xrightarrow{p}$' convergence in probability, and '$\xrightarrow{w}_p$' weak convergence in probability, in each case as the sample size diverges to positive infinity; '$1(.)$' denotes the indicator function, '$x := y$' ('$x =: y$') indicates that x is defined by y (y is defined by x), and $x \overset{a}{=} y$ indicates that the objects x and y are asymptotically equivalent; finally, $\mathcal{C} := C[0,1]$ is used to denote the space of continuous processes on $[0,1]$, and $\mathcal{D} := D[0,1]$ the space of right continuous with left limit (càdlàg) processes on $[0,1]$.

2 The Heteroskedastic Trend Break Model

We consider the time series process $\{y_t\}$ generated according to the following model,

$$\begin{aligned} y_t &= \alpha_0 + \beta_0 t + \gamma_0 DT_t(\tau_0) + u_t, \quad t = 1, ..., T, & (1) \\ u_t &= \rho_T u_{t-1} + \varepsilon_t, \quad t = 2, ..., T, & (2) \\ \varepsilon_t &= C(L) e_t = \sum_{j=0}^{\infty} c_j e_{t-j} & (3) \\ e_t &= \sigma_t z_t & (4) \end{aligned}$$

and where the initial condition in (2) is assumed to satisfy $T^{-1/2} u_1 \xrightarrow{p} 0$. The observation equation in (1) allows for the possibility of a trend break through the indicator variable $DT_t(\tau_0) := 1(t > \lfloor \tau_0 T \rfloor)(t - \lfloor \tau_0 T \rfloor)$. The potential trend break point is given by $\lfloor \tau_0 T \rfloor$, with associated break fraction τ_0 and trend break magnitude γ_0, the latter assumed to be fixed.

We assume τ_0 is unknown but satisfies $\tau_0 \in \Lambda$, 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. It is necessary to impose the conditions that $\tau_L > 0$ and $\tau_U < 1$ because our approach is based around an estimator of the unknown break date, τ_0, and, as such, our search set needs to be bounded away from the end points of the sample; however, as will subsequently be shown, neither the asymptotic null distribution nor the local asymptotic power functions of the resulting unit root statistic depend on the choices of τ_L and τ_U, beyond the requirement that $\tau_L \leq \tau_0 \leq \tau_U$. In (1), a break in trend occurs at time $\lfloor \tau_0 T \rfloor$ when $\gamma_0 \neq 0$, while if $\gamma_0 = 0$, no break in trend occurs. One might also consider a second model which allows for a simultaneous break in the level of the process at time $\lfloor \tau_0 T \rfloor$ in the model in (1)-(4). However, as argued by Perron and Rodríguez (2003, pp.2,4), we need not analyze this case separately because a change in intercept is a special case of a slowly evolving deterministic component (see Condition B of ERS, p.816) and, consequently, does not alter any of the large sample results presented in this paper.

Within (2), we focus on the near-integrated autoregressive model, $H_c : \rho_T := 1 - c/T$ with $0 \leq c < \infty$, and we will be concerned with testing the unit root null hypothesis, $H_0 : c = 0$, against local alternatives, H_c where $c > 0$. The following set of assumptions will also be taken to hold on (1)-(4).

Assumption $\mathcal{A}$: $\mathcal{A}_1$. *The lag polynomial satisfies* $C(z) \neq 0$ *for all* $|z| \leq 1$, *and* $\sum_{j=0}^{\infty} j|c_j| < \infty$; $\mathcal{A}_2$. $z_t \sim IID(0,1)$ *with* $E|z_t|^r < K < \infty$ *for some* $r \geq 4$; $\mathcal{A}_3$. *The volatility term* σ_t *satisfies* $\sigma_t = \omega(t/T)$, *where* $\omega(\cdot) \in \mathcal{D}$ *is non-stochastic and strictly positive. For* $t < 0$, $\sigma_t \leq \breve{\sigma} < \infty$.

Remark 2.1 Notice that $\{\varepsilon_t\}$ in (3) is a linear process in $\{e_t\}$, the latter formed as the product of two components, $\{z_t\}$ and $\{\sigma_t\}$. Since, under Assumption $\mathcal{A}_2$, $\{z_t\}$ is IID, conditionally on σ_t, the error term e_t has mean zero and time-varying variance σ_t^2. $\square$

Assumption $\mathcal{A}$ coincides with the set of conditions adopted in Cavaliere and Taylor (2008). As discussed there, Assumption $\mathcal{A}_1$ is standard in the unit root literature. Assumption $\mathcal{A}_2$ is somewhat stronger than is often seen, since it rules out certain forms of conditional heteroskedasticity, such as that arising from stationary GARCH models, in the errors. This assumption is made to simplify exposition and the results stated in this paper would continue to hold if this assumption was weakened along the lines detailed in Remark 1 of Cavaliere and Taylor (2008,pp.46-47). The key assumption for the purposes of this paper, however, is $\mathcal{A}_3$, which casts the dynamics of the innovation variance in a quite general framework, requiring the innovation variance only to be non-stochastic, bounded and to display a countable number of jumps. A detailed discussion of the class of variance processes allowed under $\mathcal{A}_3$ is given in Cavaliere and Taylor (2007). They show that this includes the single abrupt change model of, *inter alia*, Kim *et al.* (2002), variance processes displaying multiple volatility shifts, polynomially (possibly piecewise) trending volatility and smooth transition variance breaks, among other things. In the case where volatility displays jumps, these are not constrained to be located at the same point in the sample as the putative trend break, nor indeed are they required to lie within the search set, Λ. The conventional homoskedasticity assumption, as employed in HHLT, that $\sigma_t = \sigma$ for all t, also satisfies Assumption $\mathcal{A}_3$, since here $\omega(s) = \sigma$ for all s. Although Assumption $\mathcal{A}_3$ imposes the volatility process to be non-stochastic, this may be weakened along the same lines as are detailed in Remark 2 of Cavaliere and Taylor (2008,p.47).

A quantity which will play a key role in what follows is given by the following function in $\mathcal{C}$, known as the *variance profile* of the process:

$$\eta(s) := \left(\int_0^1 \omega(r)^2 \, dr \right)^{-1} \int_0^s \omega(r)^2 \, dr. \tag{5}$$

Observe that the variance profile satisfies $\eta(s) = s$ under homoskedasticity while it deviates from s in the presence of heteroskedasticity. Notice also that the quantity $\overline{\omega}^2 := \int_0^1 \omega(r)^2 \, dr$ which appears in (5), by Assumption $\mathcal{A}_3$ equals the limit of $T^{-1} \sum_{t=1}^T \sigma_t^2$, and may therefore be interpreted as the (asymptotic) average innovation variance.

Remark 2.2 Before progressing it is worth commenting that, since the variance σ_t^2 depends on T, a time series generated according to (1)-(4) with σ_t satisfying Assumption $\mathcal{A}_3$ formally constitutes a triangular array of the type $\{y_{T,t} := d_{T,t} + u_{T,t} : 1 \leq t \leq T, T \geq 2\}$, where $d_{T,t}$ is purely deterministic and $u_{T,t}$ is recursively defined as $u_{T,t} := \rho_T u_{T,t-1} + C(L)\,\sigma_{T,t} z_t$, $\sigma_{T,\lfloor sT \rfloor} := \omega(s)$. However, since the triangular array notation is not essential, for simplicity the subscript T will be suppressed in what follows. $\square$

3 Break Fraction Estimation and Unit Root Tests

In section 3.1 we first provide a brief review of the modified trend break fraction estimator proposed in HHLT together with their proposed ADF-based unit root testing procedure based on this estimator. We also show how their approach can be extended to the class of $\mathcal{M}$-type unit root tests. In section 3.2 we then derive the large sample behaviour of these quantities under non-stationary volatility which satisfies Assumption $\mathcal{A}_3$. Specifically, we first show that HHLT's trend break fraction estimator retains the same rates of consistency as demonstrated by HHLT for the homoskedastic case in cases both where a trend break occurs and where it does not. However, as we subsequently demonstrate, unit root statistics based around this estimator do not attain the same limiting distribution (either under the null or the local alternative) as obtain in the homoskedastic case, either where a trend break occurs or where it does not. We then quantify the impact of a one-time break in volatility on the asymptotic size and local power of the unit root tests via Monte Carlo simulation.

3.1 The HHLT Procedure - Homoskedastic Innovations

Under the assumption of homoskedastic innovations, that is $\sigma_t = \sigma$ for all t, HHLT develop a unit root test based on QD detrending that has both correct asymptotic size under the unit root null hypothesis and is asymptotically efficient against local alternatives, regardless of whether or not a break in trend occurs.

In order to carry out valid unit root inference in the case where a trend break is known to have occurred at some unknown point in the sample (that is where $\gamma_0 \neq 0$), an estimator of the unknown break fraction is required whose rate of consistency is sufficiently rapid for a unit root test based on that estimator to have an asymptotic distribution that is the same as if the break fraction τ_0 were known, such that the (asymptotic) critical values corresponding to a known break fraction can be used and that the test will be asymptotically efficient. This requires that the estimator obtains a rate of consistency which is faster than $O_p(T^{-1/2})$.

HHLT consider the *first difference* estimator of τ_0:

$$\tilde{\tau} := \arg\min_{\tau \in \Lambda} \tilde{\sigma}^2(\tau), \tag{6}$$

where $\tilde{\sigma}^2(\tau) := T^{-1}\sum_{t=2}^{T} \tilde{v}_t(\tau)^2$, and $\tilde{v}_t(\tau)$ are the OLS residuals from the regression

$$\Delta y_t = \beta_0 + \gamma_0 DU_t(\tau) + v_t, \quad t = 2, ..., T \tag{7}$$

where $v_t := \Delta u_t$ and $DU_t(\tau) := 1(t > \lfloor \tau T \rfloor)$. Under homoskedasticity, HHLT [Lemma 1] demonstrate that this estimator satisfies $\tilde{\tau} = \tau_0 + O_p(T^{-1})$, so that it is consistent at rate $O_p(T^{-1})$.

Applying a QD transformation to (1) at the estimated break fraction, $\tilde{\tau}$ of (6), yields the QD de-trended data

$$\tilde{u}_t := y_t - X_t(\tilde{\tau})' \hat{\theta}_{\bar{c}} \tag{8}$$

where $X_t(\tilde{\tau}) := (1,\, t,\, DT_t(\tilde{\tau}))'$ and $\hat{\theta}_{\bar{c}}$ is the vector of OLS parameter estimates from the regression of $y_{\bar{c},t}$ on $X_{\bar{c},t}(\tilde{\tau})$, where

$$y_{\bar{c},t} := \begin{cases} y_1 & t = 1 \\ y_t - \bar{\rho}_T y_{t-1} & t = 2, \ldots, T \end{cases},$$

$$X_{\bar{c},t}(\tilde{\tau}) := \begin{cases} X_1(\tilde{\tau}) & t = 1 \\ X_t(\tilde{\tau}) - \bar{\rho}_T X_{t-1}(\tilde{\tau}) & t = 2, \ldots, T \end{cases}$$

and $\bar{\rho}_T := 1 - \bar{c}/T$, where $\bar{c}$ is the QD parameter. This parameter is generally chosen to be the value of the local-to-unity parameter c at which the asymptotic Gaussian local power envelope for a given significance level has power equal to 0.50. Table 1 of HHLT reports relevant values of $\bar{c}$ for the true break fractions $\tau_0 \in \{0.15, 0.20, ..., 0.85\}$ at the nominal 0.10, 0.05 and 0.01 significance levels; these are reproduced in Table 1, along with additional $\bar{c}$ values for $\tau_0 \in \{0.05, 0.10\}$ and $\tau_0 \in \{0.90, 0.95\}$.

The QD detrended ADF test rejects for large negative values of the regression t-statistic for $\phi = 0$ in the ADF-type regression

$$\Delta\tilde{u}_t = \phi\tilde{u}_{t-1} + \sum_{j=1}^{p} \delta_{j,p}\Delta\tilde{u}_{t-j} + e_{p,t}, \quad t = p+2, ..., T. \tag{9}$$

We denote this statistic $\textit{ADF-GLS}^{tb}(\tilde{\tau}, \bar{c})$. As is standard, the lag truncation parameter, p, in (9) is assumed to satisfy the following condition:

Assumption $\mathcal{B}$. *As* $T \to \infty$, *the lag truncation parameter* p *in (9) satisfies the condition that* $1/p + p^3/T \to 0$.

Remark 3.1 Perron and Rodríguez (2003) recommend using the modified Akaike Information Criterion (MAIC) of Ng and Perron (2001) for selecting p in (9) with an upper bound $p_{\max}$ satisfying Assumption $\mathcal{B}$; see section 6 of Perron and Rodríguez (2003). □

Under homoskedasticity, HHLT [Theorem 1] show that under H_c, $\textit{ADF-GLS}^{tb}(\tilde{\tau}, \bar{c}) = \textit{ADF-GLS}^{tb}(\tau_0, \bar{c}) + o_p(1)$, where $\textit{ADF-GLS}^{tb}(\tau_0, \bar{c})$ denotes the ADF-type unit root statistic of Perron and Rodríguez (2003) evaluated at the true break date fraction, τ_0. This result with $c = 0$ shows that one can carry out the test $\textit{ADF-GLS}^{tb}(\tilde{\tau}, \bar{c})$ by using asymptotic critical values appropriate for $\textit{ADF-GLS}^{tb}(\tau_0, \bar{c})$. For $c > 0$, the result establishes that $\textit{ADF-GLS}^{tb}(\tilde{\tau}, \bar{c})$ and $\textit{ADF-GLS}^{tb}(\tau_0, \bar{c})$ have identical asymptotic local alternative power functions which both lie arbitrarily close to the Gaussian local power envelope; see Perron and Rodríguez (2003,pp.7-9). The asymptotic distribution of $\textit{ADF-GLS}^{tb}(\tau_0, \bar{c})$ under H_c is given in Theorem 1 of Perron and Rodríguez (2003,p.5), and critical values from this distribution for the grid of break fractions $\tau_0 \in \{0.05, 0.10, ..., 0.95\}$ are provided in Table 1 for the nominal 0.10, 0.05 and 0.01 significance levels. Here and throughout the paper, the asymptotic results we report were obtained from direct simulation of the relevant limiting distributions, approximating the Wiener processes using NIID(0,1) random variates, with the integrals approximated by normalized sums of 1000 steps. The simulations were conducted using the `rndKMn` function of Gauss 7.0 with 50,000 replications.

Where no trend break occurs (so that $\gamma_0 = 0$), $\tilde{\tau}$ is randomly distributed on Λ and consequently, as shown in Perron and Rodríguez (2003) and HHLT, $\textit{ADF-GLS}^{tb}(\tilde{\tau}, \bar{c})$ is asymptotically over-sized and is no longer efficient in that it always includes an unnecessary trend break dummy. Where it is known that $\gamma_0 = 0$, it is the test based on the standard QD detrended (constant plus linear trend) augmented Dickey-Fuller unit root test statistic of ERS, which we denote by $\textit{ADF-GLS}^{t}$, that is efficient.[2] As a consequence, HHLT propose modifying the first difference break fraction estimator so that it has the same asymptotic

[2] This statistic is formed in the same way as outlined above for $\textit{ADF-GLS}^{tb}(\tilde{\tau}, \bar{c})$ but replacing $\tilde{u}_t$ in (9)

properties as previously discussed when a break occurs, but converges to zero when a break does not occur, so that as a consequence the trend break dummy drops out from the QD detrending stage. This ensures that (under homoskedasticity) the resulting test has correct asymptotic size (using the asymptotic critical values from Table 1 of ERS, p.825) and is asymptotically efficient in the case where a break does not occur.

Specifically, HHLT propose the modified break fraction estimator

$$\bar{\tau} = [1 - \exp(-gT^{-1/2}W_T(\tilde{\tau}))]\tilde{\tau} \tag{10}$$

where g is an (asymptotically irrelevant) positive constant,[3] and $W_T(\tilde{\tau})$ denotes the Wald statistic for testing $H_0 : \gamma_0 = 0$ in the partially summed model $\sum_{i=1}^{t} y_i = \alpha_0 t + \beta_0 \sum_{i=1}^{t} i + \gamma_0 \sum_{i=1}^{t} DT_i(\tilde{\tau}) + s_t$, where $s_t := \sum_{i=1}^{t} u_i$, $t = 1, ..., T$. HHLT [Lemmas 2 and 3] demonstrate that if $\gamma_0 = 0$, $W_T(\tilde{\tau}) = O_p(1)$, and $\bar{\tau} = O_p(T^{-1/2})$, while if $\gamma_0 \neq 0$, $W_T(\tilde{\tau}) = O_p(T)$, and $\bar{\tau} = \tau_0 + O_p(T^{-1})$, giving the required properties.

Since the earliest a break can occur is τ_L, the HHLT statistic is given by[4]

$$t(\bar{\tau}) := \begin{cases} ADF\text{-}GLS^{\,t} & \text{if } \bar{\tau} < \tau_L \\ ADF\text{-}GLS^{\,tb}(\bar{\tau}, \bar{c}) & \text{if } \bar{\tau} \geq \tau_L. \end{cases} \tag{11}$$

The associated unit root test uses the decision rule as outlined previously for $ADF\text{-}GLS^{\,t}$ and $ADF\text{-}GLS^{\,tb}(\bar{\tau}, \bar{c})$ as appropriate to the unit root statistic selected in (11). HHLT [Theorem 2] demonstrate that if $\gamma_0 = 0$, $t(\bar{\tau}) = ADF\text{-}GLS^{t} + o_p(1)$, while if $\gamma_0 \neq 0$, $t(\bar{\tau}) = ADF\text{-}GLS^{tb}(\tau_0, \bar{c}) + o_p(1)$. Consequently, under the assumption of homoskedastic innovations, a unit root test based on $t(\bar{\tau})$ will be asymptotically correctly sized and efficient, regardless of whether or not a break occurs.

The approach proposed in HHLT uses ADF-type tests in forming the unit root procedure in (11). However, the same testing principle could equally well be implemented using the $\mathcal{M}$ unit root tests of Perron and Ng (1996), Stock (1999), Ng and Perron (2001), and for the broken trend case, Perron and Rodríguez (2003). In the linear trend environment unit root tests based on the $\mathcal{M}$ principle are generally considered to be significantly more robust to serially correlated shocks than ADF-based tests, particularly where the MAIC lag truncation rule of Ng and Perron (2001) is adopted; see, for example, Haldrup and Jansson (2006). It

by the QD de-trended data $y_t - X_t'\hat{\theta}_{\bar{c}}$ where $X_t := (1, t)'$ and $\hat{\theta}_{\bar{c}}$ is now the vector of OLS estimates from the regression of $y_{\bar{c},t}$ on $X_{\bar{c},t}$ where $X_{\bar{c},t} := X_1$ for $t = 1$ and $X_{\bar{c},t} := X_t - \bar{\rho}_T X_{t-1}$ for $t = 2, \ldots, T$. Here $\bar{c} = 13.5$ for tests run at the asymptotic 5% level; see ERS.

[3] Although asymptotically irrelevant, g, does have an impact in finite samples. Recommendations on the choice of g in practice are given in section 5.3 of HHLT and will be further discussed in section 5 of this paper.

[4] As discussed in HHLT, an asymptotically equivalent statistic can be formed by replacing $ADF\text{-}GLS^{\,tb}(\bar{\tau}, \bar{c})$ with $ADF\text{-}GLS^{\,tb}(\tilde{\tau}, \bar{c})$ in the definition of $t(\bar{\tau})$ when $\bar{\tau} \geq \tau_L$. However, HHLT note that unreported simulation experiments indicate that this does not improve upon the finite sample performance of $t(\bar{\tau})$. We found the same to be true for the corresponding $\mathcal{M}$-type tests discussed below.

therefore seems worthwhile to also explore variants of the HHLT procedure in (11) based on $\mathcal{M}$-type tests.

Consider first the case where a break occurs, so that $\gamma_0 \neq 0$. Here the trinity of $\mathcal{M}$ statistics evaluated at $\tilde{\tau}$ of (6), analogous to $ADF\text{-}GLS^{tb}(\tilde{\tau}, \bar{c})$, are given by

$$\begin{aligned} \mathcal{MZ}_{\alpha}^{tb}(\tilde{\tau}, \bar{c}) &:= \frac{T^{-1}\tilde{u}_T^2 - s_{AR}^2(p)}{2T^{-2}\sum_{t=2}^{T}\tilde{u}_{t-1}^2}, \quad \mathcal{MSB}^{tb}(\tilde{\tau}, \bar{c}) := \left(T^{-2}\sum_{t=2}^{T}\tilde{u}_{t-1}^2/s_{AR}^2(p)\right)^{1/2} \\ \mathcal{MZ}_{t}^{tb}(\tilde{\tau}, \bar{c}) &:= \mathcal{MZ}_{\alpha}^{tb}(\tilde{\tau}, \bar{c}) \times \mathcal{MSB}^{tb}(\tilde{\tau}, \bar{c}) \end{aligned} \tag{12}$$

where $s_{AR}^2(p)$ is an autoregressive estimator of the (non-normalized) spectral density at frequency zero of $\{\varepsilon_t\}$. Specifically,

$$s_{AR}^2(p) := \hat{\sigma}^2/(1-\widehat{\delta}(1))^2, \;\; \widehat{\delta}(1) := \sum_{i=1}^{p}\widehat{\delta}_{i,p} \tag{13}$$

where $\widehat{\delta}_{i,p}$, $i = 1, ..., p$, and $\hat{\sigma}^2$ are, respectively, the OLS slope and variance estimators from (9), and where p satisfies Assumption $\mathcal{B}$. In the case of $\mathcal{MZ}_{\alpha}^{tb}(\tilde{\tau}, \bar{c})$ and $\mathcal{MZ}_{t}^{tb}(\tilde{\tau}, \bar{c})$, the unit root null hypothesis is rejected for large negative values of the statistics, while a test based on $\mathcal{MSB}^{tb}(\tilde{\tau}, \bar{c})$ rejects for small values of the statistic. It is entirely straightforward to show, using the results established in HHLT and under the conditions stated in HHLT (which, recall, are the same as given in this paper but with the additional assumption of homoskedasticity, such that $\sigma_t = \sigma$ for all t) that under H_c, $\mathcal{M}^{tb}(\bar{\tau}, \bar{c}) = \mathcal{M}^{tb}(\tau_0, \bar{c}) + o_p(1)$, where $(\mathcal{M}^{tb}(\bar{\tau}, \bar{c}), \mathcal{M}^{tb}(\tau_0, \bar{c}))$ are used generically to denote either $(\mathcal{MZ}_{\alpha}^{tb}(\bar{\tau}, \bar{c}), \mathcal{MZ}_{\alpha}^{tb}(\tau_0, \bar{c}))$, $(\mathcal{MSB}^{tb}(\bar{\tau}, \bar{c}), \mathcal{MSB}^{tb}(\tau_0, \bar{c}))$ or $(\mathcal{MZ}_{t}^{tb}(\bar{\tau}, \bar{c}), \mathcal{MZ}_{t}^{tb}(\tau_0, \bar{c}))$, with $\mathcal{MZ}_{\alpha}^{tb}(\tau_0, \bar{c})$, $\mathcal{MSB}^{tb}(\tau_0, \bar{c})$ and $\mathcal{MZ}_{t}^{tb}(\tau_0, \bar{c})$ denoting the trinity of $\mathcal{M}$ test statistics of Perron and Rodríguez (2003) evaluated at the true break date fraction, τ_0, the limiting distributions for which under H_c are given in Theorem 1 of Perron and Rodríguez (2003,p.5); asymptotic critical values for these statistics are also reported in Table 1. Again, therefore, the $\mathcal{M}$-type tests based on the statistics in (12) will be correctly sized in the limit (when using the appropriate asymptotic critical values from Table 1) and will be asymptotically efficient.

Where a break does not occur (such that $\gamma_0 = 0$), just like the test based on $ADF\text{-}GLS^{tb}(\tilde{\tau}, \bar{c})$, the tests based on the $\mathcal{M}$ statistics in (12) will be neither correctly sized nor efficient, even in the limit, due again to the random behaviour of $\tilde{\tau}$ in the no break case. Again here to obtain correctly sized efficient tests one needs to construct the $\mathcal{M}$ statistics in (12) from data which are QD detrended allowing simply for a constant plus linear trend, as outlined in footnote 2. Denote the resulting $\mathcal{M}$ statistics by $\mathcal{MZ}_{\alpha}^{t}$, $\mathcal{MSB}^{t}$ and $\mathcal{MZ}_{t}^{t}$.

Using the modified break estimator, $\bar{\tau}$ of (10), $\mathcal{M}$-based analogues of $t(\bar{\tau})$ of (11) can be formed as

$$\mathcal{M}(\bar{\tau}) := \begin{cases} \mathcal{M}^{t} & \text{if } \bar{\tau} < \tau_L \\ \mathcal{M}^{tb}(\bar{\tau}, \bar{c}) & \text{if } \bar{\tau} \geq \tau_L. \end{cases} \tag{14}$$

where $\mathcal{M}^t$ is used in a generic sense, in the same way as $\mathcal{M}^{tb}(\bar{\tau}, \bar{c})$, to denote either $\mathcal{MZ}_\alpha^t$, $\mathcal{MSB}^t$ or $\mathcal{MZ}_t^t$. With an obvious notation, we denote these three statistics by $\mathcal{MZ}_\alpha(\bar{\tau})$, $\mathcal{MSB}(\bar{\tau})$ and $\mathcal{MZ}_t(\bar{\tau})$. Again using results established in HHLT and under the conditions stated in HHLT it is straightforward to demonstrate that, in the generic notation used above, if $\gamma_0 = 0$, $\mathcal{M}(\bar{\tau}) = \mathcal{M}^t + o_p(1)$, while if $\gamma_0 \neq 0$, $\mathcal{M}(\bar{\tau}) = \mathcal{M}^{tb}(\tau_0, \bar{c}) + o_p(1)$. Consequently, under homoskedastic innovations, unit root tests based on $\mathcal{M}(\bar{\tau})$ will again be asymptotically correctly sized and efficient, regardless of whether or not a break occurs.

We now turn to establishing the large sample behaviour of both the first difference and modified estimators, $\tilde{\tau}$ of (6) and $\bar{\tau}$ of (10) respectively, and the associated unit root test based on $t(\bar{\tau})$ of (11), along with the corresponding $\mathcal{M}$-based tests in (14), when the innovations display non-stationary volatility of the form considered in Assumption $\mathcal{A}_3$.

3.2 Behaviour of the HHLT Procedure under Non-stationary Volatility

In this section we first discuss the behaviour of the break fraction estimators $\tilde{\tau}$ of (6) and $\bar{\tau}$ of (10) in cases where the volatility process σ_t is permitted to be generated by any process satisfying Assumption $\mathcal{A}_3$. These results are detailed in the following lemma.

Lemma 1 *Let y_t be generated according to (1)-(4) under H_c, and let Assumption $\mathcal{A}$ hold. Then: (i) where $\gamma_0 \neq 0$, $\tilde{\tau} = \tau_0 + O_p(T^{-1})$, and $\bar{\tau} = \tau_0 + O_p(T^{-1})$; (ii) where $\gamma_0 = 0$, $\tilde{\tau}$ has a well defined asymptotic distribution with support equal to $\Lambda = [\tau_L, \tau_U]$, while $\bar{\tau} = O_p(T^{-1/2})$.*

The results of Lemma 1 (which also hold in the stable autoregressive case) establish that both the first difference estimator, $\tilde{\tau}$ of (6), and HHLT's modified version thereof, $\bar{\tau}$ of (10), behave exactly the same in large samples in the presence of non-stationary volatility satisfying Assumption $\mathcal{A}_3$ as they do under the constant volatility assumption of HHLT. The key implication of this is that $\bar{\tau}$ converges in probability to zero at rate $O_p(T^{-1/2})$ when there is no break in trend, but is consistent for the true break fraction, τ_0, at rate $O_p(T^{-1})$ when a break occurs. This result is crucial to our being able to employ wild bootstrap-based implementations of the tests outlined in section 3.1, as we shall see later.

Although, as Lemma 1 demonstrates, the large sample behaviour of the break fraction estimators $\tilde{\tau}$ and $\bar{\tau}$ is invariant to the presence or otherwise of non-stationary volatility in the innovations, we now show in Theorem 1 that the same is not true for the limiting distributions of the unit root tests from section 3.1.

Theorem 1 *Let y_t be generated according to (1)-(4) under H_c, and let Assumptions $\mathcal{A}$ and*

$\mathcal{B}$ *hold. Then: (i) if* $\gamma_0 = 0$,

$$\begin{aligned}
\mathcal{MZ}_\alpha(\bar{\tau}) &\stackrel{w}{\to} \frac{K_c^\eta(1)^2 - 1}{2\left(\int_0^1 K_c^\eta(r)^2 dr\right)} =: \bar{\xi}_1^{c,\bar{c},t,\eta} \\
\mathcal{MSB}(\bar{\tau}) &\stackrel{w}{\to} \left(\int_0^1 K_c^\eta(r)^2 dr\right)^{1/2} =: \bar{\xi}_2^{c,\bar{c},t,\eta} \\
\mathcal{MZ}_t(\bar{\tau}) \stackrel{a}{=} t(\bar{\tau}) &\stackrel{w}{\to} \frac{K_c^\eta(1)^2 - 1}{2\left(\int_0^1 K_c^\eta(r)^2 dr\right)^{1/2}} =: \bar{\xi}_3^{c,\bar{c},t,\eta}
\end{aligned}$$

and, (ii) if $\gamma_0 \neq 0$,

$$\begin{aligned}
\mathcal{MZ}_\alpha(\bar{\tau}) &\stackrel{w}{\to} \frac{V_1^\eta(1,\tau_0)^2 - 2V_2^\eta(1,\tau_0) - 1}{2\left(\int_0^1 V_1^\eta(r,\tau_0)^2 dr - 2\int_{\tau_0}^1 V_2^\eta(r,\tau_0) dr\right)} =: \bar{\xi}_1^{c,\bar{c},tb,\eta}(\tau_0) \\
\mathcal{MSB}(\bar{\tau}) &\stackrel{w}{\to} \left(\int_0^1 V_1^\eta(r,\tau_0)^2 dr - 2\int_{\tau_0}^1 V_2^\eta(r,\tau_0) dr\right)^{1/2} =: \bar{\xi}_2^{c,\bar{c},tb,\eta}(\tau_0) \\
\mathcal{MZ}_\alpha^t(\bar{\tau}) \stackrel{a}{=} t(\bar{\tau}) &\stackrel{w}{\to} \frac{V_1^\eta(1,\tau_0)^2 - 2V_2^\eta(1,\tau_0) - 1}{2\left(\int_0^1 V_1^\eta(r,\tau_0)^2 dr - 2\int_{\tau_0}^1 V_2^\eta(r,\tau_0) dr\right)^{1/2}} =: \bar{\xi}_3^{c,\bar{c},tb,\eta}(\tau_0)
\end{aligned}$$

where

$$\begin{aligned}
K_c^\eta(r) &:= W_c^\eta(r) - r\left\{\bar{c}^* W_c^\eta(1) + 3(1-\bar{c}^*)\int_0^1 r W_c^\eta(r) dr\right\} \\
V_1^\eta(r,\tau_0) &:= W_c^\eta(r) - r b_3 \\
V_2^\eta(r,\tau_0) &:= b_4(r-\tau_0)\left\{W_c^\eta(r) - r b_3 - \frac{1}{2}(r-\tau_0) b_4\right\}
\end{aligned}$$

with $W_c^\eta(r) := \int_0^r e^{-(r-s)c} dW(\eta(s))$, *where* $W(s)$ *is a standard Brownian motion and* $\eta(\cdot)$ *is the variance profile of the volatility process defined in (5),* $\bar{c}^* := (1+\bar{c})/(1+\bar{c}+\bar{c}^2/3)$, $b_3 := \lambda_1 b_1 + \lambda_2 b_2$, $b_4 := \lambda_2 b_1 + \lambda_3 b_2$ *with* $b_1 := (1+\bar{c})W_c^\eta(1) + \bar{c}^2 \int_0^1 r W_c^\eta(r) dr$ *and* $b_2 := (1+\bar{c}-\tau_0\bar{c})W_c^\eta(1) + \bar{c}^2 \int_{\tau_0}^1 W_c^\eta(r)(r-\tau_0) dr - W_c^\eta(\tau_0)$, *and the constants* $\lambda_1, ..., \lambda_3$ *are as defined in Theorem 1 of Perron and Rodríguez (2003,p.5) but replacing* $\bar{c}$ *by* $-\bar{c}$ *in their expressions.*

Remark 3.2 As can be seen from comparing the representations given in Theorem 1 with the corresponding representations in Theorem 1 of Perron and Rodríguez (2003,p.5) for the trend break case ($\gamma_0 \neq 0$) and ERS for the linear trend case ($\gamma_0 = 0$), the asymptotic null distributions (those pertaining to $c = 0$) of the unit root statistics have the usual form with the exception that the standard limiting Brownian motion, $W_0(s)$, is replaced by the Brownian functional $W_0^\eta(s)$, the latter reducing to the former only where the process is homoskedastic; that is, where $\omega(\cdot)$ is a constant function. The process $W_0^\eta(s)$ is known as a *variance-transformed* Brownian motion, i.e. a Brownian motion under a modification of the time domain; see, for example, Davidson (1994). This difference has serious implications for the size of the associated unit root tests, with the standard tabulated critical values (either those given in ERS or those provided in Table 1) no longer appropriate in either the linear or broken trend environments. □

Remark 3.3 It is also clear from the representations in Theorem 1 for $c > 0$ that the asymptotic local power functions of the tests in both the linear and broken trend cases will also be affected by non-stationary volatility (even if critical values from the correct null distributions were used) since, as with the null case, it is only where $\omega(\cdot)$ is constant that the representations given in parts (i) and (ii) of Theorem 1 reduce to the corresponding representations given in ERS and Perron and Rodríguez (2003) respectively. Specifically, the (size-adjusted) asymptotic local power functions of the tests, in each case at significance level ε, are given by $\pi_j(c, \bar{c}, t, \eta)$ and $\pi_j(c, \bar{c}, tb, \eta, \tau_0)$, $j = 1, 2, 3$, in the linear and broken trend cases, respectively, where

$$\pi_j(c, \bar{c}, \cdot, \eta) := P(\bar{\xi}_j^{c,\bar{c},\cdot,\eta} < b_j(\bar{c}, \cdot, \eta)), \;\; j = 1, 2, 3 \tag{15}$$

and

$$\pi_j(c, \bar{c}, \cdot, \eta, \cdot) := P(\bar{\xi}_j^{c,\bar{c},\cdot,\eta}(\cdot) < b_j(\bar{c}, \cdot, \eta, \cdot)), \;\; j = 1, 2, 3 \tag{16}$$

where $b_j(\bar{c}, \cdot, \eta)$ and $b_j(\bar{c}, \cdot, \eta, \cdot)$ respectively satisfy $P(\bar{\xi}_j^{0,\bar{c},\cdot,\eta} < b_j(\bar{c}, \cdot, \eta)) = \varepsilon$ and $P(\bar{\xi}_j^{0,\bar{c},\cdot,\eta}(\cdot) < b_j(\bar{c}, \cdot, \eta, \cdot)) = \varepsilon$, $j = 1, 2, 3$, where the index $j = 1$ refers to the test based on $\mathcal{MZ}_\alpha(\bar{\tau})$, $j = 2$ to that based on $\mathcal{MSB}(\bar{\tau})$, and $j = 3$ to both tests based on $\mathcal{MZ}_\alpha^t(\bar{\tau})$ and on $t(\bar{\tau})$. Moreover, unlike in the homoskedastic case, these cannot necessarily be expected to lie close to the Gaussian asymptotic local power envelope in either the linear or broken trend environments. □

Remark 3.4 As in HHLT, we have assumed thus far that, γ_0, the magnitude of the trend break in (1) is either zero or non-zero but independent of the sample size, T. It is straightforward, however, to show that the results in Lemma 1 and Theorem 1 can be extended to cover the so-called *shrinking break* case where the break magnitude is local to zero in the sample size; that is, the case where $\gamma_0 = \gamma_T := \kappa T^{-d}$, where $d \geq 0$ and κ is a non-zero constant. As noted in the proof of part (i) of Lemma 1, $\tilde{\tau} - \tau_0$ is of $O_p((T\gamma_T)^{-1})$ and, consequently, our requirement that $\tilde{\tau}$ be consistent for τ_0 at a rate faster than $O_p(T^{-1/2})$ is satisfied in the shrinking break case, provided $0 \leq d < 1/2$. In order to establish the properties of the modified break estimator, $\bar{\tau}$ of (10), in the shrinking break case we must also consider the behaviour of $W_T(\tilde{\tau})$ of (10) here. It is entirely straightforward to demonstrate from the results given in part (ii) of Lemma 1 that $W_T(\tilde{\tau})$ is of $O_p(T^{1-2d})$ here, so that this statistic will continue to diverge in the shrinking break case, again provided $0 \leq d < 1/2$. Given that $W_T(\tilde{\tau})$ is scaled by $T^{-1/2}$ in (10) this requires that $0 \leq d < 1/4$ for the results stated in Lemma 1 for $\bar{\tau}$, and hence the results stated in Theorem 1, to also hold in the shrinking break case. However, the $T^{-1/2}$ scaling used on $W_T(\tilde{\tau})$ in (10) is not essential and could be relaxed to $T^{d-1/2}$ in which case the results given in Lemma 1 for $\bar{\tau}$ together with the results stated in Theorem 1 would then also hold for any $0 \leq d < 1/2$. The results given for our proposed bootstrap statistics in Theorem 2, below, also extend in the same way to cover the shrinking break case. $\square$

To conclude this section we now quantify the impact of a one-time change in volatility on the asymptotic size and local power of the tests of section 3.1. Figure 1 reports the asymptotic size of nominal 0.05-level $\mathcal{M}(\bar{\tau})$ and $t(\bar{\tau})$ tests for a single abrupt shift in volatility from σ_0 to σ_1 at time $\lfloor \tau_\sigma T \rfloor$; i.e. for the volatility function

$$\omega(s) = \sigma_0 + (\sigma_1 - \sigma_0)1(s > \tau_\sigma), \quad s \in [0,1] \tag{17}$$

with $\tau_\sigma \in [0,1]$. Results are reported for $\sigma_1/\sigma_0 \in \{1/10, 1/9, ..., 1/2, 1, 2, 3, ..., 10\}$ and $\tau_\sigma \in \{0.3, 0.5, 0.7\}$, allowing for positive and negative breaks in volatility at a range of timings (the setting $\sigma_1/\sigma_0 = 1$ giving the constant volatility case). Four deterministic specifications are considered: (a) no break in trend, (b) break in trend with $\tau_0 = 0.3$, (c) break in trend with $\tau_0 = 0.5$, and (d) break in trend with $\tau_0 = 0.7$. Together with the settings for the volatility shifts, cases (b)-(d) allow for situations both where the shift in volatility coincides with the break in trend, and where the breaks occur at different points in time.

While the $\mathcal{M}(\bar{\tau})$ and $t(\bar{\tau})$ tests are correctly sized under constant volatility, a pattern of asymptotic over-size is clearly evident when a shift in volatility occurs. The size distortions increase with the magnitude of the shift in volatility, and can be quite severe, up to around 40% in the worst of the cases considered. The $\mathcal{MZ}_\alpha(\bar{\tau})$, $\mathcal{MSB}(\bar{\tau})$, and $\mathcal{MZ}_t(\bar{\tau}), t(\bar{\tau})$ tests are affected in different ways by the non-constant volatility, but these differences are nonetheless quite small, especially when considered in relation to the size distortions induced by the volatility shift. The most severe size biases occur when the timings of the break in trend and the break in volatility are either both early ($\tau_0 = \tau_\sigma = 0.3$) or both late ($\tau_0 = \tau_\sigma = 0.7$), although substantial size distortions are also observed for other combinations of break timings, and also when no break in trend is present. The impact of a volatility shift on test size is also seen to be strongly dependent on the direction of the shift: when the volatility break occurs early (late), a negative (positive) shift generates relatively greater size distortions than when the shift is positive (negative).

In Figures 2-4 the asymptotic size-corrected local power of the $\mathcal{MZ}_t(\bar{\tau}), t(\bar{\tau})$ tests are reported for the same four deterministic specifications as above.[5] Figures 2, 3 and 4 present powers when the volatility undergoes an abrupt shift, with, respectively, $\tau_\sigma = 0.3$, 0.5 and 0.7. In each case, we report results for constant volatility ($\sigma_1/\sigma_0 = 1$) as well as the changing

[5]The size-adjusted powers for the $\mathcal{MZ}_\alpha(\bar{\tau})$ and $\mathcal{MSB}(\bar{\tau})$ tests were found to be almost identical to those of $\mathcal{MZ}_t(\bar{\tau}), t(\bar{\tau})$ and are consequently omitted.

volatility settings $\sigma_1/\sigma_0 \in \{1/10, 1/5, 5, 10\}$. Relative to the baseline case of homoskedasticity, it can be seen from Figures 2-4 that a shift in volatility can have a substantial and detrimental impact on the size-adjusted power of the unit root tests. Unsurprisingly, the cases where power is most dramatically affected are those where size was found to be most distorted, i.e. early negative and late positive volatility shifts, where the timing of the breaks in trend and volatility coincide.

4 Bootstrap Unit Root Tests

As demonstrated in the previous section, non-stationary volatility introduces a time deformation aspect to the limiting distributions of the unit root statistics under H_c, which alters their form *vis-à-vis* the homoskedastic case. In section 4.1 we propose bootstrap analogues of the unit root tests from section 3.1. In section 4.2 we show that the bootstrap tests allow us to retrieve asymptotically correct p-values under the null and share the same asymptotic local power functions as the corresponding (asymptotically size-adjusted) standard tests.

Our proposed bootstrap algorithm draws on the 'wild bootstrap' literature (see, *inter alia*, Liu, 1988, and Mammen, 1993) and, as we demonstrate, allows us to construct bootstrap unit root tests that are robust under the null to non-stationary volatility of the form given in Assumption $\mathcal{A}_3$. In the context of the present problem, the wild bootstrap scheme is required, rather than the standard residual re-sampling schemes used for other bootstrap unit root tests proposed in the literature because, unlike these, the wild bootstrap can replicate the pattern of heteroskedasticity present in the shocks; see Remark 4.1. The wild bootstrap has also recently been successfully applied by Gonçalves and Kilian (2004,2007) to the problem of carrying out inference in stationary autoregressive models driven by heterogenous shocks, and by Cavaliere and Taylor (2008) to construct bootstrap versions of the $\mathcal{MZ}_\alpha^t$, $\mathcal{MSB}^t$ and $\mathcal{MZ}_t^t$ tests of section 3.1 which are asymptotically robust under the unit root null to non-stationary volatility of the form described in Assumption $\mathcal{A}_3$.

4.1 The Bootstrap Algorithm

Implementation of the bootstrap is straightforward and requires only the OLS de-trended data based around the first difference break date estimate, $\tilde{\tau}$ of (6), which, crucially, we know from Lemma 1 to be T-consistent for the true break fraction τ_0 where a trend break occurs, even in the presence of non-stationary volatility. Although this estimator is randomly distributed over Λ when no break occurs, this has no impact on the behaviour of the resulting bootstrap tests. The following steps constitute our bootstrap algorithm:

Algorithm 1 (Wild Bootstrap Unit Root Tests)

Step (i) *OLS de-trend y_t under the null hypothesis using the first difference break fraction estimator, $\tilde{\tau}$ of (6); that is, $\hat{\varepsilon}_t := \Delta y_t - \hat{\beta}_0 - \hat{\gamma}_0 DU_t(\tilde{\tau})$, $t = 2, ..., T$, where $\hat{\beta}_0$ and $\hat{\gamma}_0$ are the OLS estimates of β_0 and γ_0, respectively, obtained from regressing Δy_t on a constant and $DU_t(\tilde{\tau})$, $t = 2, ..., T$.*[6]

Step (ii) *Generate T bootstrap innovations ε_t^*, $t = 1, ..., T$, according to the device $\varepsilon_t^* := \hat{\varepsilon}_t w_t$, $t = 2, ..., T$, and $\varepsilon_1^* := 0$, where $\{w_t\}_{t=1}^T$ denotes an independent $N(0, 1)$ sequence.*

Step (iii) *Construct the bootstrap sample as the partial sum process*[7]

$$y_t^* := \sum_{i=1}^{t} \varepsilon_i^*, \quad t = 1, ..., T. \tag{18}$$

Step (iv) *According to the value of the modified break fraction estimator, $\bar{\tau}$ of (10), obtained from the original sample data y_t, $t = 1, ..., T$, obtain the bootstrap test statistics, as follows:*

(i) if $\bar{\tau} < \tau_L$, $t(\bar{\tau})^* :=$ ADF-GLS^{t*}, $\mathcal{MZ}_\alpha(\bar{\tau})^* := \mathcal{MZ}_\alpha^{t*}$, $\mathcal{MSB}(\bar{\tau})^* := \mathcal{MSB}^{t*}$ and $\mathcal{MZ}_t(\bar{\tau})^* := \mathcal{MZ}_t^{t*}$;

(ii) if $\bar{\tau} \geq \tau_L$, $t(\bar{\tau})^* :=$ ADF-GLS$^{tb*}(\bar{\tau}, \bar{c})$, $\mathcal{MZ}_\alpha(\bar{\tau})^* := \mathcal{MZ}_\alpha^{tb*}(\bar{\tau}, \bar{c})$, $\mathcal{MSB}(\bar{\tau})^* := \mathcal{MSB}^{tb*}(\bar{\tau}, \bar{c})$ and $\mathcal{MZ}_t(\bar{\tau})^* := \mathcal{MZ}_t^{tb*}(\bar{\tau}, \bar{c})$,

where, for example, ADF-GLS$^{tb*}(\bar{\tau}, \bar{c})$ denotes the statistic ADF-GLS$^{tb}(\bar{\tau}, \bar{c})$ constructed as outlined in section 3.1, but applied to the bootstrap data, y_t^*, rather than the original data, y_t, and using the break fraction estimator $\bar{\tau}$ obtained from the original data, y_t, with the same obvious notational convention also used for the other bootstrap statistics.

Step (v) *Bootstrap p-values are computed as: $p_{1,T}^* := G_{1,T}^*(t(\bar{\tau}))$, $p_{2,T}^* := G_{2,T}^*(\mathcal{MZ}_\alpha(\bar{\tau}))$, $p_{3,T}^* := G_{3,T}^*(\mathcal{MSB}(\bar{\tau}))$, and $p_{4,T}^* := G_{4,T}^*(\mathcal{MZ}_t(\bar{\tau}))$, where $G_{j,T}^*(\cdot)$, $j = 1, ..., 4$, denotes the conditional (on the original sample data) cumulative density function (cdf) of $t(\bar{\tau})^*$,*

[6] An asymptotically equivalent procedure is to de-trend using the modified estimator, $\bar{\tau}$ of (10), such that we use the residuals from a regression of Δy_t on just a constant when $\bar{\tau} < \tau_L$ or, when $\bar{\tau} \geq \tau_L$, from the regression of Δy_t on a constant and $DU_t(\check{\tau})$, where $\check{\tau}$ can be either $\bar{\tau}$ or $\tilde{\tau}$ (cf. footnote 4). Unreported Monte Carlo simulations suggest no material difference between the finite sample size and power properties of these two approaches.

[7] Notice that we do not include the estimated deterministic component in constructing the bootstrap sample. Doing so, by replacing $\sum_{i=1}^t \varepsilon_i^*$ in (18) with either $\sum_{i=1}^t \{\varepsilon_i^* + \bar{\beta}_0\}$ when $\bar{\tau} < \tau_L$, where $\bar{\beta}_0 := (T-1)^{-1}(y_T - y_1)$, or $\sum_{i=1}^t \{\varepsilon_i^* + \hat{\beta}_0 + \hat{\gamma}_0 DU_t(\check{\tau})\}$ when $\bar{\tau} \geq \tau_L$, where $\check{\tau}$ can be either $\bar{\tau}$ or $\tilde{\tau}$ (cf. footnote 4), could be used as an alternative bootstrap scheme, with the bootstrap break fraction estimators, say $\tilde{\tau}^*$ and $\bar{\tau}^*$ with an obvious notation, then being estimated from the bootstrap sample and used in the subsequent steps of the algorithm in place of $\bar{\tau}$ and $\tilde{\tau}$, respectively. This would be asymptotically equivalent to the approach outlined in Algorithm 1. In unreported Monte Carlo simulations we found that this alternative approach did not result in superior finite sample properties overall, and so we have adopted the (computationally) simpler of the two.

$\mathcal{MZ}_\alpha(\bar{\tau})^*$, $\mathcal{MSB}(\bar{\tau})^*$ *and* $\mathcal{MZ}_t(\bar{\tau})^*$ *respectively. Notice, therefore, that bootstrap tests, run at the* ξ, $0 < \xi < 1$, *significance level, based on* $t(\bar{\tau})$, $\mathcal{MZ}_\alpha(\bar{\tau})$, $\mathcal{MSB}(\bar{\tau})$ *and* $\mathcal{MZ}_t(\bar{\tau})$ *are then defined such that they reject the unit root null hypothesis,* $H_0 : \rho_T = 1$ *if* $p^*_{j,T} < \xi$, $j = 1, ..., 4$, *respectively.*

Remark 4.1 Notice that the bootstrap innovations, ε^*_t, replicate the pattern of heteroskedasticity present in the original shocks since, conditionally on $\hat{\varepsilon}_t$, ε^*_t is independent over time with zero mean and variance $\hat{\varepsilon}^2_t$. The fact that the bootstrap sample has uncorrelated increments also means that the true bootstrap lag order, p^*_0 say, is zero and, consequently, the lag truncation used in the bootstrap analogue of (9), p^* say, need not equal p, the lag order used in constructing the original statistic. In what follows we set $p^* = 0$ (unless otherwise stated) although the results stated in Theorem 2 are valid for any p^* such that $p^*/T^{1/3} \to 0$ as $T \to \infty$. $\square$

Remark 4.2 As in Cavaliere and Taylor (2008), the unit root null is imposed on the re-sampling scheme used in step (iii) of Algorithm 1. Notice that the arguments made in Paparoditis and Politis (2003) as to why it is preferable on power grounds to not impose the null hypothesis when constructing the residuals to be used for re-sampling, do not apply here. This is because, conditionally on the original data, the bootstrap innovations, ε^*_t from step (ii) are serially uncorrelated, regardless of any serial correlation present in the fitted residuals $\hat{\varepsilon}_t$. As an aside, because wild bootstrap innovations are (conditionally) serially uncorrelated it is often the case that improved finite sample performance can be obtained by incorporating an additional re-colouring (or sieve) device in step (iii) of Algorithm 1; see Cavaliere and Taylor (2009) for details. We experimented with this, both where the underlying innovations, ε_t, were serially correlated and where they were not, and found that such re-colouring did not improve on the finite sample performance of the bootstrap tests outlined in Algorithm 1. $\square$

Remark 4.3 As discussed in Davidson and Flachaire (2008), in some cases improved accuracy can be obtained when using the wild bootstrap by replacing the Gaussian distribution used for generating the pseudo-data by an asymmetric distribution with $E(w_t) = 0$, $E\left(w^2_t\right) = 1$ and $E\left(w^3_t\right) = 1$ (Liu, 1988), two examples being Mammen's (1993) two-point distribution: $P(w_t = -0.5(\sqrt{5} - 1) = 0.5(\sqrt{5} + 1)/\sqrt{5} = p$, $P(w_t = 0.5(\sqrt{5} + 1)) = 1 - p$, and the Rademacher distribution: $P(w_t = 1) = 1/2 = P(w_t = -1)$. We found only minor differences between the finite sample properties of the bootstrap unit root tests based on the Gaussian or the Mammen or Rademacher distributions. $\square$

Remark 4.4 In practice the cdfs $G^*_{j,T}$, $j = 1, ..., 4$, will be unknown but can be approximated in the usual way through numerical simulation; see, *inter alia*, Hansen (1996) and Andrews and Buchinsky (2001). Using HHLT's $t(\bar{\tau})$ statistic to illustrate, this is done by generating M such (conditionally) independent bootstrap statistics, say $t_m(\bar{\tau})^*$, $m = 1, ..., M$, computed as for $t(\bar{\tau})^*$ above but from $y^*_{m,t} := \sum_{i=1}^t \varepsilon^*_{m,i}$, $t = 1, ..., T$, $\varepsilon^*_{m,t} := \hat{\varepsilon}_t w_{m:t}$, $t = 2, ..., T$, and $\varepsilon^*_{m,1} := 0$, with $\{\{w_{m:t}\}_{t=1}^T\}_{m=1}^M$ a doubly independent $N(0,1)$ sequence. The simulated bootstrap p-value is then computed as $\tilde{p}^*_{1,T} := M^{-1}\sum_{m=1}^M \mathbb{1}\left(t_m(\bar{\tau})^* \leq t(\bar{\tau})\right)$, and is such that $\tilde{p}^*_{1,T} \overset{a.s.}{\to} p^*_{1,T}$ as $M \to \infty$. An approximate standard error for $\tilde{p}^*_{1,T}$ is given by $(\tilde{p}^*_{1,T}(1-\tilde{p}^*_{1,T})/M)^{1/2}$; see Hansen (1996, p.419). $\square$

4.2 Asymptotic Properties

In this section we discuss the asymptotic properties of the wild bootstrap unit root tests. We show that under the unit root null, and for any volatility process satisfying Assumption $\mathcal{A}_3$, the bootstrap statistics converge to the same asymptotic distribution as the test statistics computed on the original data; that is, the wild bootstrap allows us to replicate the correct first-order asymptotic null distributions of the test statistics of interest. We also demonstrate the asymptotic properties of the bootstrap statistics under near-integrated alternatives.

In Theorem 2 we now show that the bootstrap statistics from section 4.1 allow us to retrieve the correct asymptotic null distributions for the tests from section 3.1, and, hence, that the corresponding p-values are asymptotically pivotal. This theorem is derived under the condition that the bootstrap lag truncation parameter, $p^* = 0$; cf. Remark 4.1:

Theorem 2 *Let y_t be generated according to (1)-(4) under H_c and let Assumption $\mathcal{A}$ hold. Let the bootstrap sample be generated as detailed in Algorithm 1 with $p^* = 0$. Then: (i) if $\gamma_0 = 0$, $\mathcal{MZ}_\alpha(\bar{\tau})^* \overset{w}{\to}_p \bar{\xi}_1^{0,\bar{c},t,\eta}$, $\mathcal{MSB}(\bar{\tau})^* \overset{w}{\to}_p \bar{\xi}_2^{0,\bar{c},t,\eta}$, $\mathcal{MZ}_t(\bar{\tau})^* \overset{w}{\to}_p \bar{\xi}_3^{0,\bar{c},t,\eta}$ and $t(\bar{\tau}) \overset{w}{\to}_p \bar{\xi}_3^{0,\bar{c},t,\eta}$; and (ii) if $\gamma_0 \neq 0$, $\mathcal{MZ}_\alpha(\bar{\tau})^* \overset{w}{\to}_p \bar{\xi}_1^{0,\bar{c},tb,\eta}(\tau_0)$, $\mathcal{MSB}(\bar{\tau})^* \overset{w}{\to}_p \bar{\xi}_2^{0,\bar{c},tb,\eta}(\tau_0)$, $\mathcal{MZ}_t(\bar{\tau})^* \overset{w}{\to}_p \bar{\xi}_3^{0,\bar{c},tb,\eta}(\tau_0)$ and $t(\bar{\tau}) \overset{w}{\to}_p \bar{\xi}_3^{0,\bar{c},tb,\eta}(\tau_0)$. Moreover, if $\rho_T = 1$, $p^*_{j,T} \overset{w}{\to} U[0,1]$, $j = 1, ..., 4$.*

Theorem 2 demonstrates the usefulness of the wild bootstrap: as the number of observations diverges, the bootstrapped statistics have the same null distribution as the corresponding original test statistic and, consequently, the bootstrap p-values are uniformly distributed under the null hypothesis, leading to tests with asymptotically correct size. As Theorem 2 also shows, these limiting distributions also hold for the bootstrapped statistics under local alternatives. This immediately implies that the bootstrap unit root tests will have the same asymptotic local power functions as the original unit root tests, provided the latter were run using critical values from the appropriate limiting null distribution; that is, the asymptotic local power functions given in (15) and (16). The asymptotic theory therefore predicts that the bootstrap tests should have power approximately equal to the size-adjusted power of the standard tests. An obvious and important implication of this result is that under homoskedasticity ($\sigma_t = \sigma$ for all t) the wild bootstrap tests will not suffer from any loss of (asymptotic local) efficiency, relative to the standard tests.

Figure 1 also reports results for the asymptotic size of the wild bootstrap tests $\mathcal{MZ}_\alpha(\bar{\tau})^*$, $\mathcal{MSB}(\bar{\tau})^*$, and $\mathcal{MZ}_t(\bar{\tau})^*, t(\bar{\tau})^*$. Since these tests are correctly sized asymptotically, the sizes in Figure 1 are equal to 0.05 for all of the deterministic and volatility settings considered. The contrast with the corresponding $\mathcal{MZ}_\alpha(\bar{\tau})$, $\mathcal{MSB}(\bar{\tau})$, and $\mathcal{MZ}_t(\bar{\tau}), t(\bar{\tau})$ tests is striking, and serves to highlight the asymptotic robustness of the wild bootstrap tests to non-constant volatility. In addition, the asymptotic local power of the bootstrap tests is identical to that of the original tests when the latter are size-adjusted, thus the asymptotic power curves presented in Figures 2-4 can also be interpreted as the power curves for $\mathcal{MZ}_\alpha(\bar{\tau})^*$, $\mathcal{MSB}(\bar{\tau})^*$, and $\mathcal{MZ}_t(\bar{\tau})^*, t(\bar{\tau})^*$. The impact of non-constant volatility on the power of the tests can again be seen; while reductions in power are observed relative to the baseline case of constant volatility, the wild bootstrap tests attain the same asymptotic local power as the (infeasible) size-corrected original tests.

5 Finite Sample Simulations

In this section we investigate the finite sample size and power properties of the standard $\mathcal{M}(\bar{\tau})$ and $t(\bar{\tau})$ tests (using nominal 0.05-level asymptotic critical values), along with their wild boostrap analogues $\mathcal{M}(\bar{\tau})^*$ and $t(\bar{\tau})^*$. In order to compute the test statistics, choices must be made for the parameter g in (10). For the ADF-based tests $t(\bar{\tau})$ and $t(\bar{\tau})^*$, we set $g = 3$, opting for the better sized of the two procedures recommended by HHLT. For the $\mathcal{M}(\bar{\tau})$ and $\mathcal{M}(\bar{\tau})^*$ tests, unreported simulations similar to those in HHLT suggested that the best finite sample size/power trade-off was obtained for $g = 6$, hence we use that value for each of the $\mathcal{M}$-based tests in what follows. We also need to specify the range of allowable break fractions, Λ. The main results reported in this section pertain to tests calculated for $\Lambda = [0.15, 0.85]$; that is, 15% trimming. This degree of trimming coincides with that proposed by Banerjee *et al.* (1992) and Andrews (1993) and, more recently, the trimming used in the simulation experiments of Perron and Rodríguez (2003), Kim and Perron (2009) and HHLT. The sensitivity of the tests' finite sample behaviour to the choice of Λ is, however, also investigated. Moreover, all of the results in this paper can be repeated for other choices of Λ and a suite of Gauss programs to do this is available from the authors on request.

When accounting for potentially autocorrelated innovations, the lag truncation parameter, p, used in the ADF-type regressions and the computation of $s^2_{AR}(p)$, is selected according to the MAIC lag selection rule of Ng and Perron (2001) using the modification suggested by Perron and Qu (2007) - hereafter denoted by p_{MAIC} - with an upper bound of $p_{\max} := \lfloor 12(T/100)^{1/4} \rfloor$. In the implementation of the bootstrap algorithm, we set $p^* = 0$; cf. Remark 4.1.[8] All the simulations reported in this section were conducted using 10,000 Monte Carlo

[8] We also experimented with setting the value of p^* used in calculating the bootstrap statistics equal to that selected by MAIC for the corresponding original test statistic, and also re-selecting p^* according to MAIC for each bootstrap sample. However, we found that neither of these approaches resulted in a superior size/power trade-off across the range of DGPs considered, even where re-colouring was employed as outlined in Remark

replications and $M = 499$ bootstrap replications, again using the `rndKMn` function of Gauss 7.0.

Figure 5 reports finite sample size results for data generated according to (1)-(4) with $c = 0$ for the shift in volatility model where σ_t shifts from σ_0 to σ_1 at time $\lfloor \tau_\sigma T \rfloor$ (we let $\sigma_0 = 1$ without loss of generality), with $\sigma_1/\sigma_0 \in \{1/10, 1/9, ..., 1/2, 1, 2, 3, ..., 10\}$ as before. We set $z_t \sim$ NIID(0,1), $u_1 = \varepsilon_1$ and $C(L) = 1$, abstracting at this stage from the effects of autocorrelated innovations. The size behaviour of the standard unit root tests $\mathcal{MSB}(\bar{\tau})$, $\mathcal{MZ}_t(\bar{\tau})$ and $t(\bar{\tau})$ is considered, both when the true lag order is assumed known (i.e. $p = 0$ in the computation of the test statistics), and also the feasible versions where the lag order is treated as unknown and determined according to p_{MAIC}.[9] We also report results for the corresponding feasible wild bootstrap tests $\mathcal{MSB}(\bar{\tau})^*$, $\mathcal{MZ}_t(\bar{\tau})^*$ and $t(\bar{\tau})^*$, where the statistics are again calculated using $p = p_{\text{MAIC}}$ ($p^* = 0$ is used in constructing the bootstrap statistics based on the re-sampled data, as outlined above). We use the same four deterministic specifications as in the asymptotic simulations, and in each case we consider a single timing for the break in volatility, namely: (a) no break in trend, $\tau_\sigma = 0.7$, (b) break in trend, $\tau_0 = \tau_\sigma = 0.3$, (c) break in trend, $\tau_0 = \tau_\sigma = 0.5$, and (d) break in trend, $\tau_0 = \tau_\sigma = 0.7$. When a break in trend is present in the DGP, in order to standardize its magnitude in the presence of different volatility specifications, we set $\gamma_0 = \gamma_0' \bar{\sigma}$, where $\bar{\sigma}$ is the average standard deviation across the sample period (i.e., $\bar{\sigma} := T^{-1} \sum_{t=1}^{T} \sigma_t$); we let $\gamma_0' = 1$ throughout. Finally, we let $\alpha_0 = \beta_0 = 0$ without loss of generality, and consider the sample size $T = 150$.

The pattern of finite sample size behaviour for the $\mathcal{M}(\bar{\tau})$ and $t(\bar{\tau})$ tests when $p = 0$ closely resembles the asymptotic results for these tests presented in Figure 1. As might be expected, the size distortions are more severe for $T = 150$ than in the limit, particularly in the case of no break. The overall picture is one of substantial over-size, emphasizing the lack of robustness of these procedures to non-stationary volatility. When lag augmentation is incorporated into the original statistics ($p = p_{\text{MAIC}}$), the size distortions are reduced, in many cases even below the level observed in the limit; it appears, therefore, that the MAIC procedure results in an over-specification of the lag order when non-stationary volatility is present, and this property (somewhat artificially) mitigates the extent to which the tests display over-size in finite samples. However, substantial size distortions are still observed, with the $\mathcal{M}(\bar{\tau})$ and $t(\bar{\tau})$ tests lacking adequate size control when shifts in volatility are present. On the other hand, the feasible wild bootstrap tests ($\mathcal{MSB}(\bar{\tau})^*$, $\mathcal{MZ}_t(\bar{\tau})^*$ and $t(\bar{\tau})^*$ with $p = p_{\text{MAIC}}$) display very attractive size properties. When a break in trend occurs, the tests are conservative, while in the no trend break case, size is close to the nominal level, with only

4.2.

[9]The sizes of $\mathcal{MZ}_\alpha(\bar{\tau})$ are similar to the other $\mathcal{M}(\bar{\tau})$ tests, and always lie between the sizes of $\mathcal{MSB}(\bar{\tau})$ and $\mathcal{MZ}_t(\bar{\tau})$; results for $\mathcal{MZ}_\alpha(\bar{\tau})$ are thus omitted for the sake of clarity in the figures.

a little over-size observed (size is always less than 0.08). The contrast with the original tests is marked, with the wild bootstrap tests clearly exhibiting reliable size properties for both homoskedastic and non-stationary volatility models, in finite samples as well as in the limit.

Figures 6-7 present finite sample power comparisons of the original and bootstrap tests, again for $T = 150$. As in the asymptotic analysis, we focus on the $\mathcal{MZ}_t(\bar{\tau})$ and $t(\bar{\tau})$ tests, along with their bootstrap analogues, $\mathcal{MZ}_t(\bar{\tau})^*$ and $t(\bar{\tau})^*$, owing to the fact that the three $\mathcal{M}$-based tests again display very similar size-adjusted powers. Results are provided for $c \in \{0, 1, ..., 50\}$ for the same four deterministic specifications as before, each for the case of homoskedastic errors (Figure 6) and an illustrative volatility shift example (Figure 7). We report results for tests using both the infeasible setting of $p = 0$ and the feasible version $p = p_{\text{MAIC}}$. To enable meaningful comparisons, the powers of the original tests are size-adjusted so as to ensure that a particular test has size equal to that of the corresponding wild bootstrap test.

On comparing the results of Figure 6 with those of Figure 7, it can be seen that the power of the original tests is adversely affected by the presence of volatility shifts, as was seen in the limit. Further, when lag selection according to MAIC is employed in the computation of the test statistics, the powers are reduced relative to when p is (infeasibly) set to zero. This is consistent with the fact that tests using $p = p_{\text{MAIC}}$ have lower sizes than those based on $p = 0$, and represents the cost of having to select the lag order in practice, particularly when the volatility is non-stationary. However, the most important comparison in Figures 6-7 for our purposes is the power of the wild bootstrap tests relative to the (size-adjusted) power of the original tests. Here, the asymptotic prediction that the bootstrap tests should have power approximately equal to the size-adjusted power of the standard tests is borne out in finite samples. For both constant volatility and non-stationary volatility specifications, the power functions of the bootstrap tests follow the size-adjusted power curves of the original tests very closely, and in some circumstances the bootstrap tests actually achieve power in excess of the size-adjusted power of the corresponding standard tests. The bootstrap tests therefore have very attractive finite sample properties: there is little cost involved in applying the bootstrap tests when the volatility function is in fact constant, while obvious gains are displayed when changes in volatility occur; in the latter case, the wild bootstrap tests achieve reliable finite sample size (unlike their standard test counterparts), and at the same time have power very similar to the size-adjusted power of the corresponding standard tests.

We next consider the size behaviour of the original and bootstrap tests when the shocks are autocorrelated, i.e. when $C(L) \neq 1$ in (3). We now let the ε_t be generated according to either an AR(1) process: $\varepsilon_t = \phi\varepsilon_{t-1} + e_t$ (with $\varepsilon_1 = e_1$), or an MA(1) process: $\varepsilon_t = e_t - \theta e_{t-1}$ (again with $\varepsilon_1 = e_1$). When a break in trend is present, we additionally standardize the break magnitude by the long-run standard deviation of ε_t, i.e. $\gamma_0 := \gamma_0'\bar{\sigma}/(1 - \phi)$ and $\gamma_0 := \gamma_0'\bar{\sigma}(1 - \theta)$ for the AR(1) and MA(1) processes, respectively. Here, we restrict attention to feasible tests that determine the lag order to be used in the statistics according to p_{MAIC}. Table 2 reports results for a number of ϕ and θ settings for the same eight DGP settings considered in Figures 6-7, but with $c = 0$. When the volatility is constant, the wild bootstrap tests have sizes similar to those of the corresponding standard tests, and although some over-size is observed in the presence of autocorrelated shocks, size is on the whole reasonably well controlled in small samples, particularly for $t(\bar{\tau})$ and $t(\bar{\tau})^*$. For cases where a shift in volatility occurs, the wild bootstrap tests continue to control size across the range of DGPs considered (in fact, the sizes are generally lower than in the corresponding homoskedastic cases), but now the standard tests are mostly over-sized, as expected. Comparing the different variants of the wild bootstrap tests, $\mathcal{MZ}_\alpha(\bar{\tau})^*$, $\mathcal{MSB}(\bar{\tau})^*$ and $\mathcal{MZ}_t(\bar{\tau})^*$ have almost identical sizes, while $t(\bar{\tau})^*$ displays less upward size distortion. Since $t(\bar{\tau})^*$ loses little in terms of power relative to the $\mathcal{M}$-based alternatives, this test appears to have the most appealing finite sample properties of the tests considered in this paper.

Finally, we investigate the sensitivity of the tests' finite sample size behaviour to different degrees of trimming; i.e., alternative settings for $\Lambda = [\tau_L, \tau_U]$. Table 3 reports results for a subset of the DGPs examined above, with $\Lambda = [0.10, 0.90]$ and $\Lambda = [0.05, 0.95]$, i.e. 10% and 5% trimming, respectively. Of the eight DGPs previously considered, the amount of trimming is most likely to affect test size for the cases where no break in trend is present, due to the important role of τ_L in determining whether the linear or broken trend versions of the tests are applied (see (11) and (14)). We therefore report results for the two DGPs where only a linear trend is present. On comparing the results of the two panels of Table 3, along with the corresponding entries in Table 2, it is clear that the degree of trimming has little effect on the finite sample size of the tests (sizes generally decrease very slightly with the degree of trimming), highlighting a reassuring degree of robustness to this setting. Unreported results for the other six DGPs, for which a break in trend is present, confirm that the finite sample sizes are even less sensitive to the different trimming choices than in the no trend break cases.

6 Conclusions

In this paper we have explored the impact that non-stationary volatility has on unit root tests which allow for a possible break in trend. We have focused on the ADF-based test of HHLT and extensions thereof using the corresponding $\mathcal{M}$-type tests. Numerical evidence was presented which showed that non-stationary volatility can have potentially serious implications for the reliability of these tests with size often being substantially above the nominal level. This was shown to be a feature of the limiting distributions of the statistics. To rectify this problem, we have proposed wild bootstrap-based implementations of the tests, these having proved to be highly successful in other unit root testing applications. The proposed bootstrap tests have the considerable advantage that they are not tied to a given parametric model of volatility within the class of non-stationary volatility processes considered. The asymptotic validity of our proposed bootstrap tests within the class of non-stationary volatility considered was demonstrated. Monte Carlo simulation evidence for the case of a one-time change in volatility was also reported which suggested that the proposed bootstrap unit root tests perform well in finite samples avoiding the large oversize problems that can occur with the standard tests, yet emulating the finite sample power properties of (infeasible) size-adjusted implementations of the standard tests.

Acknowledgements

We thank Peter Philips, Tim Vogelsang, two anonymous referees and participants at the Granger Centre Conference on Bootstrap and Numerical Methods in Time Series, held in Nottingham on September 10–11, 2008 for useful comments on earlier versions of this paper. We also gratefully acknowledge the financial support provided by the Economic and Social Research Council of the UK under research grant RES-000-22-3882.

A Appendix

This appendix is organized as follows. In section A.1 we provide the mathematical proofs of Theorem 1 and the related Lemma 1. This section also contains some preliminary lemmas that are used in the sequel. Section A.2 contains proofs related to the bootstrap tests. Some useful results related to bootstrap tests for the case of no trends breaks or a trend break at a known date are also reported. Detailed proofs are reported only for the tests based on $\mathcal{MZ}_\alpha$ (as well as for its bootstrap counterparts). Proofs for tests based on $\mathcal{MZ}_t$, $\mathcal{MSB}$ and *ADF-GLS* are similar and omitted for brevity. Finally, in this appendix we make use of the following notation: we let $\sigma^{*2} := \max\{\sup_s \omega(s), \breve{\sigma}\}$ and $\bar{\omega}^2 := \int_0^1 \omega(s)^2 ds$; moreover, we let $\lambda_\omega^2 := \sigma^2 C(1)^2$ denote the (asymptotic) average long run variance of ε_t.

A.1 Proof of Lemma 1, Theorem 1 and Related Results

A.1.1 Preliminary Lemmas

The following lemmas present the asymptotic distribution of the $\mathcal{MZ}_\alpha^t$ statistic in the case of no trend breaks ($\gamma_0 = 0$) and of the $\mathcal{MZ}_\alpha^{tb}(\tau_0)$ in the general case. Notice that the latter statistic corresponds to the case where the date of the possible break is known.

Lemma A.1 *Let y_t be generated according to (1)-(4) under H_c, and let Assumptions $\mathcal{A}$ and $\mathcal{B}$ hold. Then, if $\gamma_0 = 0$, $\mathcal{MZ}_\alpha^t \xrightarrow{w} \bar{\xi}_1^{c,\bar{c},t,\eta}$.*

Proof. The theorem follows by adapting the same arguments as in the proof of Theorem 1 in Cavaliere and Taylor (2007) to the case of GLS detrended data.

Lemma A.2 *Let y_t be generated according to (1)-(4) under H_c, and let Assumptions $\mathcal{A}$ and $\mathcal{B}$ hold. Then,* $\mathcal{MZ}_\alpha^{tb}(\tau_0) \xrightarrow{w} \xi_1^{c,\bar{c},tb,\eta}(\tau_0)$.

Proof. The theorem follows by adapting the same arguments as in the proof of Theorem 1 in Perron and Rodríguez (2003) to the case of heteroskedastic data. To this end, one has simply to amend their Lemma A.1 to the following:
(a) $T^{-1/2}u_{\lfloor T\cdot\rfloor} \xrightarrow{w} \bar{\omega}W_c^\eta(\cdot)$; (b) $T^{-3/2}\sum_{t=1}^T u_t \xrightarrow{w} \bar{\omega}\int_0^1 W_c^\eta(s)\,ds$; (c) $T^{-2}\sum_{t=1}^T u_t^2 \xrightarrow{w} \bar{\omega}^2\int_0^1 W_c^\eta(s)^2\,ds$.
The result in part (a) follows as in Remark 3.1 of Cavaliere and Taylor (2007), while parts (b) and (c) follow from (a) and the CMT.

A.1.2 Proof of Lemma 1

Part (i). We begin by proving the consistency of $\tilde{\tau}$ when $\gamma_0 \neq 0$. When $c = 0$, since $v_t := \Delta u_t = \varepsilon_t$ the proof follows by adapting Proposition 3 of Bai (1994) to independent, but not identically distributed errors. In order to extend Proposition 3 of Bai (1994) to the present framework it suffices to show that in the presence of heteroskedasticity, Bai's generalization of Hajek-Renyi inequality (see Bai, 1994, Proposition 1) becomes as follows:

$$\Pr\left(\max_{m\leq k\leq n} c_k\left|\sum_{i=1}^k \varepsilon_i\right|\right) \leq C_0\frac{\sigma^{*2}}{\alpha^2}\left(mc_m^2 + \sum_{i=m+1}^n c_i^2\right) \tag{A.1}$$

with $\sigma^{*2} := \max\{\sup_s \omega(s), \breve{\sigma}\}$ and C_0 as defined in Bai (1994, p.470). As in Bai (1994,p.457), this inequality implies that

$$\Pr\left(\sup_{k\geq m}\frac{1}{k}\left|\sum_{i=1}^k \varepsilon_t\right|\right) \leq \frac{C_1}{\alpha^2 m}. \tag{A.2}$$

Using (A.1) and (A.2) it is straightforward to see that the proof of Proposition 3 in Bai (1994) holds under our assumptions, hence implying, for the case $c = 0$, that $\tilde{\tau} - \tau_0 = O_p\left((T\gamma_0)^{-1}\right)$. The foregoing results generalise to the case where $c > 0$ using the same arguments as are made in HHLT, proof of Lemma 1.

In order to prove the result for $\bar{\tau}$ we need to establish that $T^{-1}W_T(\tau_0)$ has a well-defined distribution; see HHLT, proofs of Lemma 2 and Lemma 3(i). As in HHLT, consider the partial sums of the DGP in (1), which can be written as $w_t = Z_{1t}'\theta_{1,0} + z_{2,\tau_0}'\gamma_0 + s_t$, $t = 1, ..., T$, where $w_t := \sum_{i=1}^t y_i$, $Z_{1t} := \sum_{i=1}^t (1,i)'$, $z_{2,\tau t} := \sum_{i=1}^t DT_i(\tau_0)$, $s_t := \sum_{i=1}^t u_i$ and $\theta_{1,0} := (\alpha_0, \beta_0)'$. In matrix form we can write $w = Z_1\theta_{1,0} + z_{2,\tau}\gamma_0 + s = Z_{\tau_0}\theta_0 + s$, where $Z_\tau = (Z_1, z_{2,\tau})$. We then have that

$$W_T(\tau) = \frac{\left(z_{2,\tau}'\bar{P}_1 w\right)^2}{\left(z_{2,\tau}'\bar{P}_1 z_{2,\tau}\right)\left(w'\bar{P}_\tau w\right)}$$

where $\bar{P}_1 := I_T - Z_1 (Z_1' Z_1)^{-1} Z_1'$ and $\bar{P}_\tau := I_T - Z_\tau (Z_\tau' Z_\tau)^{-1} Z_\tau'$. When $\tau = \tau_0$, we have

$$W_T(\tau_0) = \frac{\left(\gamma_0 z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0} + z_{2,\tau_0}' \bar{P}_1 s\right)^2}{\left(z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0}\right)\left(s' \bar{P}_1 s\right) - \left(z_{2,\tau_0}' \bar{P}_1 s\right)^2}$$

which implies that

$$T^{-1} W_T(\tau_0) = \frac{\left(\gamma_0 T^{-5} z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0} + T^{-5} z_{2,\tau_0}' \bar{P}_1 s\right)^2}{\left(T^{-5} z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0}\right)\left(T^{-4} s' \bar{P}_1 s\right) - \left(T^{-9/2} z_{2,\tau_0}' \bar{P}_1 s\right)^2}.$$

Let $Z_1(r) := \left(r, \frac{1}{2} r^2\right)'$, $Z_{2,\tau}(r) := \frac{1}{2}(r-\tau)^2 \mathbb{I}(r \geq \tau)$ and $Z_\tau(r) := \left(Z_1(r)', Z_{2,\tau}(r)\right)'$. As in HHLT, we then have that

$$T^{-5} z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0} \quad \rightarrow \quad \int_0^1 Z_{2.1,\tau_0}(r)^2 \, dr \tag{A.3}$$

and that $T^{-5} z_{2,\tau_0}' \bar{P}_1 s = o_p(1)$, where $Z_{2.1,\tau_0}(r)$ is the residual process from a projection of $Z_{2,\tau_0}(r)$ on $Z_1(r)$. To deal with the remaining terms, we make use of the FCLT from Theorem 1(i) of Cavaliere and Taylor (2007), which establishes that

$$\frac{1}{T^{1/2}} \sum_{t=1}^{\lfloor Tr \rfloor} \varepsilon_t \xrightarrow{w} \lambda_\omega W^\eta(r) \tag{A.4}$$

where W^η is the time-change Brownian motion $W^\eta(r) := W(\eta(r))$, W denoting a standard BM. Using standard arguments we have that $(u_{\lfloor Tr \rfloor}, s_{\lfloor Tr \rfloor}) \xrightarrow{w} (W_c^\eta(r), \mathbb{W}_c^\eta(r))$, with $W_c^\eta(r) := \int_0^r e^{-(r-s)c} dW^\eta(s)$ and $\mathbb{W}_c^\eta(r) := \int_0^r W_c^\eta(s)\, ds$. The CMT further implies

$$T^{-4} s' \bar{P}_1 s \xrightarrow{w} \lambda_\omega^2 \int_0^1 S_1(r)^2 \, dr \tag{A.5}$$

$$T^{-9/2} z_{2,\tau_0}' \bar{P}_1 s \xrightarrow{w} \lambda_\omega \int_0^1 Z_{2.1,\tau_0}(r) S_1(r) \, dr \tag{A.6}$$

where S_1 is the residual process from a projection of $\mathbb{W}_c^\eta$ on Z_1. Hence,

$$T^{-1} W_T(\tau_0) \xrightarrow{d} \frac{\gamma_0^2 \int_0^1 Z_{2.1,\tau_0}(r)^2 \, dr}{\int_0^1 Z_{2.1,\tau_0}(r)^2 \, dr \int_0^1 S_1(r)^2 \, dr - \left(\int_0^1 Z_{2.1,\tau_0}(r) S_1(r) \, dr\right)^2} = O_p(1),$$

as required.

Part (ii). When $\gamma_0 = 0$ we can apply Theorem 3.1 of Nunes et al. (1995) to get $\tilde{\tau} = O_p(1)$. Then, a sufficient condition for $\bar{\tau}$ to be of $O_p(T^{-1/2})$ is that $W_T(\tau_0) = O_p(1)$; cf. HHLT, proof of Lemma 3(i). But this immediately follows from (A.3), (A.5), (A.6), and by noting that, under $\gamma_0 = 0$,

$$W_T(\tau_0) = \frac{\left(T^{-9/2} z_{2,\tau_0}' \bar{P}_1 s\right)^2}{\left(T^{-5} z_{2,\tau_0}' \bar{P}_1 z_{2,\tau_0}\right)\left(T^{-4} s' \bar{P}_1 s\right) - \left(T^{-9/2} z_{2,\tau_0}' \bar{P}_1 s\right)^2} = O_p(1) \ .$$

A.1.3 Proof of Theorem 1

By using Lemmas A.1 and A.2 above, the stated results follow immediately by establishing that: (a) if $\gamma_0 = 0$, then $\mathcal{MZ}_\alpha(\bar{\tau}) - \mathcal{MZ}_\alpha^t \xrightarrow{p} 0$; and, (b) if $\gamma_0 \neq 0$, then $\mathcal{MZ}_\alpha(\bar{\tau}) - \mathcal{MZ}_\alpha^{tb}(\tau_0, \bar{c}) \xrightarrow{p} 0$.

Proof of (a). This follows from Lemma 1, which implies that $\Pr(\bar{\tau} < \tau_T) \to 1$ and, hence, that $\mathcal{MZ}_\alpha(\bar{\tau}) - \mathcal{MZ}_\alpha^t \xrightarrow{p} 0$.

Proof of (b). The proof obtains by showing that, provided $\tilde{\tau}$ is any $O_p(T^{-1})$ consistent estimator of τ_0, it holds that $\mathcal{MZ}_\alpha^{tb}(\tilde{\tau}, \bar{c}) = \mathcal{MZ}_\alpha^{tb}(\tau_0, \bar{c}) + o_p(1)$. We now establish this.

As in HHLT, we make use of the matrix notation

$$
\begin{aligned}
y &= X\theta_0 + u \\
y_{\bar{c}} &= X_{\bar{c}}\theta_0 + u_{\bar{c}},
\end{aligned}
\tag{A.7}
$$

where $\theta_0 := (\alpha_0, \beta_0, \gamma_0)'$, $y := (y_1, ..., y_T)'$, $X := (X_1(\tau_0), \ldots, X_T(\tau_0))'$, $u := (u_1, ..., u_T)'$, $y_{\bar{c}} := (y_1, y_2 - \bar{\rho}_T y_1 ..., y_T - \bar{\rho}_T y_{T-1})'$, $X_{\bar{c}} := (X_1(\tau_0), X_2(\tau_0) - \bar{\rho}X_1(\tau_0), \ldots, X_T(\tau_0) - \bar{\rho}X_{T-1}(\tau_0))'$ and $u_{\bar{c}} := (u_1, u_2 - \bar{\rho}_T u_1 ..., u_T - \bar{\rho}_T u_{T-1})'$. The quasi-GLS estimator of θ_0 obtained from (A.7) (with τ_0 known) is denoted $\hat{\theta}_{\bar{c}}$; the corresponding de-trended series is $\hat{u} := y - X\hat{\theta}_{\bar{c}}$. The $\mathcal{MZ}_\alpha$ test statistic for $\hat{u}$ is

$$
\mathcal{MZ}_\alpha^{tb}(\tau_0, \bar{c}) := \frac{\hat{u}_T^2/T - \hat{s}_{AR}^2(p)}{2\hat{u}_{-1}'\hat{u}_{-1}/T}
$$

with $\hat{u}_{-1} := (\hat{u}_1, ..., \hat{u}_{T-1})'$. The estimator $\hat{s}_{AR}^2(p) := \hat{\sigma}^2/(1 - \hat{\delta}'\iota)^2$ (ι being the p-dimensional unit vector) is obtained from the regression

$$
\begin{aligned}
\Delta\hat{u}_t &= \phi\hat{u}_{t-1} + \delta'\hat{U}_{p,t} + e_{p,t} \ (t = p+1, ..., T) \\
&= a'\hat{Z}_t + e_{p,t} \ ,
\end{aligned}
\tag{A.8}
$$

where $\hat{Z}_t := (\hat{u}_{t-1}, \hat{U}_{p,t}')'$. Similarly, when $\tilde{\tau}$ is used instead of τ_0, we define the matrices $\tilde{X} := (X_1(\tilde{\tau}), \ldots, X_T(\tilde{\tau}))'$ and $\tilde{X}_{\bar{c}} := (X_{\bar{c},1}(\tilde{\tau}), \ldots, X_{\bar{c},T}(\tilde{\tau}))'$. The OLS estimator from a regression of $y_{\bar{c}}$ on $\tilde{X}_{\bar{c}}$ is denoted $\tilde{\theta}_{\bar{c}}$ and the corresponding de-trended series is $\tilde{u} := y - \tilde{X}\tilde{\theta}_{\bar{c}}$. The test statistic on $\tilde{u}$ is

$$
\mathcal{MZ}_\alpha(\tilde{\tau}, \bar{c}) := \frac{\tilde{u}_T^2/T - \tilde{s}_{AR}^2(p)}{2\tilde{u}_{-1}'\tilde{u}_{-1}/T}
$$

with $\tilde{s}_{AR}^2(p) := \tilde{\sigma}^2/(1 - \tilde{\delta}'\iota)^2$ estimated from the regression

$$
\begin{aligned}
\Delta\tilde{u}_t &= \phi\tilde{u}_{t-1} + \delta'\tilde{U}_{p,t} + e_{p,t} \\
&= a'\tilde{Z}_t + e_{p,t}
\end{aligned}
\tag{A.9}
$$

with $\tilde{Z}_t := \left(\tilde{u}_{t-1}, \tilde{U}_{p,t}'\right)'$.

We need to establish some preliminary results first. As in HHLT, in order to prove the theorem it is useful to establish that, for $D_T := \text{diag}\left(1, T^{1/2}, T^{1/2}\right)$, $\hat{\tau} - \tau_0 = O_p\left(T^{-1}\right)$ implies that

$$D_T(\tilde{\theta}_{\bar{c}} - \hat{\theta}_{\bar{c}}) = O_p\left(T^{-1/2}\right). \tag{A.10}$$

Furthermore, we need to show that, uniformly for all $s \in [0,1]$,

$$T^{-1/2}\left(\tilde{u}_{\lfloor Ts \rfloor} - \hat{u}_{\lfloor Ts \rfloor}\right) = O_p\left(T^{-1/2}\right) \tag{A.11}$$

and that, uniformly in $i, j = 1, \dots p$,

$$T^{-1}\sum_{t=p+1}^{T} \Delta\tilde{u}_{t-i}\Delta\tilde{u}_{t-j} - T^{-1}\sum_{t=p+1}^{T} \Delta\hat{u}_{t-i}\Delta\hat{u}_{t-j} = O_p\left(T^{-1}\right) \tag{A.12}$$

$$T^{-1}\sum_{t=p+1}^{T} \tilde{u}_{t-1}\Delta\tilde{u}_{t-i} - T^{-1}\sum_{t=p+1}^{T} \hat{u}_{t-1}\Delta\hat{u}_{t-i} = O_p\left(T^{-1/2}\right). \tag{A.13}$$

These results, together with the following rates

$$\left\| \left(T^{-1}\hat{U}_p'\hat{U}_p\right)^{-1} \right\| = O_p(1) \tag{A.14}$$

$$\left\| T^{-1}\hat{u}_{-1}'\hat{U}_p \right\| = O_p\left(p^{1/2}\right) \tag{A.15}$$

implies that Lemma 6 in HHLT holds under the conditions of Theorem 1 here.

To prove (A.10), we find from HHLT, that

$$D_T^{-1}X_{\bar{c}}'X_{\bar{c}}D_T^{-1} = \sum_{t=1}^{T} D_T^{-1}X_{\bar{c},t}(\tau_0)\,X_{\bar{c},t}(\tau_0)'\,D_T^{-1} \to \begin{pmatrix} 1 & 0 \\ 0 & \int_0^1 H_{\bar{c},\tau_0}(s)\,H_{\bar{c},\tau_0}(s)'\,ds \end{pmatrix}$$

with $H_{\bar{c},\tau_0}(s) := (1 + \bar{c}s, 1(s > \tau_0)(1 + \bar{c}(s - \tau_0)))'$. Furthermore, since (for $t > 1$) $u_{\bar{c},t} = \Delta u_t + \frac{\bar{c}}{T}u_{t-1} = \varepsilon_t - \frac{c}{T}u_{t-1} + \frac{\bar{c}}{T}u_{t-1}$, by standard arguments (specifically, the FCLT in (A.4) and the CMT) it holds that $D_T^{-1}X_{\bar{c}}'u_{\bar{c}} = O_p(1)$. This then implies that $D_T(\hat{\theta}_{\bar{c}} - \theta_0) = O_p(1)$ and, using the same arguments as in HHLT (proof of Lemma 4), that $D_T(\tilde{\theta}_{\bar{c}} - \hat{\theta}_{\bar{c}}) = O_p\left(T^{-1/2}\right)$, as required.

The proofs of (A.11)-(A.13) obtain as in HHLT, proof of Lemma 5, which does not require unconditional homoskedasticity. Finally, (A.14) and (A.15) follow from Cavaliere and Taylor (2007, Lemma 3).

We can now make use of these preliminary results to prove the main result in (b). First, notice that (A.11) and the fact that $\sup|T^{-1/2}\hat{u}_T|$ and $\sup|T^{-1/2}\tilde{u}_T|$ are both of $O_p(1)$ imply that $T^{-1}(\hat{u}_T^2 - \tilde{u}_T^2) = O_p(T^{-1/2})$ and, similarly, that

$$\frac{\tilde{u}_{-1}'\tilde{u}_{-1}}{T^2} - \frac{\hat{u}_{-1}'\hat{u}_{-1}}{T^2} = O_p(T^{-1/2}). \tag{A.16}$$

Hence, it is left to prove that $\tilde{s}^2_{AR}(p) - \hat{s}^2_{AR}(p) = o_p(1)$, which follows if we can show that $\tilde{\sigma}^2 - \hat{\sigma}^2 = o_p(1)$ and that $|\tilde{\delta}'\iota - \hat{\delta}'\iota| = o_p(1)$.

Regarding the difference $\tilde{\sigma}^2 - \hat{\sigma}^2$, notice that

$$\tilde{\sigma}^2 - \hat{\sigma}^2 = T^{-1}\left(\Delta\tilde{u}'\tilde{P}\Delta\tilde{u} - \Delta\hat{u}'\hat{P}\Delta\hat{u}\right) + T^{-1}\left(\frac{\left(\hat{u}'_{-1}\hat{P}\Delta\hat{u}\right)^2}{\hat{u}'_{-1}\hat{P}\hat{u}_{-1}} - \frac{\left(\tilde{u}'_{-1}\tilde{P}\Delta\tilde{u}\right)^2}{\tilde{u}'_{-1}\tilde{P}\tilde{u}_{-1}}\right)$$

with $\Delta\hat{u} := (\Delta\hat{u}_2, ..., \Delta\hat{u}_T)'$, $\hat{U}_p := (\hat{U}_{p,2}, \ldots, \hat{U}_{p,T})'$, $\hat{P} := I_{T-p-1} - \hat{U}_p\left(\hat{U}'_p\hat{U}_p\right)^{-1}\hat{U}'_p$; $\Delta\tilde{u}$, $\tilde{u}_{-1}$, $\tilde{U}_p$, $\tilde{P}$ are defined similarly. First, as in HHLT, (A.13), (A.14) and Lemma 6 in HHLT implies that $T^{-1}\left(\tilde{u}'_{-1}\tilde{P}\Delta\tilde{u} - \hat{u}'_{-1}\hat{P}\Delta\hat{u}\right) \xrightarrow{p} 0$ and, by using (A.16), that $T^{-1}\left(\tilde{u}'_{-1}\tilde{P}\tilde{u}_{-1} - \hat{u}'_{-1}\hat{P}\hat{u}_{-1}\right) \xrightarrow{p} 0$. Similarly, (A.12), (A.14) and Lemma 6 in HHLT imply that $T^{-1}\left(\Delta\tilde{u}'\tilde{P}\Delta\tilde{u} - \Delta\hat{u}'\hat{P}\Delta\hat{u}\right) \xrightarrow{p} 0$. Taken together, these results yield the result that $\tilde{\sigma}^2 - \hat{\sigma}^2 \to 0$, in probability.

It now remains to prove that $|\tilde{\delta}'\iota - \hat{\delta}'\iota| = o_p(1)$. Letting $G_T := diag(T, T^{1/2}I_p)$, we show that under the stated conditions, $||G_T(\tilde{a}' - \hat{a})|| = o_p(p^{1/2}/T^{1/2})$. Since $\hat{\delta} = (0, I_p)\hat{a}$, this result implies that $\iota'|\tilde{\delta} - \hat{\delta}| \leq p^{1/2}||\tilde{\delta} - \hat{\delta}|| = o_p(1)$, as $p = o(T^{1/2})$, hence completing the proof. To that end, notice that

$$\begin{aligned} G_T(\tilde{a} - \hat{a}) &= \left(G_T^{-1}\tilde{Z}'\tilde{Z}G_T^{-1}\right)^{-1} G_T^{-1}\tilde{Z}'\Delta\tilde{u} - \left(G_T^{-1}\hat{Z}'\hat{Z}G_T^{-1}\right)^{-1} G_T^{-1}\hat{Z}'\Delta\hat{u} \\ &= \left(\left(G_T^{-1}\tilde{Z}'\tilde{Z}G_T^{-1}\right)^{-1} - \left(G_T^{-1}\hat{Z}'\hat{Z}G_T^{-1}\right)^{-1}\right) G_T^{-1}\tilde{Z}'\Delta\tilde{u} \\ &\quad + \left(G_T^{-1}\hat{Z}'\hat{Z}G_T^{-1}\right)^{-1}\left(G_T^{-1}\tilde{Z}'\Delta\tilde{u} - G_T^{-1}\hat{Z}'\Delta\hat{u}\right) \end{aligned}$$

Standard manipulations (Berk, 1974, proof of Lemma 3) establish that, up to an $o_p(1)$ term,

$$\left\|\left(G_T^{-1}\tilde{Z}'\tilde{Z}G_T^{-1}\right)^{-1} - \left(G_T^{-1}\hat{Z}'\hat{Z}G_T^{-1}\right)^{-1}\right\| \leq \text{const}\left\|\left(G_T^{-1}\tilde{Z}'\tilde{Z}G_T^{-1}\right) - \left(G_T^{-1}\hat{Z}'\hat{Z}G_T^{-1}\right)\right\| = O_p\left(T^{-1/2}\right).$$

To show the last result, let $q := \left(G_T^{-1}\tilde{Z}'\tilde{Z}G_T^{-1}\right) - \left(G_T^{-1}\hat{Z}'\hat{Z}G_T^{-1}\right)$. The upper left element of q, say q_{11}, satisfies $q_{11} = T^{-2}\tilde{u}'_{-1}\tilde{u}_{-1} - T^{-2}\hat{u}'_{-1}\hat{u}_{-1} = O_p(T^{-1/2})$; see (A.16). The lower right block is given by $q_{22} = T^{-1}\tilde{U}'_p\tilde{U}_p - T^{-1}\hat{U}'_p\hat{U}_p$ and satisfies $||q_{22}|| = O_p(pT^{-1})$, as in HHLT, proof of Lemma 6. Similarly, the lower left block of q, is such that

$$||q_{21}|| = ||T^{-3/2}\tilde{U}'_p\tilde{u}_{-1} - T^{-3/2}\hat{U}'_p\hat{u}_{-1}|| = O_p\left(p^{1/2}T^{-1}\right),$$

which implies, as $p = o(T^{1/2})$, that $||q|| = O_p(T^{-1/2})$.

Regarding the term $G_T^{-1}\tilde{Z}'\Delta\tilde{u} - G_T^{-1}\hat{Z}'\Delta\hat{u}$, we can again proceed as in HHLT, proof of Lemma 6, to show that

$$\left\|G_T^{-1}\tilde{Z}'\Delta\tilde{u} - G_T^{-1}\hat{Z}'\Delta\hat{u}\right\| = \left\|\begin{array}{c} \frac{1}{T}\tilde{u}'_{-1}\Delta\tilde{u} - \frac{1}{T}\hat{u}'_{-1}\Delta\hat{u} \\ \frac{1}{T^{1/2}}\tilde{U}'_p\Delta\tilde{u} - \frac{1}{T^{1/2}}\hat{U}'_p\Delta\hat{u} \end{array}\right\| = O_p(p^{1/2}/T^{1/2}).$$

Finally, similarly to HHLT (Lemmas 5 and 6) one can show that $||(G_T^{-1}\hat{Z}'\hat{Z}G_T^{-1})^{-1}||$ and $||G_T^{-1}\tilde{Z}'\Delta\tilde{u}||$ are both of $O_p(1)$, which leads to the desired result, $||G_T(\tilde{a} - \hat{a})|| = O_p(p^{1/2}/T^{1/2})$.

A.2 Proof of Theorem 2 and Related Results

In this section we provide the proof of Theorem 2. Before doing so, we need to establish some preliminary results useful for determining the asymptotic distribution of the various bootstrap statistics; see section A.2.1. In section A.2.2 we provide some results for the bootstrap versions of the $\mathcal{MZ}_\alpha^t$ and $\mathcal{MZ}_\alpha^{tb}(\tau_0, \bar{c})$ statistics. The proof of Theorem 2 is given in section A.2.3.

We denote the probability distribution induced by the bootstrap (i.e., conditional on the original data) by P^*. Recall that $\xrightarrow{w}_p$ denotes weak convergence in probability; i.e., $X_T^* \xrightarrow{w}_p X$ if $\sup_{x\in\mathbb{R}} |P^*(X_T \leq x) - P(X \leq x)| \to 0$, in probability. We also make use of the following notation (see Chang and Park, 2003, p.386). We say that $X_T^* = o_p^*(1)$ if $P^*(|X_T^*| > \epsilon) \xrightarrow{p} 0$ for any $\epsilon > 0$. Similarly, $X_T^* = O_p^*(1)$ means that for any $\epsilon > 0$ we can find a constant M such that for T large enough, $P^*(|X_T^*| > M) < \epsilon$, in probability. Notice that most of the standard results for o_p and O_p extend to o_p^* and O_p^*, see e.g. Lemma 1 in Chang and Park (2003). Results involving both o_p,O_p and o_p^*,O_p^* are also possible. For instance, if $X_T^* = o_p^*(1)$ and $X_T = o_p(1)$, then it is easy to show that $X_T^* X_T = o_p^*(1)$. Similarly, if $X_T^* = O_p^*(1)$ and $X_T = o_p(1)$, then it can be shown that $X_T^* X_T = o_p^*(1)$.

A.2.1 Preliminary Lemmata

The following Lemma contains a bootstrap FCLT which is necessary in order to derive the asymptotic distributions of the bootstrap statistics considered. The Lemma can be proved as in Cavaliere and Taylor (2008), proof of Theorem 2.

Lemma A.3 *Let* $S_T^*(\cdot) := T^{-1/2} u_{\lfloor T\cdot \rfloor}^* = T^{-1/2} \sum_{t=1}^{\lfloor T\cdot \rfloor} \varepsilon_t^*$. *Moreover, let* $u_{t,\bar{c}}^* := u_t^* - (1 - \bar{c}/T) u_{t-1}^*$ *for* $t = 2, ..., T$ *and* $u_{t,\bar{c}}^* := u_t^*$ *for* $t = 1$. *Finally, let* $S_{T,\bar{c}}^*(\cdot) := \sum_{t=1}^{\lfloor T\cdot \rfloor} u_{t,\bar{c}}^*$. *Then,* $S_T^*(\cdot) \xrightarrow{w}_p W^\eta(\cdot)$ *and* $S_{T,\bar{c}}^*(\cdot) \xrightarrow{w}_p W^\eta(\cdot) + \bar{c} \int_0^{\cdot} W^\eta(s)\, ds$.

A.2.2 Bootstrap Statistics when there is no Uncertainty about the Break date

Theorem A.3 below shows that a wild bootstrap version of the $\mathcal{MZ}_\alpha^t$ test mimics the correct asymptotic distribution if there are no trend breaks ($\gamma_0 = 0$). Similarly, in Theorem A.4 below it is proved that a wild bootstrap version of the $\mathcal{MZ}_\alpha^{tb}(\tau_0)$ (i.e., the break date is known) test delivers the correct asymptotic distribution.

Theorem A.3 (Cavaliere and Taylor, 2008, Theorem 2) *Under the conditions of Theorem 2(i),* $\mathcal{MZ}_\alpha^{t*} \xrightarrow{w}_p \bar{\xi}_1^{0,\bar{c},t,\eta}$.

Proof. See Cavaliere and Taylor (2008).

Theorem A.4 *Under the conditions of Theorem 2(ii),* $\mathcal{MZ}_\alpha^{tb*}(\tau_0) \xrightarrow{w}_p \bar{\xi}_1^{0,\bar{c},tb,\eta}(\tau_0)$.

Proof. The theorem can be proved by using Lemma A.3 and by extending the proof of Theorem 2 in Cavaliere and Taylor (2008) to the case of a broken linear trend. This extension is tedious but straightforward and is therefore omitted for brevity.

A.2.3 Proof of Theorem 2

Part (i). Let $\gamma_0 = 0$. We need to establish that

$$\left| P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) - P(\bar{\xi}_1^{0,\bar{c},t,\eta} \leq x) \right| \xrightarrow{p} 0 \tag{A.17}$$

since, by Polya's theorem, the continuity of the limiting cdf $P(\bar{\xi}_1^{0,\bar{c},t,\eta} \leq \cdot)$ implies that (A.17) holds uniformly over all $x \in \mathbb{R}$. We have that

$$\begin{aligned} \left| P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) - P(\bar{\xi}_1^{0,\bar{c},t,\eta} \leq x) \right| \quad &\leq \quad \left| P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) - P \left(\mathcal{MZ}_\alpha^{t*} \leq x \right) \right| \\ &\quad + \left| P^* \left(\mathcal{MZ}_\alpha^{t*} \leq x \right) - P(\bar{\xi}_1^{0,\bar{c},t,\eta} \leq x) \right|. \end{aligned}$$

The first term in the right member of the preceding inequality is of $o_p(1)$. Specifically, since, conditionally on the data, $\mathbb{I}\left(\bar{\tau} < \tau_L\right)$ is non-random, we have that

$$\begin{aligned} P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) \quad &= \quad P^* \left(\mathcal{MZ}_\alpha^{t*} \leq x \right) \mathbb{I}\left(\bar{\tau} < \tau_L\right) + P^* \left(\mathcal{MZ}_\alpha^{tb*}(\bar{\tau}, \bar{c}) \leq x \right) \mathbb{I}\left(\bar{\tau} \geq \tau_L\right) \\ &= \quad P^* \left(\mathcal{MZ}_\alpha^{t*} \leq x \right) + \left(P^* \left(\mathcal{MZ}_\alpha^{tb*}(\bar{\tau}, \bar{c}) \leq x \right) - P^* \left(\mathcal{MZ}_\alpha^{t*} \leq x \right) \right) \mathbb{I}\left(\bar{\tau} \geq \tau_L\right). \end{aligned}$$

Since $\bar{\tau} \xrightarrow{p} 0$ when $\gamma_0 = 0$ (Lemma 1), it holds that $\mathbb{I}\left(\bar{\tau} \geq \tau_L\right) \xrightarrow{p} 0$ and, hence, that

$$\left| P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) - P^* \left(\mathcal{MZ}_\alpha^{t*} \leq x \right) \right| \leq \mathbb{I}\left(\bar{\tau} \geq \tau_L\right) \xrightarrow{p} 0.$$

This in turn implies that $\left| P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) - P \left(\bar{\xi}_1^{0,\bar{c},t,\eta,0} \leq x \right) \right| = \left| P^* \left(\mathcal{MZ}_\alpha^{t*} \leq x \right) - P \left(\bar{\xi}_1^{0,\bar{c},t,\eta,0} \leq x \right) \right|$ $+ o_p(1) = o_p(1)$, since, by Theorem A.3, $\left| P^* \left(\mathcal{MZ}_\alpha^{t*} \leq x \right) - P(\bar{\xi}_1^{0,\bar{c},t,\eta} \leq x) \right| \to 0$, in probability. This proves Theorem 2(i).

Part (ii). Let $\gamma_0 \neq 0$. As in the proof of part (i), we need to establish that

$$\left| P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) - P(\bar{\xi}_1^{0,\bar{c},tb,\eta} \left(\tau_0 \right) \leq x) \right| \xrightarrow{p} 0. \tag{A.18}$$

Using Theorem A.4 and the same arguments as in the proof of part (i) it is straightforward to see that

$$\begin{aligned} &\left| P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) - P \left(\bar{\xi}_1^{0,\bar{c},tb,\eta} \left(\tau_0 \right) \leq x \right) \right| \leq \left| P^* \left(\mathcal{MZ}_\alpha \left(\bar{\tau} \right)^* \leq x \right) - P^* \left(\mathcal{MZ}_\alpha^{tb*} \left(\bar{\tau} \right) \leq x \right) \right| \\ &+ \left| P^* \left(\mathcal{MZ}_\alpha^{tb*} \left(\bar{\tau} \right) \leq x \right) - P^* \left(\mathcal{MZ}_\alpha^{tb*} \left(\tau_0 \right) \leq x \right) \right| + \left| P^* \left(\mathcal{MZ}_\alpha^{tb*} \left(\tau_0 \right) \leq x \right) - P \left(\bar{\xi}_1^{0,\bar{c},tb,\eta} \left(\tau_0 \right) \leq x \right) \right| \\ &= \left| P^* \left(\mathcal{MZ}_\alpha^{tb*} \left(\bar{\tau} \right) \leq x \right) - P^* \left(\mathcal{MZ}_\alpha^{tb*} \left(\tau_0 \right) \leq x \right) \right| + o_p(1). \end{aligned}$$

Hence, (A.18) follows by proving that $\left| P^* \left(\mathcal{MZ}_\alpha^{tb*} \left(\bar{\tau} \right) \leq x \right) - P^* \left(\mathcal{MZ}_\alpha^{tb*} \left(\tau_0 \right) \leq x \right) \right| \to 0$, in probability. This is implied by the following lemma.

Lemma A.4 *Under the conditions of Theorem 2(ii),* $P^*\left(\left|\mathcal{MZ}_\alpha^{tb*}(\bar{\tau})-\mathcal{MZ}_\alpha^{tb*}(\tau_0)\right|>\epsilon\right)\to 0$*, in probability.*

Proof of Lemma A.4. The proof closely follows the proof of Theorem 1(ii), the only difference being that we are now dealing with bootstrap statistics. In matrix form, the bootstrap sample can be written as $y^*=u^*$ or, after taking the quasi-GLS differences, as $y^*_{\bar{c}}=u^*_{\bar{c}}$. As previously, let $X:=(X_1(\tau_0),\ldots,X_T(\tau_0))'$, $X_{\bar{c}}:=(X_{\bar{c},1}(\tau),\ldots,X_{\bar{c},T}(\tau))'$, $\bar{X}:=(X_1(\bar{\tau}),\ldots,X_T(\bar{\tau}))'$ and $\bar{X}_{\bar{c}}:=(X_{\bar{c},1}(\bar{\tau}),\ldots,X_{\bar{c},T}(\bar{\tau}))'$. The bootstrap sample, detrended at τ_0, is $\hat{u}^*:=y^*-X\hat{\theta}^*_{\bar{c}}$, with $\hat{\theta}^*_{\bar{c}}$ the OLS estimator of the regression coefficients obtained by regressing of $y^*_{\bar{c}}$ on $X_{\bar{c}}$. Similarly, the bootstrap sample, detrended at $\bar{\tau}$, is $\bar{u}^*:=y^*-\bar{X}\bar{\theta}^*_{\bar{c}}$, with $\bar{\theta}^*_{\bar{c}}$ the OLS estimator of the regression coefficients obtained by regressing of $y^*_{\bar{c}}$ on $\bar{X}_{\bar{c}}$. The two statistics of interest are

$$\mathcal{MZ}_\alpha^{tb*}(\tau_0,\bar{c}):=\frac{\hat{u}_T^{*2}/T-\hat{s}^{*2}_{AR}(0)}{2\hat{u}'^*_{-1}\hat{u}^*_{-1}/T},\quad \mathcal{MZ}_\alpha^{tb*}(\bar{\tau},\bar{c}):=\frac{\bar{u}_T^{*2}/T-\bar{s}^{*2}_{AR}(0)}{2\bar{u}'^*_{-1}\bar{u}^*_{-1}/T}$$

with $\hat{u}^*_{-1}:=\left(\hat{u}^*_1,...,\hat{u}^*_{T-1}\right)'$, $\bar{u}^*_{-1}:=\left(\bar{u}^*_1,...,\bar{u}^*_{T-1}\right)'$ and $\hat{s}^{*2}_{AR}(0),\bar{s}^{*2}_{AR}(0)$ the usual long run variance estimators (for $p^*=0$), computed on the bootstrap sample.

As for the proof of Theorem 1, a useful result is that, as T diverges to positive infinity,

$$D_T(\bar{\theta}^*_{\bar{c}}-\hat{\theta}^*_{\bar{c}})=O^*_p\left(T^{-1/2}\right) \tag{A.19}$$

which implies that

$$T^{-1/2}\max_{t=1,\ldots,T}(\hat{u}^*_t-\bar{u}^*_t)=O^*_p\left(T^{-1/2}\right)\ . \tag{A.20}$$

Using results in HHLT (proof of Lemma 4), see also the proof of Theorem 1, to show (A.19) we need to prove that $D_T^{-1}\bar{X}'_{\bar{c}}u^*_{\bar{c}}-D_T^{-1}X'_{\bar{c}}u^*_{\bar{c}}=O_p\left(T^{-1/2}\right)$. But this follows from

$$D_T^{-1}\bar{X}'_{\bar{c}}u^*_{\bar{c}}-D_T^{-1}X'_{\bar{c}}u^*_{\bar{c}}=\left(0,0,T^{-1/2}\sum_{t=1}^{T}\left(DT_{\bar{c},t}(\bar{\tau})-DT_{\bar{c},t}(\tau_0)\right)u^*_{\bar{c},t}\right)=O^*_p\left(T^{-1/2}\right) \tag{A.21}$$

since, by letting (without loss of generality) $\bar{\tau}<\tau_0$, it holds that

$$\begin{aligned}
&T^{-1/2}\sum_{t=1}^{T}\left(DT_{\bar{c},t}(\bar{\tau})-DT_{\bar{c},t}(\tau_0)\right)u^*_{\bar{c},t}\\
=\ &T^{-1/2}\sum_{\lfloor\bar{\tau}T\rfloor<t\le\lfloor\tau_0T\rfloor}\left(1+\bar{c}\left(t-1-\lfloor\bar{\tau}T\rfloor\right)/T\right)u^*_{\bar{c},t}+\bar{c}\frac{\lfloor\tau_0T\rfloor-\lfloor\bar{\tau}T\rfloor}{T}T^{-1/2}\sum_{\lfloor\tau_0T\rfloor<t\le T}u^*_{\bar{c},t}\\
=\ &O^*_p\left(T^{-1/2}\right)+O^*_p\left(T^{-1}\right)=O^*_p\left(T^{-1/2}\right)\text{ as }\lfloor\tau_0T\rfloor-\lfloor\bar{\tau}T\rfloor\text{ is of }O_p(1)\,.
\end{aligned}$$

Equation (A.20) implies (cf. proof of Theorem 1) that

$$T^{-1}\left(\hat{u}_T^{*2}-\bar{u}_T^{*2}\right)=o^*_p(1)\ . \tag{A.22}$$

Moreover, since

$$T^{-2}\left(\hat{u}^{*\prime}_{-1}\hat{u}^{*}_{-1}-\bar{u}^{*\prime}_{-1}\bar{u}^{*}_{-1}\right)=T^{-2}\left(\bar{\theta}^{\prime}_{\bar{c}}\bar{X}^{\prime}_{\bar{c}}u^{*}_{\bar{c}}-\hat{\theta}^{*\prime}_{\bar{c}}X^{\prime}_{\bar{c}}u^{*}_{\bar{c}}\right)+o^{*}_{p}(1)$$
$$=T^{-2}\hat{\theta}^{*\prime}_{\bar{c}}D_T\left(D_T^{-1}\bar{X}^{\prime}_{\bar{c}}u^{*}_{\bar{c}}-D_T^{-1}X^{\prime}_{\bar{c}}u^{*}_{\bar{c}}\right)-T^{-2}\left(D_T\left(\hat{\theta}^{*}_{\bar{c}}-\bar{\theta}_{\bar{c}}\right)\right)^{\prime}D_T^{-1}\bar{X}^{\prime}_{\bar{c}}u^{*}_{\bar{c}}+o^{*}_{p}(1),$$

the results in (A.19), (A.21), $D_T\hat{\theta}^{*}_{\bar{c}}=O^{*}_{p}(1)$ and $D_T^{-1}\bar{X}^{\prime}_{\bar{c}}u^{*}_{\bar{c}}=O^{*}_{p}\left(T^{1/2}\right)$ together yield that

$$T^{-2}\left(\hat{u}^{*\prime}_{-1}\hat{u}^{*}_{-1}-\bar{u}^{*\prime}_{-1}\bar{u}^{*}_{-1}\right)=o^{*}_{p}(1)\ . \tag{A.23}$$

The proof is therefore completed by showing that $\hat{s}^{*2}_{AR}(0)-\bar{s}^{*2}_{AR}(0)=o^{*}_{p}(1)$. Now,

$$\hat{s}^{*2}_{AR}(0)-\bar{s}^{*2}_{AR}(0)=T^{-1}\left(\Delta\hat{u}^{*\prime}\Delta\hat{u}^{*}-\Delta\bar{u}^{*\prime}\Delta\bar{u}^{*}\right)+T^{-1}\left(\frac{\left(\bar{u}^{*\prime}_{-1}\Delta\bar{u}^{*}\right)^2}{\bar{u}^{*\prime}_{-1}\bar{u}^{*}_{-1}}-\frac{\left(\hat{u}^{*\prime}_{-1}\Delta\hat{u}^{*}\right)^2}{\hat{u}^{*\prime}_{-1}\hat{u}^{*}_{-1}}\right) \tag{A.24}$$

For the term $\hat{u}^{*\prime}_{-1}\Delta\hat{u}^{*}_{-1}-\bar{u}^{*\prime}_{-1}\Delta\bar{u}^{*}_{-1}$, we proceed as in HHLT, proof of Lemma 5(iii), to show

$$T^{-1}\left(\hat{u}^{*\prime}_{-1}\Delta\hat{u}^{*}_{-1}-\bar{u}^{*\prime}_{-1}\Delta\bar{u}^{*}_{-1}\right)=o^{*}_{p}(1) \tag{A.25}$$

follows from (A.19) and from the fact that

$$T^{-3/2}\sum_{t=\lfloor\bar{\tau}T\rfloor+1}^{\lfloor\tau_0T\rfloor}\hat{u}^{*}_{t-1}=o^{*}_{p}(1)\ . \tag{A.26}$$

The result in (A.26) can be proved by noticing that $T^{-3/2}\sum_{t=\lfloor\bar{\tau}T\rfloor+1}^{\lfloor\tau_0T\rfloor}\hat{u}^{*}_{t-1}\leq T^{-1}(\lfloor\tau_0T\rfloor-\lfloor\bar{\tau}T\rfloor)\max_{t=0,\ldots,T}T^{-1/2}|\hat{u}^{*}_{t}|=o^{*}_{p}(1)$, since $\bar{\tau}-\tau_0=o_p(1)$ and $\max_{t=0,\ldots,T}T^{-1/2}|\hat{u}^{*}_{t}|)=O^{*}_{p}(1)$. Hence, (A.25) holds.

For the term $T^{-1}\left(\Delta\hat{u}^{*\prime}\Delta\hat{u}^{*}-\Delta\bar{u}^{*\prime}\Delta\bar{u}^{*}\right)$, we can make use of the decomposition

$$T^{-1}\left(\Delta\hat{u}^{*\prime}\Delta\hat{u}^{*}-\Delta\bar{u}^{*\prime}\Delta\bar{u}^{*}\right)=T^{-1}\sum_{t=1}^{T}\left(\Delta\bar{u}^{*}_{t}-\Delta\hat{u}^{*}_{t}\right)\Delta\bar{u}^{*}_{t}+T^{-1}\sum_{t=1}^{T}\Delta\hat{u}^{*}_{t}\left(\Delta\bar{u}^{*}_{t}-\Delta\hat{u}^{*}_{t}\right). \tag{A.27}$$

For the first term in the right member of (A.27), straightforward calculations (cf. HHLT, proof of Lemma 5(ii)) lead to the equality

$$\begin{aligned}T^{-1}\sum_{t=1}^{T}\left(\Delta\bar{u}^{*}_{t}-\Delta\hat{u}^{*}_{t}\right)\Delta\bar{u}^{*}_{t} \quad &= \quad \left(\hat{\beta}^{*}_{\bar{c}}-\bar{\beta}^{*}_{\bar{c}}\right)T^{-1}\left(\bar{u}^{*}_{T}-\bar{u}^{*}_{p-1}\right)+\left(\hat{\gamma}^{*}_{\bar{c}}-\bar{\gamma}^{*}_{\bar{c}}\right)T^{-1}\left(\bar{u}^{*}_{T}-\bar{u}^{*}_{\lfloor\tau_0T\rfloor}\right)\\ &\quad -\left(\bar{\gamma}_{\bar{c}}-\gamma_0\right)T^{-1}\left(\bar{u}^{*}_{\lfloor\tau_0T\rfloor}-\bar{u}^{*}_{\lfloor\bar{\tau}T\rfloor}\right)\end{aligned}$$

which is obviously of $o^{*}_{p}(1)$, as required. The second term in the right member of (A.27) can be handled similarly, hence implying that $T^{-1}\left(\Delta\hat{u}^{*\prime}\Delta\hat{u}^{*}-\Delta\bar{u}^{*\prime}\Delta\bar{u}^{*}\right)=o^{*}_{p}(1)$. Hence, (A.24) holds. This result, together with (A.22) and (A.23), imply that $\mathcal{MZ}^{tb*}(\bar{\tau},\bar{c})=\mathcal{MZ}^{tb*}(\tau_0,\bar{c})+o^{*}_{p}(1)$, which proves the lemma and completes the proof.

References

Andrews, D.W.K., and Buchinsky, M. (2001). Evaluation of a three-step method for choosing the number of bootstrap repetitions. Journal of Econometrics 103, 345-386.

Bai, J. (1994). Least squares estimation of a shift in a linear process. Journal of Time Series Analysis 15, 453–472.

Banerjee, A., Lumsdaine, R. and 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.

Berk, K.N. (1974). Consistent autoregressive spectral estimates. Annals of Statistics 2, 489–502.

Busetti F. and Taylor, A.M.R. (2003). Testing against stochastic trend in the presence of variance shifts. Journal of Business and Economic Statistics 21, 510–531.

Carrion-i-Silvestre, J.L., Kim, D. and 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.

Cavaliere G. (2004). Unit root tests under time-varying variances. Econometric Reviews, 23, 259–292.

Cavaliere G. and Taylor, A.M.R. (2005). Stationarity tests under time-varying variances. Econometric Theory 21, 1112–1129.

Cavaliere, G. and Taylor, A.M.R. (2007). Testing for unit roots in time series models with non-stationary volatility. Journal of Econometrics 140, 919-947.

Cavaliere G. and Taylor, A.M.R. (2008). Bootstrap unit root tests for time series with non-stationary volatility. Econometric Theory 24, 43-71.

Cavaliere G. and Taylor, A.M.R. (2009). Bootstrap M Unit Root Tests. Econometric Reviews 28, 393-421.

Chang, Y. and Park, Y,J. (2003). A sieve bootstrap for the test of a unit root. Journal of Time Series Analysis 24, 379-400.

Christiano, L.J. (1992). Searching for a Break in GNP. Journal of Business and Economic Statistics 10, 237-50.

Davidson J. (1994). Stochastic limit theory. Oxford: Oxford University Press.

Davidson, R. and Flachaire, E. (2008). The wild bootstrap, tamed at last. Journal of Econometrics 146, 162-169.

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

Giné, E. and Zinn, J. (1990). Bootstrapping general empirical measures. Annals of Probability 18, 851–869.

Gonçalves, S. and Kilian, L. (2004). Bootstrapping autoregressions with conditional heteroskedasticity of unknown form. Journal of Econometrics 123, 89-120.

Gonçalves, S. and Kilian, L. (2007). Asymptotic and bootstrap inference for $AR(\infty)$ processes with conditional heteroskedasticity. Econometric Reviews 26, 609-641.

Haldrup, N., and Jansson, M. 2006. Improving Power and Size in Unit Root Testing. Palgrave Handbooks of Econometrics: Vol. 1 Econometric Theory, Chapter 7. T. C. Mills and K. Patterson (eds.). Palgrave MacMillan, Basingstoke.

Hansen, B.E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413–430.

Harris, D., Harvey, D.I., Leybourne, S.J. and Taylor, A.M.R. (2009). Testing for a Unit Root in the Presence of a Possible Break in Trend. Econometric Theory 25, 1545-1588.

Kim, D. and 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.

Kim T.H., Leybourne, S. and Newbold, P. (2002). Unit root tests with a break in innovation variance. Journal of Econometrics 109, 365–387.

Liu, R.Y. (1988). Bootstrap procedures under some non i.i.d. models. Annals of Statistics 16, 1696–1708.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. Annals of Statistics 21, 255–285.

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

Nunes, L.C., Kuan, C.-M. and Newbold, P. (1995). Spurious Break. Econometric Theory 11, 736-749.

Paparoditis, E.P. and Politis, D.N. (2003). Residual-based block bootstrap for unit root testing. Econometrica 71, 813-855.

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. and Ng, S. (1996). Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. Review of Economic Studies 63, 435–463.

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

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

Sensier, M. and van Dijk, D. (2004). Testing for volatility changes in U.S. macroeconomic time series. Review of Economics and Statistics 86, 833-839.

Stock, J.H. (1999), A class of tests for integration and cointegration. Engle, R.F. and White, H. (eds.), Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W.J. Granger, Oxford: Oxford University Press, 137–167.

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

Stock, J. and Watson, M.W. (2005). Implications of dynamic factor analysis for VAR models, NBER Working Paper #11467.

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

Vogelsang, T.J. and Perron, P. (1998). Additional tests for a unit root allowing the possibility of breaks in the trend function. International Economic Review 39, 1073-1100.

Zivot, E. and 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. Asymptotic critical values and QD de-trending parameters

	0.10 level				0.05 level				0.01 level			
		critical values				critical values				critical values		
τ_0	$\bar{c}$	$\mathcal{MZ}_{\alpha}^{tb}(\tau_0,\bar{c})$	$\mathcal{MSB}_{\alpha}^{tb}(\tau_0,\bar{c})$	$ADF\text{-}GLS^{tb}(\tau_0,\bar{c})$, $\mathcal{MZ}_{t}^{tb}(\tau_0,\bar{c})$	$\bar{c}$	$\mathcal{MZ}_{\alpha}^{tb}(\tau_0,\bar{c})$	$\mathcal{MSB}_{\alpha}^{tb}(\tau_0,\bar{c})$	$ADF\text{-}GLS^{tb}(\tau_0,\bar{c})$, $\mathcal{MZ}_{t}^{tb}(\tau_0,\bar{c})$	$\bar{c}$	$\mathcal{MZ}_{\alpha}^{tb}(\tau_0,\bar{c})$	$\mathcal{MSB}_{tb}^{\alpha}(\tau_0,\bar{c})$	$ADF\text{-}GLS^{tb}(\tau_0,\bar{c})$, $\mathcal{MZ}_{t}^{tb}(\tau_0,\bar{c})$
0.05	11.8	−17.33	0.168	−2.91	15.8	−21.27	0.151	−3.23	24.2	−29.60	0.129	−3.82
0.10	12.8	−18.66	0.162	−3.02	16.8	−22.54	0.147	−3.32	25.4	−30.79	0.126	−3.89
0.15	13.4	−19.35	0.159	−3.08	17.6	−23.06	0.146	−3.37	26.2	−31.46	0.125	−3.94
0.20	13.8	−19.83	0.157	−3.12	17.8	−23.49	0.144	−3.40	26.6	−31.77	0.125	−3.96
0.25	14.0	−20.09	0.156	−3.14	18.2	−23.83	0.143	−3.42	26.6	−32.00	0.124	−3.97
0.30	14.2	−20.23	0.156	−3.15	18.4	−23.95	0.143	−3.43	26.8	−32.22	0.124	−3.99
0.35	14.4	−20.21	0.156	−3.15	18.6	−23.98	0.143	−3.43	27.0	−32.17	0.124	−3.99
0.40	14.4	−20.17	0.156	−3.15	18.4	−23.95	0.143	−3.44	27.0	−32.35	0.123	−3.99
0.45	14.4	−20.11	0.156	−3.15	18.4	−23.90	0.143	−3.44	26.6	−32.14	0.124	−3.99
0.50	14.2	−19.96	0.157	−3.14	18.2	−23.77	0.144	−3.42	26.8	−31.87	0.124	−3.97
0.55	14.0	−19.78	0.158	−3.13	18.0	−23.52	0.145	−3.41	26.6	−31.75	0.125	−3.96
0.60	13.8	−19.47	0.159	−3.10	17.6	−23.25	0.146	−3.39	26.0	−31.49	0.125	−3.95
0.65	13.4	−19.08	0.161	−3.07	17.4	−22.90	0.147	−3.37	25.8	−31.00	0.126	−3.92
0.70	13.2	−18.56	0.163	−3.03	17.0	−22.46	0.148	−3.34	25.4	−30.49	0.127	−3.88
0.75	12.6	−18.03	0.166	−2.99	16.6	−21.83	0.151	−3.29	25.0	−29.83	0.129	−3.85
0.80	12.2	−17.45	0.169	−2.95	16.0	−21.16	0.153	−3.24	24.4	−29.11	0.131	−3.80
0.85	11.6	−16.63	0.173	−2.88	15.2	−20.24	0.157	−3.17	23.6	−28.28	0.133	−3.75
0.90	11.2	−15.73	0.178	−2.80	14.6	−19.30	0.161	−3.10	22.6	−27.05	0.136	−3.67
0.95	10.6	−14.66	0.184	−2.70	14.0	−18.03	0.166	−2.99	21.6	−25.66	0.139	−3.58

Table 2. Finite sample size; ARMA shocks; $T = 150$; shift in volatility.

DGP	ϕ	θ	$\mathcal{MZ}_\alpha(\bar{\tau})$	$\mathcal{MSB}_\alpha(\bar{\tau})$	$\mathcal{MZ}_t(\bar{\tau})$	$t(\bar{\tau})$	$\mathcal{MZ}_\alpha(\bar{\tau})^*$	$\mathcal{MSB}_\alpha(\bar{\tau})^*$	$\mathcal{MZ}_t(\bar{\tau})^*$	$t(\bar{\tau})^*$
(i)	0.0	0.0	0.064	0.064	0.063	0.090	0.079	0.078	0.079	0.067
	0.5	0.0	0.119	0.121	0.116	0.088	0.130	0.130	0.130	0.062
	−0.5	0.0	0.046	0.045	0.047	0.080	0.051	0.050	0.052	0.056
	0.0	0.5	0.099	0.098	0.101	0.122	0.110	0.108	0.111	0.092
	0.0	−0.5	0.093	0.094	0.090	0.060	0.108	0.109	0.107	0.040
(ii)	0.0	0.0	0.166	0.180	0.156	0.129	0.064	0.065	0.061	0.025
	0.5	0.0	0.241	0.264	0.225	0.112	0.112	0.114	0.109	0.017
	−0.5	0.0	0.125	0.134	0.121	0.114	0.041	0.041	0.040	0.018
	0.0	0.5	0.160	0.168	0.154	0.141	0.071	0.069	0.071	0.043
	0.0	−0.5	0.215	0.235	0.204	0.103	0.095	0.098	0.093	0.013
(iii)	0.0	0.0	0.026	0.026	0.026	0.050	0.041	0.040	0.041	0.038
	0.5	0.0	0.058	0.060	0.057	0.046	0.075	0.076	0.074	0.031
	−0.5	0.0	0.015	0.015	0.017	0.045	0.024	0.023	0.025	0.030
	0.0	0.5	0.053	0.053	0.053	0.078	0.067	0.066	0.068	0.060
	0.0	−0.5	0.047	0.048	0.046	0.030	0.062	0.063	0.061	0.021
(iv)	0.0	0.0	0.083	0.079	0.086	0.102	0.026	0.026	0.026	0.014
	0.5	0.0	0.104	0.099	0.106	0.094	0.042	0.042	0.042	0.011
	−0.5	0.0	0.055	0.052	0.056	0.099	0.016	0.015	0.016	0.012
	0.0	0.5	0.082	0.079	0.085	0.134	0.036	0.035	0.036	0.033
	0.0	−0.5	0.102	0.097	0.105	0.093	0.041	0.041	0.041	0.009
(v)	0.0	0.0	0.024	0.024	0.025	0.051	0.035	0.035	0.036	0.034
	0.5	0.0	0.058	0.060	0.058	0.046	0.069	0.069	0.068	0.030
	−0.5	0.0	0.017	0.017	0.018	0.046	0.024	0.023	0.025	0.029
	0.0	0.5	0.053	0.051	0.053	0.081	0.065	0.063	0.066	0.056
	0.0	−0.5	0.045	0.046	0.045	0.031	0.056	0.056	0.057	0.019
(vi)	0.0	0.0	0.047	0.045	0.050	0.065	0.030	0.030	0.030	0.022
	0.5	0.0	0.073	0.069	0.074	0.062	0.045	0.045	0.045	0.015
	−0.5	0.0	0.031	0.028	0.032	0.059	0.016	0.015	0.016	0.018
	0.0	0.5	0.055	0.052	0.057	0.081	0.038	0.038	0.038	0.034
	0.0	−0.5	0.066	0.063	0.069	0.051	0.041	0.041	0.042	0.012
(vii)	0.0	0.0	0.030	0.030	0.029	0.053	0.040	0.040	0.040	0.036
	0.5	0.0	0.060	0.061	0.058	0.047	0.067	0.067	0.065	0.028
	−0.5	0.0	0.020	0.019	0.019	0.046	0.023	0.023	0.023	0.027
	0.0	0.5	0.056	0.055	0.055	0.080	0.062	0.062	0.062	0.054
	0.0	−0.5	0.047	0.047	0.047	0.034	0.056	0.056	0.056	0.020
(viii)	0.0	0.0	0.146	0.154	0.137	0.150	0.035	0.035	0.036	0.018
	0.5	0.0	0.209	0.223	0.198	0.119	0.071	0.075	0.069	0.007
	−0.5	0.0	0.104	0.109	0.100	0.133	0.023	0.024	0.023	0.012
	0.0	0.5	0.133	0.138	0.129	0.158	0.046	0.048	0.047	0.035
	0.0	−0.5	0.186	0.198	0.178	0.107	0.055	0.058	0.055	0.005

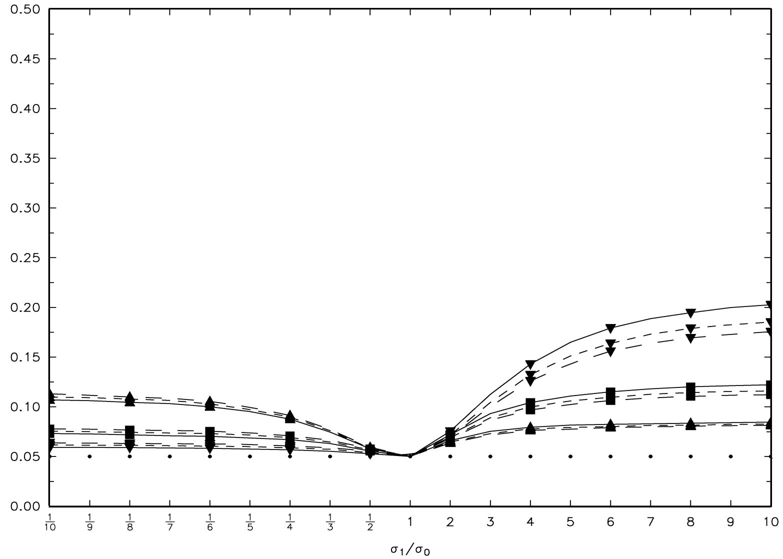
Note: The numbered DGPs denote: (i) No break in trend, $\sigma_1/\sigma_0 = 1$; (ii) No break in trend, $\sigma_1/\sigma_0 = 5$, $\tau_\sigma = 0.7$; (iii) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.3$, $\sigma_1/\sigma_0 = 1$; (iv) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.3$, $\sigma_1/\sigma_0 = 1/5$, $\tau_\sigma = 0.3$; (v) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.5$, $\sigma_1/\sigma_0 = 1$; (vi) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.5$, $\sigma_1/\sigma_0 = 1/5$, $\tau_\sigma = 0.5$; (vii) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.7$, $\sigma_1/\sigma_0 = 1$; (viii) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.7$, $\sigma_1/\sigma_0 = 5$, $\tau_\sigma = 0.7$.

Table 3. Finite sample size; ARMA shocks; $T = 150$; shift in volatility; alternative trimming settings.

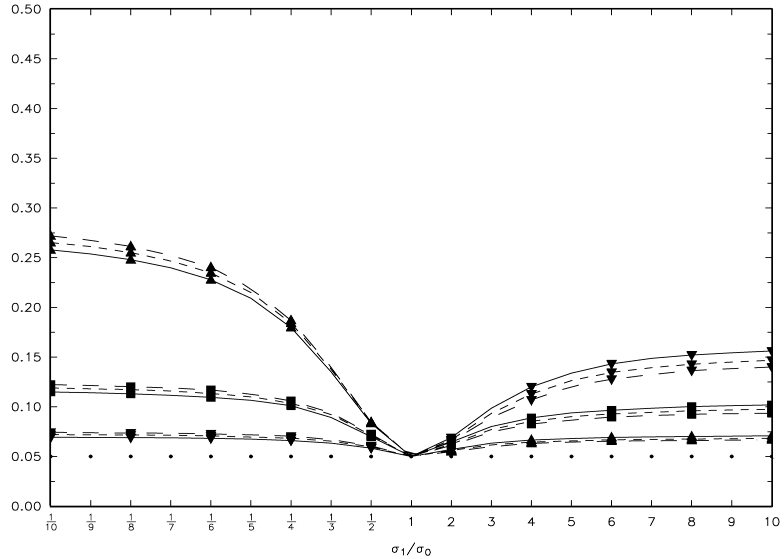
Panel A. $\Lambda = [0.10, 0.90]$										
DGP	ϕ	θ	$\mathcal{MZ}_\alpha(\bar{\tau})$	$\mathcal{MSB}_\alpha(\bar{\tau})$	$\mathcal{MZ}_t(\bar{\tau})$	$t(\bar{\tau})$	$\mathcal{MZ}_\alpha(\bar{\tau})^*$	$\mathcal{MSB}_\alpha(\bar{\tau})^*$	$\mathcal{MZ}_t(\bar{\tau})^*$	$t(\bar{\tau})^*$
(i)	0.0	0.0	0.062	0.062	0.062	0.089	0.078	0.076	0.077	0.067
	0.5	0.0	0.114	0.116	0.112	0.085	0.127	0.126	0.127	0.059
	−0.5	0.0	0.045	0.043	0.045	0.078	0.048	0.048	0.049	0.054
	0.0	0.5	0.096	0.094	0.098	0.117	0.106	0.104	0.107	0.089
	0.0	−0.5	0.089	0.089	0.087	0.059	0.103	0.104	0.103	0.040
(ii)	0.0	0.0	0.155	0.166	0.146	0.116	0.062	0.064	0.060	0.023
	0.5	0.0	0.222	0.241	0.208	0.100	0.110	0.113	0.109	0.016
	−0.5	0.0	0.116	0.125	0.112	0.101	0.042	0.041	0.041	0.017
	0.0	0.5	0.145	0.152	0.139	0.125	0.066	0.065	0.067	0.040
	0.0	−0.5	0.203	0.218	0.193	0.094	0.096	0.098	0.094	0.012

Panel B. $\Lambda = [0.05, 0.95]$										
DGP	ϕ	θ	$\mathcal{MZ}_\alpha(\bar{\tau})$	$\mathcal{MSB}_\alpha(\bar{\tau})$	$\mathcal{MZ}_t(\bar{\tau})$	$t(\bar{\tau})$	$\mathcal{MZ}_\alpha(\bar{\tau})^*$	$\mathcal{MSB}_\alpha(\bar{\tau})^*$	$\mathcal{MZ}_t(\bar{\tau})^*$	$t(\bar{\tau})^*$
(i)	0.0	0.0	0.060	0.059	0.060	0.084	0.073	0.072	0.074	0.065
	0.5	0.0	0.107	0.107	0.106	0.081	0.121	0.119	0.120	0.056
	−0.5	0.0	0.042	0.040	0.043	0.072	0.047	0.045	0.048	0.052
	0.0	0.5	0.082	0.081	0.084	0.103	0.093	0.092	0.094	0.080
	0.0	−0.5	0.084	0.084	0.084	0.056	0.098	0.099	0.098	0.037
(ii)	0.0	0.0	0.134	0.144	0.127	0.099	0.058	0.058	0.056	0.020
	0.5	0.0	0.194	0.209	0.181	0.086	0.100	0.100	0.098	0.014
	−0.5	0.0	0.099	0.106	0.096	0.086	0.039	0.037	0.038	0.014
	0.0	0.5	0.118	0.124	0.114	0.102	0.057	0.056	0.057	0.033
	0.0	−0.5	0.177	0.189	0.169	0.080	0.086	0.088	0.085	0.012

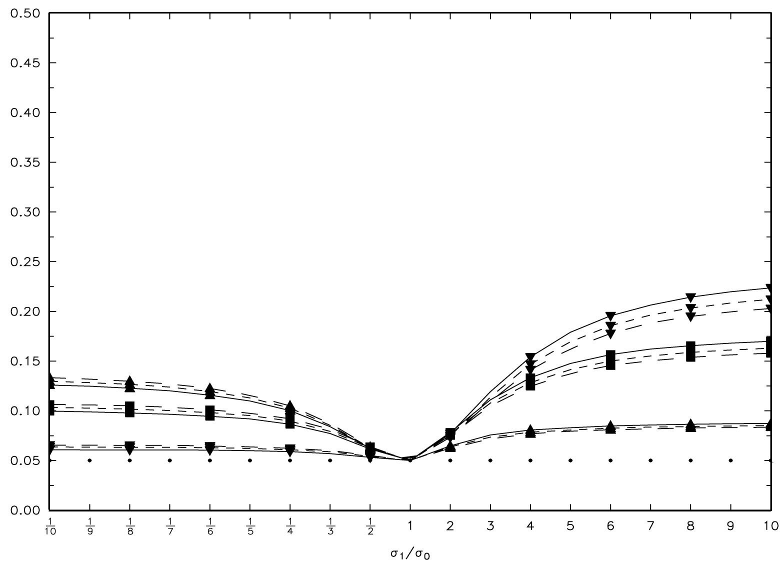
Note: The numbered DGPs denote: (i) No break in trend, $\sigma_1/\sigma_0 = 1$; (ii) No break in trend, $\sigma_1/\sigma_0 = 5$, $\tau_\sigma = 0.7$.



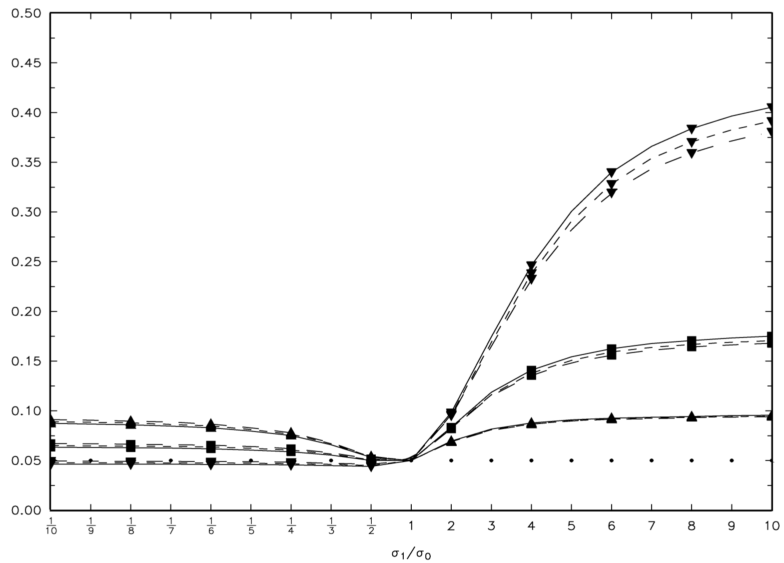
(a) No break in trend



(b) Break in trend, $\tau_0 = 0.3$

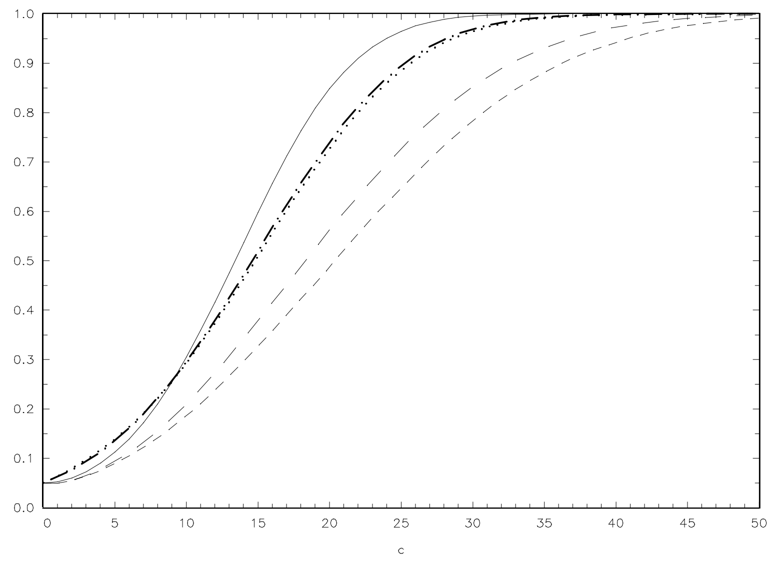


(c) Break in trend, $\tau_0 = 0.5$

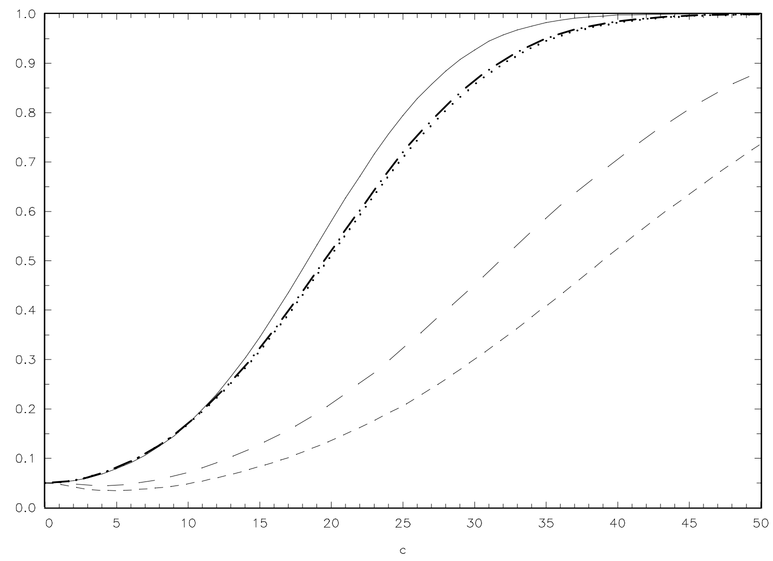


(d) Break in trend, $\tau_0 = 0.7$

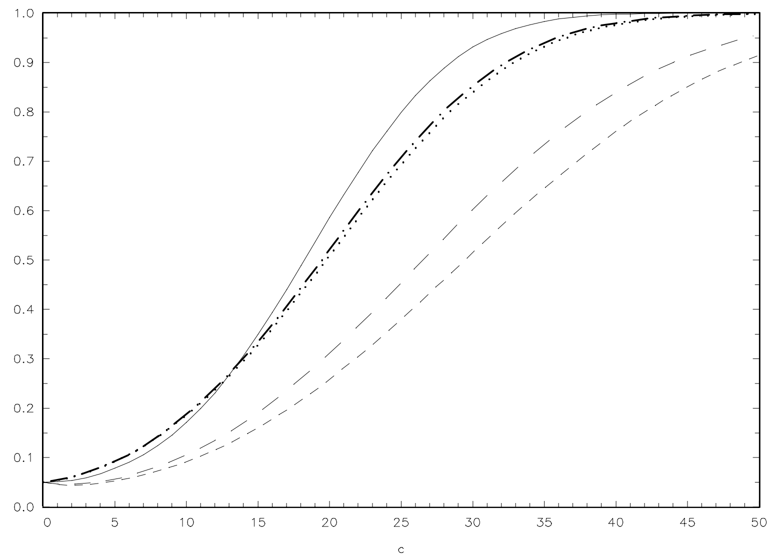
Figure 1. Asymptotic size; shift in volatility. $\mathcal{MSB}(\bar{\tau})$: —— ; $\mathcal{MZ}_\alpha(\bar{\tau})$: - - - ; $\mathcal{MZ}_t(\bar{\tau}), t(\bar{\tau})$: – – ; $\mathcal{MSB}(\bar{\tau})^*, \mathcal{MZ}_\alpha(\bar{\tau})^*, \mathcal{MZ}_t(\bar{\tau})^*, t(\bar{\tau})^*$: $\cdots$. Lines marked with triangles, squares and inverted triangles denote $\tau_\sigma = 0.3$, $\tau_\sigma = 0.5$ and $\tau_\sigma = 0.7$, respectively.



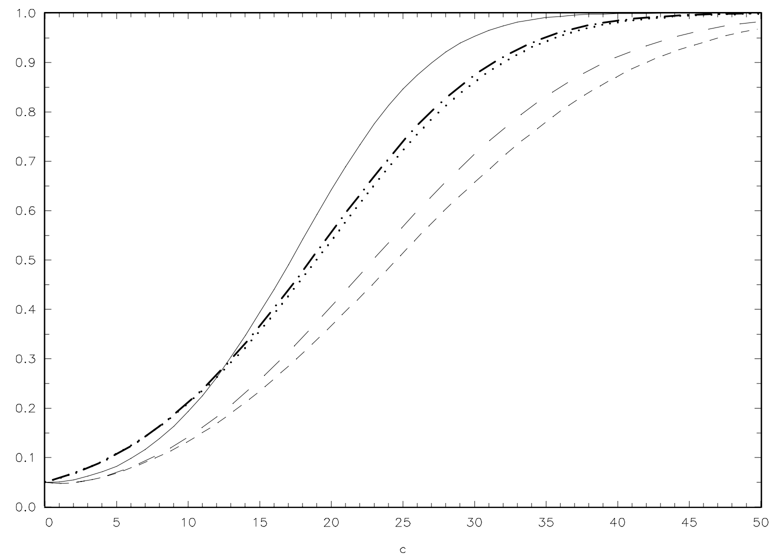
(a) No break in trend



(b) Break in trend, $\tau_0 = 0.3$

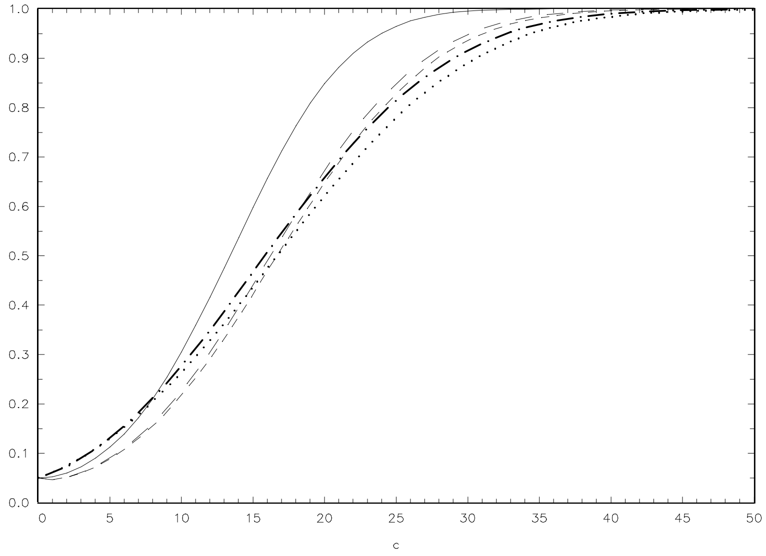


(c) Break in trend, $\tau_0 = 0.5$

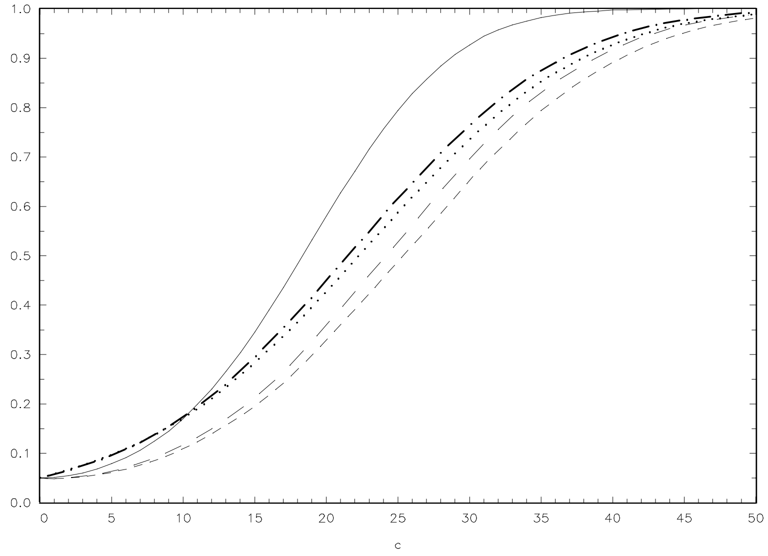


(d) Break in trend, $\tau_0 = 0.7$

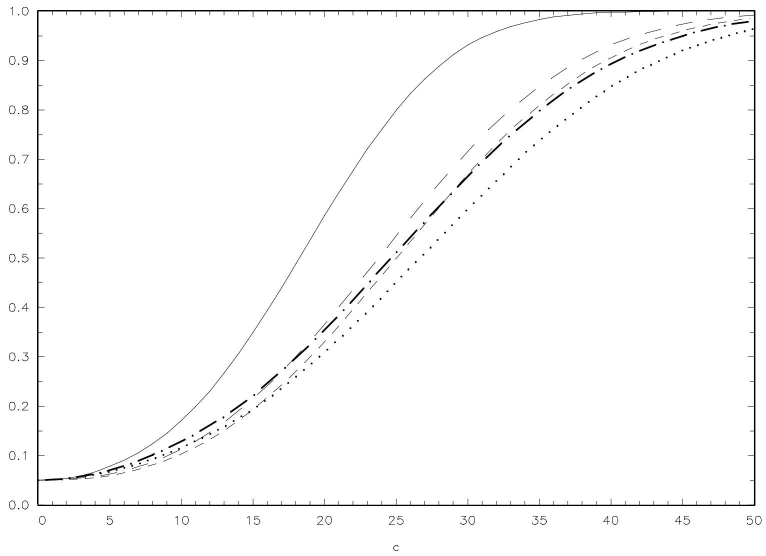
Figure 2. Asymptotic size-adjusted local power of $\mathcal{MZ}_t(\bar{\tau}), t(\bar{\tau})$; shift in volatility, $\tau_\sigma = 0.3$.
$\sigma_1/\sigma_0 = 1$: —— ; $\sigma_1/\sigma_0 = 1/10$: - - - ; $\sigma_1/\sigma_0 = 1/5$: — — ; $\sigma_1/\sigma_0 = 5$: — · — ; $\sigma_1/\sigma_0 = 10$: $\cdots$.



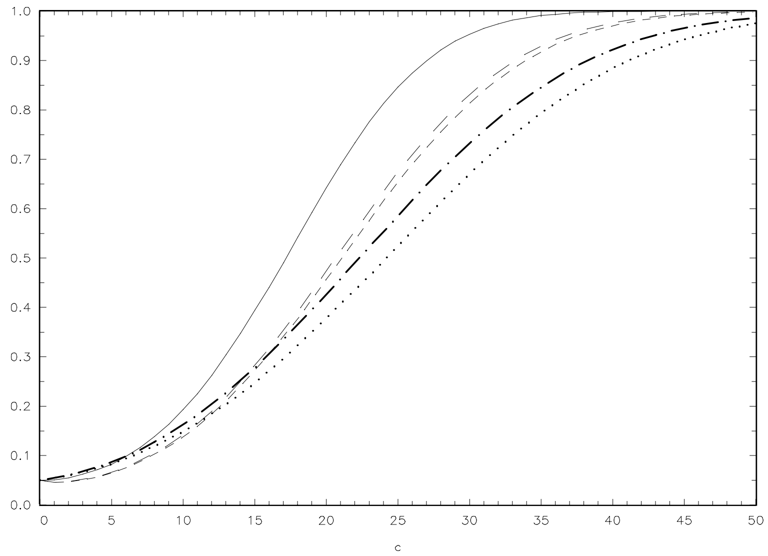
(a) No break in trend



(b) Break in trend, $\tau_0 = 0.3$

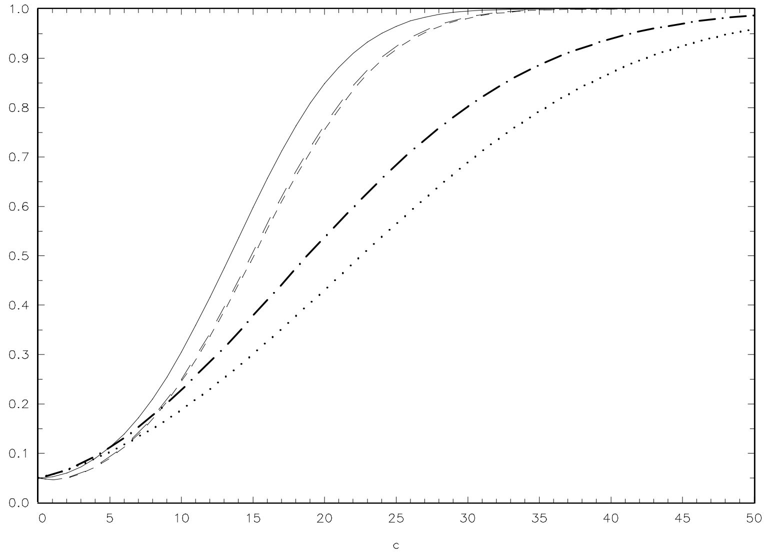


(c) Break in trend, $\tau_0 = 0.5$

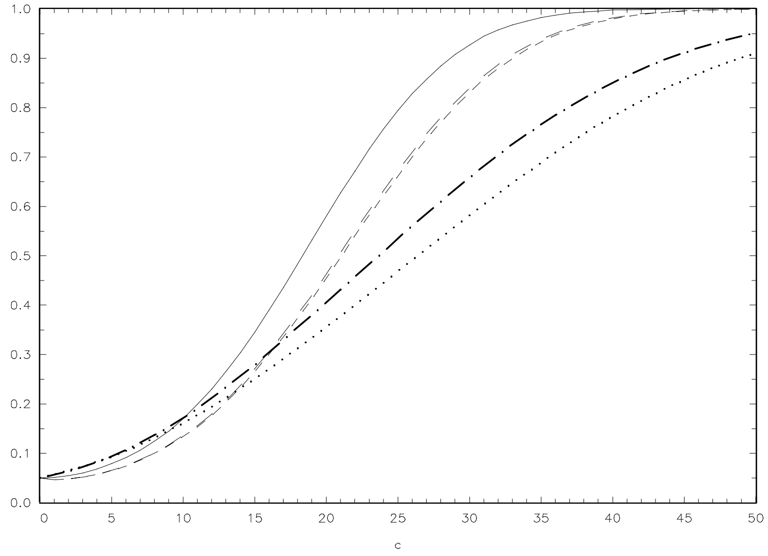


(d) Break in trend, $\tau_0 = 0.7$

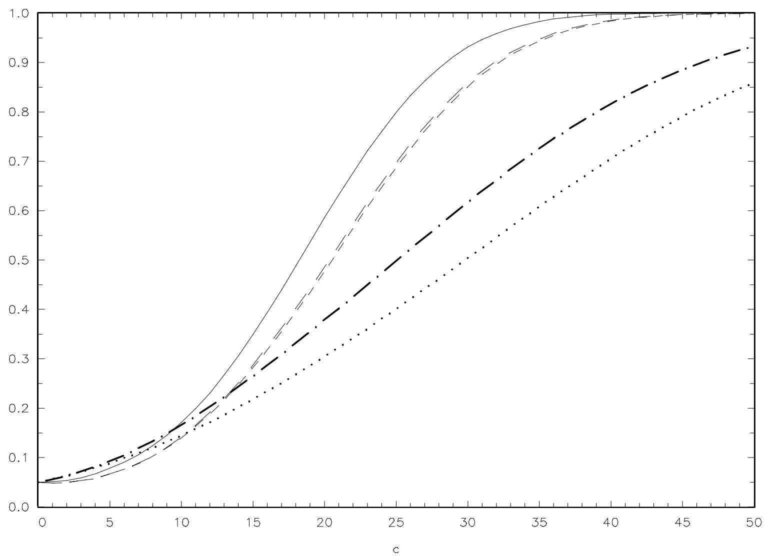
Figure 3. Asymptotic size-adjusted local power of $\mathcal{MZ}_t(\bar{\tau}), t(\bar{\tau})$; shift in volatility, $\tau_\sigma = 0.5$.
$\sigma_1/\sigma_0 = 1$: —— ; $\sigma_1/\sigma_0 = 1/10$: - - - ; $\sigma_1/\sigma_0 = 1/5$: $--$; $\sigma_1/\sigma_0 = 5$: $-\cdot-$; $\sigma_1/\sigma_0 = 10$: $\cdots$.



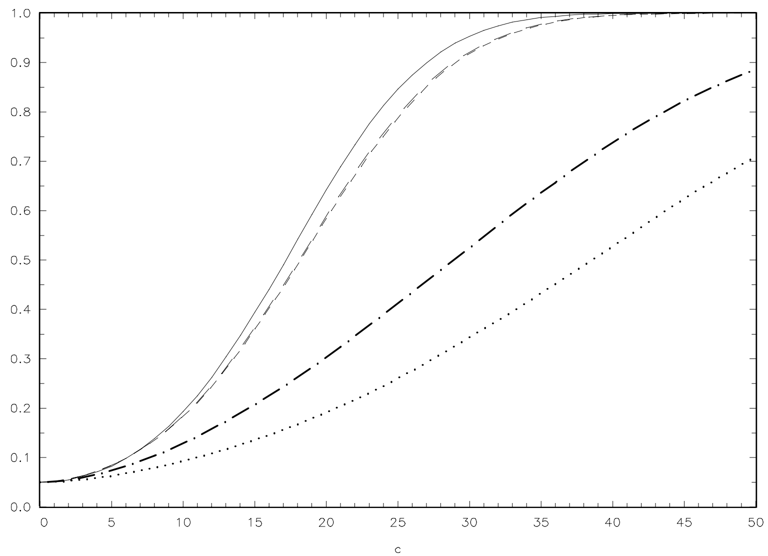
(a) No break in trend



(b) Break in trend, $\tau_0 = 0.3$

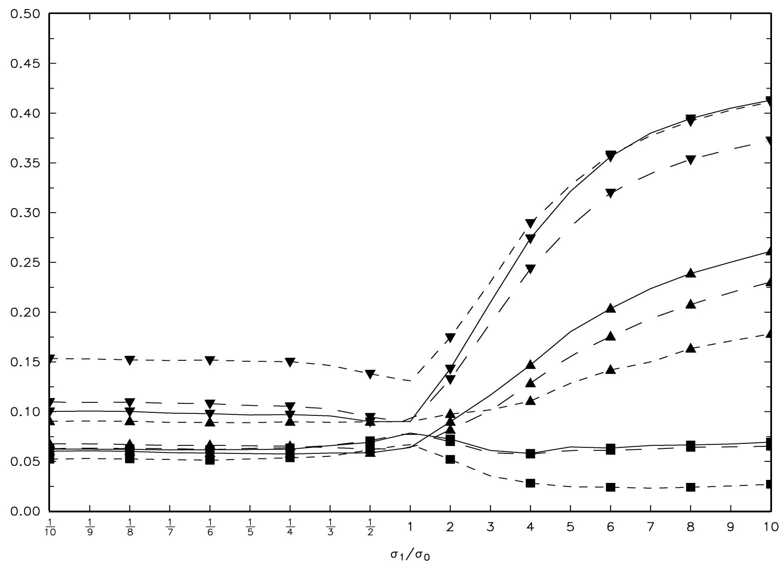


(c) Break in trend, $\tau_0 = 0.5$

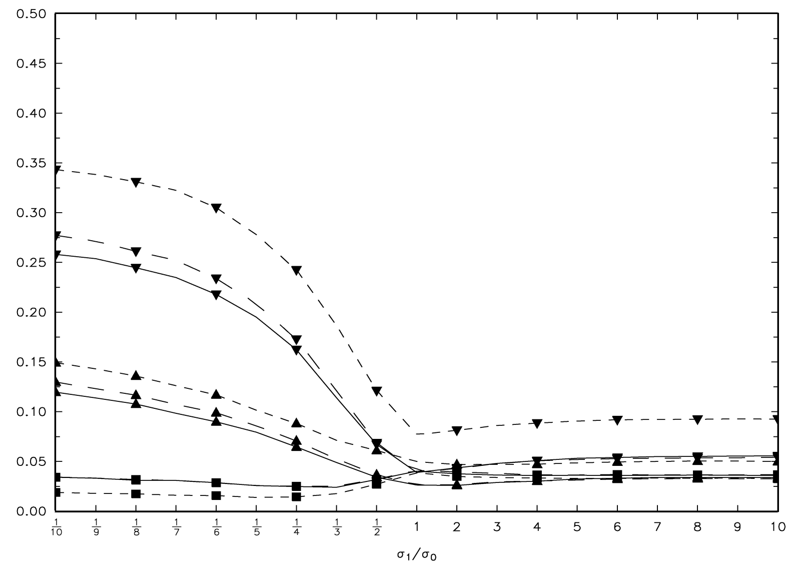


(d) Break in trend, $\tau_0 = 0.7$

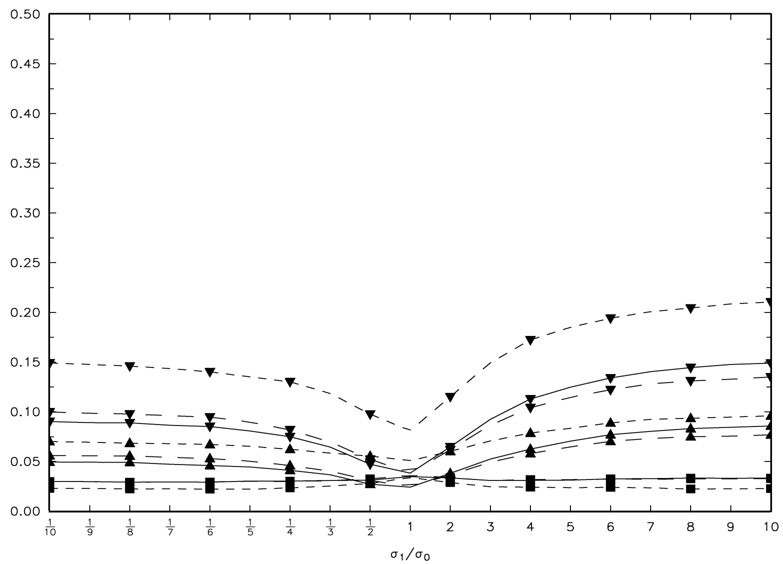
Figure 4. Asymptotic size-adjusted local power of $\mathcal{MZ}_t(\bar{\tau}), t(\bar{\tau})$; shift in volatility, $\tau_\sigma = 0.7$.
$\sigma_1/\sigma_0 = 1$: —— ; $\sigma_1/\sigma_0 = 1/10$: - - - ; $\sigma_1/\sigma_0 = 1/5$: − − ; $\sigma_1/\sigma_0 = 5$: − · − ; $\sigma_1/\sigma_0 = 10$: $\cdots$.



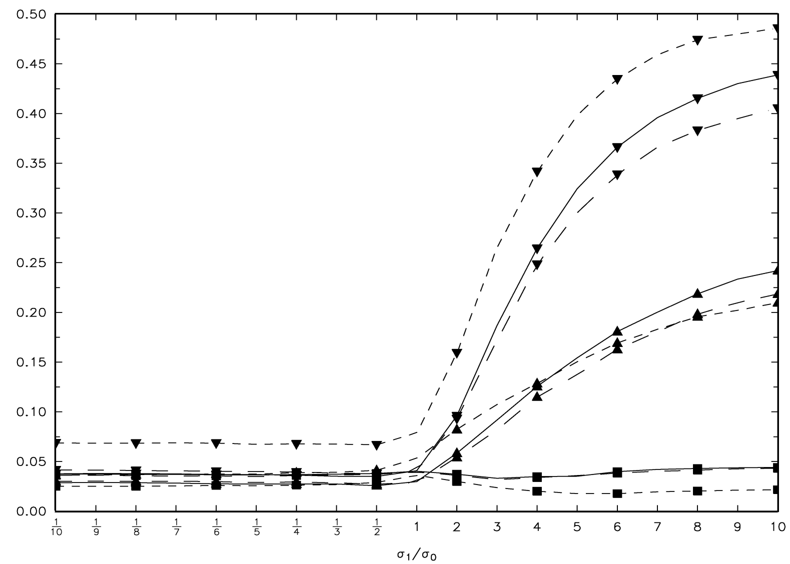
(a) No break in trend, $\tau_\sigma = 0.7$



(b) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.3$, $\tau_\sigma = 0.3$

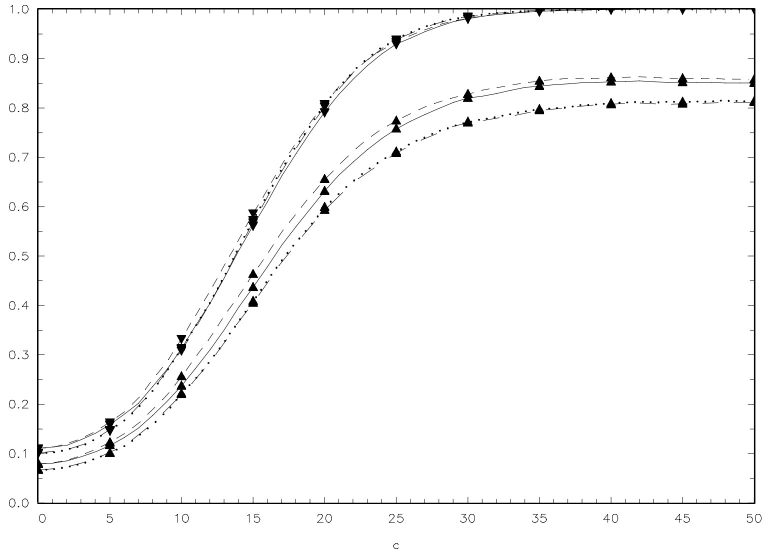


(c) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.5$, $\tau_\sigma = 0.5$

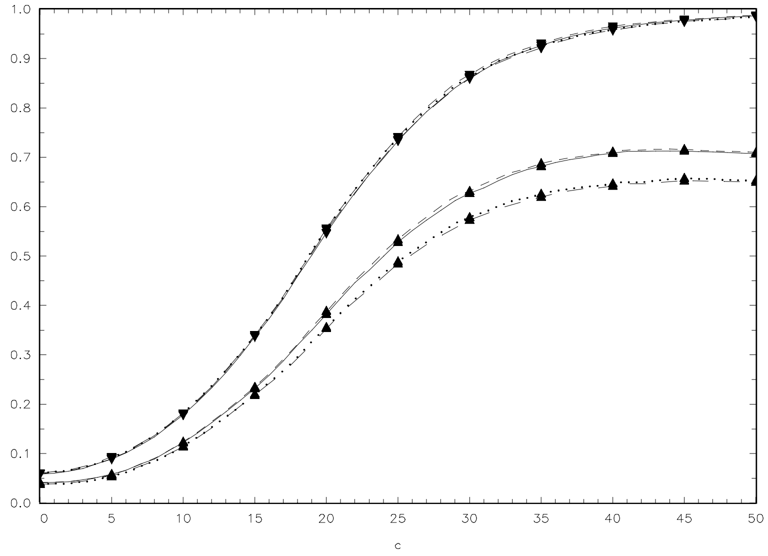


(d) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.7$, $\tau_\sigma = 0.7$

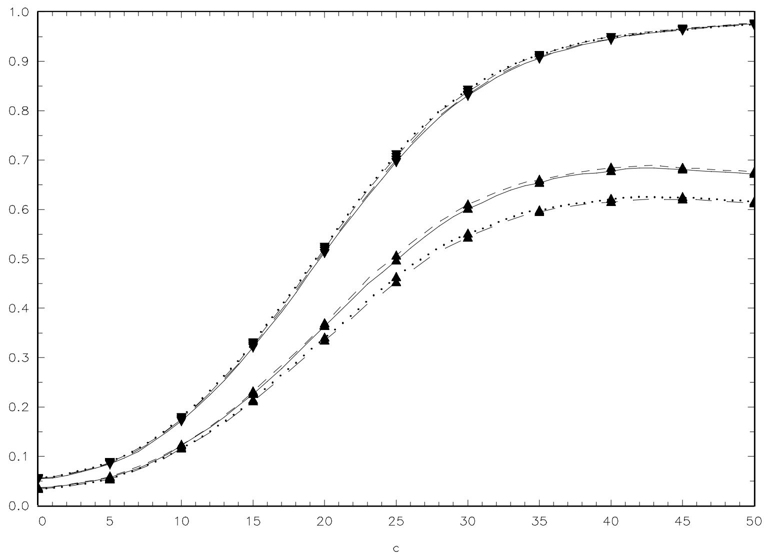
Figure 5. Finite sample size; $T = 150$; shift in volatility. $\mathcal{MSB}$: —— ; $\mathcal{MZ}_t$: – – ; t: - - - - .
Lines marked with (i) inverted triangles, (ii) triangles and (iii) squares denote, respectively,
(i) $\mathcal{M}(\bar{\tau})$ and $t(\bar{\tau})$ with $p = 0$, (ii) $\mathcal{M}(\bar{\tau})$ and $t(\bar{\tau})$ with $p = p_{\text{MAIC}}$, and (iii) $\mathcal{M}(\bar{\tau})^*$ and $t(\bar{\tau})^*$ with $p = p_{\text{MAIC}}$.



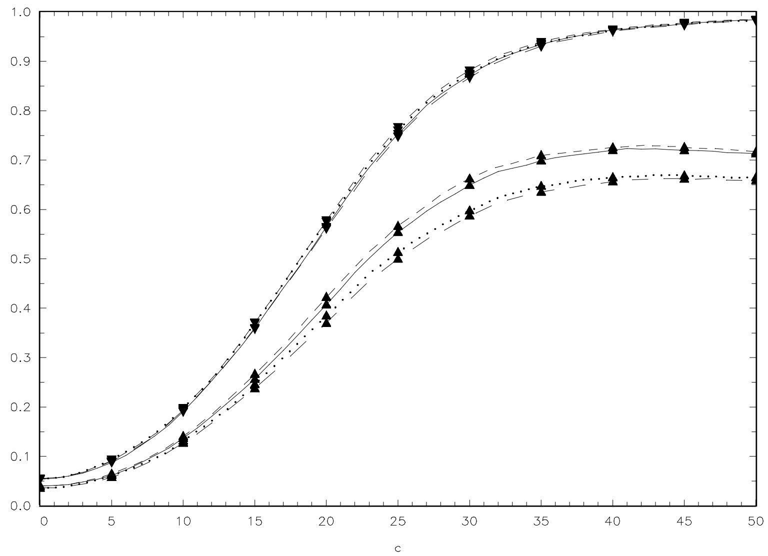
(a) No break in trend



(b) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.3$

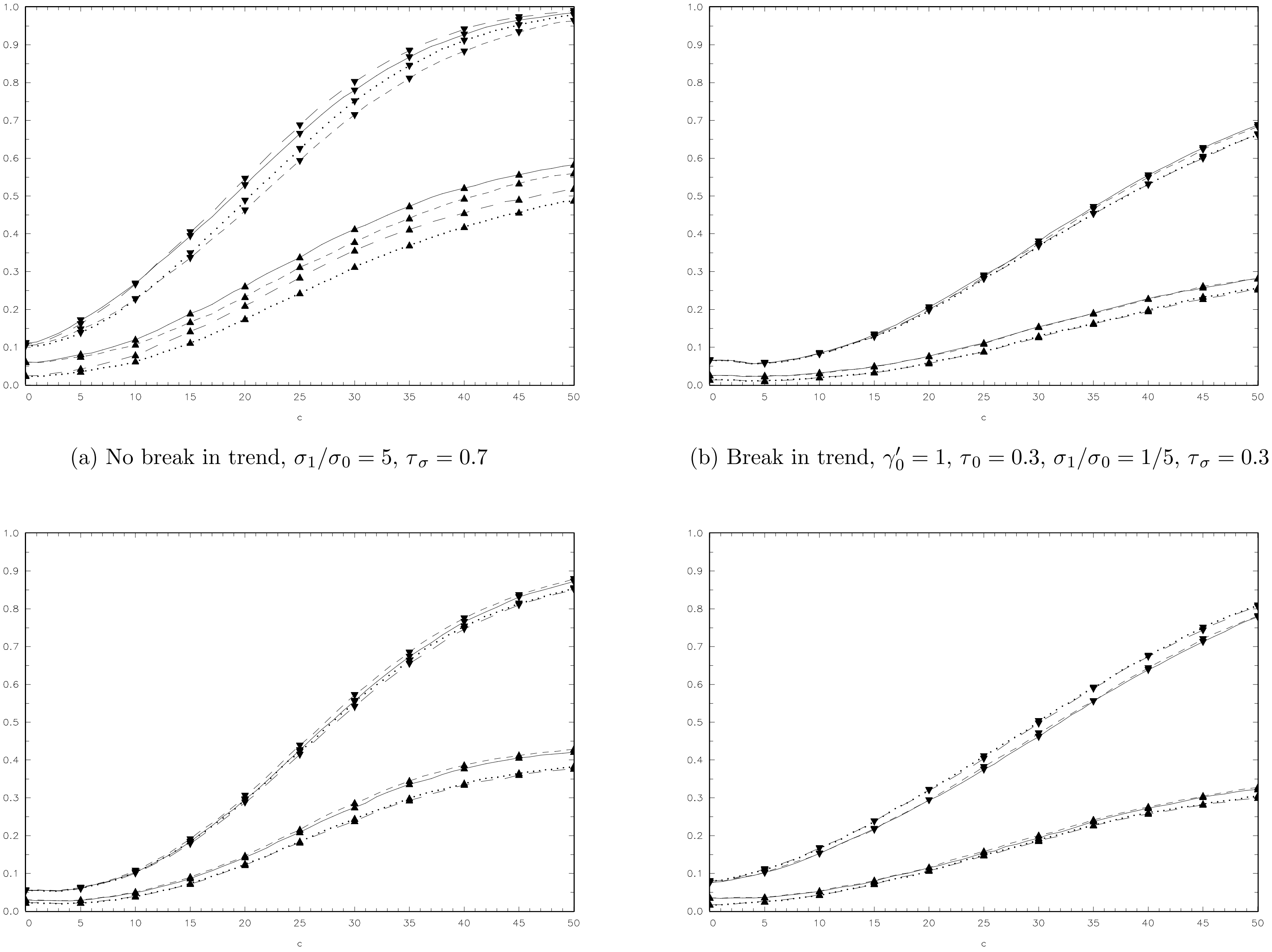


(c) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.5$



(d) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.7$

Figure 6. Finite sample power comparisons; $T = 150$; constant volatility. $\mathcal{MZ}_t(\bar{\tau})$: - - - ; $t(\bar{\tau})$: $\cdots$; $\mathcal{MZ}_t(\bar{\tau})^*$: —— ; $t(\bar{\tau})^*$: $- -$. Lines marked with inverted triangles and triangles denote, respectively, tests conducted with $p = 0$ and $p = p_{\text{MAIC}}$.



(a) No break in trend, $\sigma_1/\sigma_0 = 5$, $\tau_\sigma = 0.7$

(b) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.3$, $\sigma_1/\sigma_0 = 1/5$, $\tau_\sigma = 0.3$

(c) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.5$, $\sigma_1/\sigma_0 = 1/5$, $\tau_\sigma = 0.5$

(d) Break in trend, $\gamma_0' = 1$, $\tau_0 = 0.7$, $\sigma_1/\sigma_0 = 5$, $\tau_\sigma = 0.7$

Figure 7. Finite sample power comparisons; $T = 150$; shift in volatility. $\mathcal{MZ}_t(\bar{\tau})$: - - - ; $t(\bar{\tau})$: $\cdots$; $\mathcal{MZ}_t(\bar{\tau})^*$: —— ; $t(\bar{\tau})^*$: $- -$. Lines marked with inverted triangles and triangles denote, respectively, tests conducted with $p = 0$ and $p = p_{\text{MAIC}}$.